Journal of Applied Biology & Biotechnology Vol. 1 (04), pp. 016-023, Nov-Dec, 2013
Available online at http://www.jabonline.in


# Identification, evaluation and optimization of a minimum simple sequence repeat marker set for triticale breeding

W.C. Botes and D. Bitalo

Department of Genetics, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa.

**ARTICLE INFO**

*Article history:*
Received on: 26/11/2013
Revised on: 08/12/2013
Accepted on: 14/12/2013
Available online: 31/12/2013

***Key words:***
Wheat, rye, express sequence tags, Polymorphism information content, genetic diversity.

**ABSTRACT**

Hexaploid triticale (x *Triticosecale* Wittmack) is a man-made hybrid that consists of 3 genomes, A, B and R, the A- and B-genome from wheat and the R-genome from rye. Simple sequence repeat (SSR) markers are transferable between closely related species. The aim of this study was to exploit this transferability in order to identify, evaluate and optimize a minimum SSR marker set for cultivar identification, marker assisted backcrosses, and germplasm assessment in triticale. A set of 29 express sequence tags (EST) and genomic SSR primer pairs developed in rye were used to test the genetic variability among eight triticale cultivars and two rye cultivars. Only 23 of the primer sets produced amplicons and expressed a 79.3% transferability. Fifty-ESTs and genomic SSR markers were used to analyse the wheat (A and B) genomes, and achieved an84% transferability. Following the identification and evaluation of the markers two EST R-genome SSRs were selected, and five wheat SSRs (two EST and three genomic) based on discriminatory ability, which corresponded with available pedigree information.

## 1. INTRODUCTION

The distribution and organisation of microsatellites in a number of plant genomes have been studied using fluorescence *in situ* hybridization (FISH) and in-gel hybridization techniques [1]. Using FISH Röder *et al.* [2] showed that there was a regular occurrence of $(GAA)_n$ repeats in barley and other *Triticeae* species and Gupta *et al.*[1] using in-gel hybridization showed that there were numerous microsatellites in long stretches of repeated DNA of bread wheat. However, later studies showed that FISH and in-gel hybridization techniques only detected the extended regions of the repetitive DNA, and did not focus on the regions of the DNA uniquely prone to containing microsatellite motifs [1]. However, on the basis of genetic and physical mapping techniques, short regions of repetitive DNA were amplified using PCR amplification. Hexaploid triticale (x *Triticosecale* spp. Wittmack ex A. Camus 1927) is a man-made hybrid that consists of 3 genomes, A, B and R, the A- and B-genome from wheat and the R-genome from rye. Genetic maps of wheat constructed using SSRs showed that the microsatellites were not clustered in particular regions, but were evenly distributed in the different regions of the wheat chromosomes [3]. The average distances between two SSRs for wheat were recorded as 704 kb between two $(AC)_n$ repeat motifs and 440 kb between two $(GA)_n$ repeat motifs [4]. Rye has been noted to be a vital genetic source for wheat improvement programs due to its agronomical traits; the resistance genes *Pm8* to powdery mildew (*Erysiphe graminis* f. sp. *Tritici*), *Lr26* to leaf rust (*Puccinia recondite* f. sp. *Tritici*), *Sr31* to stem rust (*Pucciniagraminis* f. sp. *Tritici*) and *Yr9* to yellow rust (*Puccinia stiiformis* f. sp. *Tritici*) which map on chromosome 1RS [5], contributed to the success of wheat cultivars bearing the 1BL.1RS translocation. These and other traits of rye that cannot be analysed in other related crop species, necessitate a more detailed analysis of the rye genome using microsatellite markers in order to recognise economically important regions with molecular markers [6]. These maps are also necessary in triticale breeding programs as rye is a constituent of the synthesised hybrid [7]. The pioneer linkage maps of rye were based on RFLP markers [8, 9], and later improved by combining morphological genes and selected anchor probes of rye and other cereals like wheat and barley [10] to saturate the rye linkage maps. Due to the high level of reproducibility, robustness and ease of high throughput analysis of SSRs [6], they have become the favoured marker choice for mapping genes in rye. However rye SSR databases are not as well covered as those of other cereal genomes like wheat [3]. Efforts to saturate rye linkage maps with SSR markers includes studies done by Hackauf and Wehling [6] that used public sequencing resources to access rye cDNA sequences from anthers, cold-stressed leaves and aluminium-stressed roots and used these as a resource for SSR marker development. One other effort was the utilisation of mapped wheat and barley microsatellites that may cross-amplify in rye. Khlestkina *et al.* [7] searched 8,930 EST sequences deposited in Genbank <http://www.ncbi.nlm.nih.gov> for rye microsatellite sequences with di-, tri- and tetra-nucleotide repeats and also utilised wheat microsatellite markers (WMS) of the Gatersleben

* *Corresponding Author*
*Department of Genetics, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa.*

collection [3] to saturate the rye genomic map. Such efforts opened the gateway to screening the R genome in triticale using SSR markers. The wheat genomic and EST databases have accumulated numerous sequences of SSR markers that were previously shown to be successful in assessing the genetic diversity in wheat and were also reported to exhibit transferability to triticale. Triticale shares the A and B genomes with wheat and the wheat genomic and EST-SSRs usually detect a single locus in each of those genomes [11]. These databases are publicly available via domains like the Wheat Microsatellite Consortium (WMC) primer sequences published online at <http://wheat.pw.usda.gov/ggpages/SSR/WMC/>, and the BARC markers that were developed for the United States Wheat and Barley Scab Initiative to map and characterise genes for fusarium resistance [12, 13]. Other database sources have a large deposit of ESTs developed for wheat, and the EST-SSRs can be developed from the *in silico* analysis of these databases and can improve gene-based maps and marker assisted selection (MAS) studies [14]. In wheat, *insilico* analysis of the EST-SSR bread wheat database <http://www.ncbi.nlm.nih.gov/dbEST> showed a presence of one EST-SSR every 6.2 kb [15]. Previous studies have successfully implemented SSRs developed in both wheat and rye to assess the genetic diversity among triticale.Tams *et al.* [16, 17] assessed the genetic diversity of winter triticale using SSRs from wheat and rye and concluded that these SSRs provide a powerful tool in the study of genetic diversity among triticale lines. Kuleung *et al.*[18] used 43 gSSRs developed in wheat, and 14 gSSRs and EST-SSRs developed in rye, to assess the genetic diversity and relationship of 80 hexaploid triticale accessions pooled from global breeding programs. Cluster analysis grouped the accessions into fivegroupsthat corresponded with the pedigree information, country of origin, growth habits and release year, and this was attributed to the exchange of material between breeding programs. At theSU-PBL, results of work done using wheat SSRs was used to analyse 40 wheat genotypes sourced from five breeding programs, including SU-PBL genotypes. Data generated by these markers managed to group the genotypes into seven clusters, and managed to accurately identify different cultivars [19]. This study aims to extend the use of these SSR markers and others, developed, in the wheat and rye genomes, to evaluate the genetic diversity among triticale cultivars, to identify polymorphic SSRs, and to optimise a minimum SSR marker set that would be cost effective, and amenable for routine usage in cultivar identification, marker assisted backcrosses, and germplasm assessment in the SU-PBL breeding program.

## 2. MATERIALS AND METHODS

### 2.1. Plant material

Spring triticale cultivars ('US2007', 'US2008', 'US2009', 'US2010', 'Bacchus', 'Tobie', 'Ibis' and 'Rex'), two rye cultivars ('Duiker' and 'Henoch') sourced from the SU-PBL, were evaluated in this study. A commercial wheat cultivar 'SST88' (Sensako, South Africa), was also evaluated.

### 2.2. Genomic DNA extraction

DNA was extracted following the CTAB based protocol as described by Honing [19]. All extracted DNA was then quantified and its quality checked using the NanoDrop® ND-1000 Spectrophotometer following the Thermo SCIENTIFIC user's manual guide. Each DNA sample was then diluted to 100 ng/ µl and stored at 4$^{o}$C while all stock DNA was stored at -20$^{o}$C until needed.

### 2.3. SSR marker evaluation

Fifty-seven SSR markers developed in wheat were screened for polymorphisms in this study. Röder *et al.* [3, 5] developed SSR markers that were genomic specific and mapped to the A, B and D genomes of hexaploid wheat. Clones containing microsatellite repeats GA and GT were purified from genomic phage library fragments and primers were designed for these sequenced clones and designated Xgwm for "Wheat Microsatellite." Ma *et al.*[20] screened two wheat (*Triticum aestivum* L.) genomic libraries for di-, tri-, and tetranucleotide tandem repeats. Clones containing $(AC)_n$, $(AG)_n$, $(TCT)_n$ and $(TTG)_n$ repeats were isolated and sequenced. Primers flanking these repeat sequences were designed using the software MACVECTOR (Kodak, New Haven, Conn.) and were designated Xcnl. Other genomic specific SSRs developed for the wheat genome include:

- Xwmc primers which were supplied to all members of the WMC (Wheat Microsatellite Consortium) for genotyping and characterisation [1, 21, 22];
- Xcfa (Pierre Sourdille microsatellite) primers [23]; and
- Xbarc (Beltsville Agriculture Research Center) and Xcfd (Pierre Sourdille microsatellite) [24, 25].

Zhang *et al.* [14, 26] screened over 46,000 EST contigs to identify microsatellites. The microsatellites were defined as sequences having a minimum of three repeats (with each repeat motif containing 1 to 6 nucleotides) and a total length of no less than 12 nucleotides. Primers spanning these microsatellites were then designed using PRIMER software (version 0.5, Whitehead Institute for Biomedical Research, Cambridge, Mass). Primers with an optimal length (18 to 22 bp), an optimal product size (100 to 400 bp), melting temperatures ranging from 57$^{o}$C and 63$^{o}$C and other factors were selected and these primer sequences subjected to a BLAST analysis to evade redundancy. These primers were designated cfe and each forward primer was M13-tailed with the sequence 5'-CACGACGTTGTAAAACGAC-3'. Rye EST sequences deposited in GenBank (http://www.ncbi.nlm.nih.gov) were screened for microsatellite sequences with di, tri and tetra-nucleotide repeats. Primers flanking these repeats were designed using Primer3.0 Input (version 0.4.0) software [27] and designated rye expressed microsatellite sites (REMS) [7] and *Secale cereale* microsatellite (SCM) [6]. An 18bp sequence (5'-TGT AAA ACG ACG GCC AGT-3') identical to the sequence of an M13 primer, was added to the 5'-end of each forward primer as described by Hackauf and Wehling [6] and [7]. Oligo Analyzer version 1.0.2

was employed to analyse the designed primers for self-annealing. An additional group of SSRs developed by Saal and Wricke [28], and labelled "SCM" were also screened. These primers were designed from the flanking sequences of $(GT/CA)_n$ and $(CT/GA)_n$ repeats in the genomic library using the software OLIGO [29], choosing primers between 18 and 26 bp in length and annealing temperatures between 50°C and 65°C.

### 2.4. PCR protocol

PCR reactions were run with a Thermal Cycler 2720 (Applied Biosystems) with a total reaction volume of 20µl following conditions as described in Röder *et al.* [3]; 1X Bioline $NH_4$ buffer [16mM (NH4)2SO4, 67mM Tris-HCl (pH 8.8) at 25°C, 0.01% (v/v) Tween-20], 0.2mM of each dNTP, 0.5µM of each primer, 1.5mM $MgCl_2$, 1U *Taq* polymerase (Bioline BIOTAC™) and 300ng template DNA. Thermal cycling conditions were (1) 94°C for 3 min (2) 94°C for 1 min (45 cycles) (3) annealing temperature for 1 min (4) 72°C for 2 min (4) 72°C for 10 min. PCR primers designed for the wheat genome were used in the PCR amplification of wheat and triticale cultivars while those designed for the rye genome were used in the PCR amplification of rye and triticale cultivars.

### 2.5. DNA separation and analysis

Amplification products were separated on a 6% (w/v) denaturing polyacrylamide gel (acrylamide: bisacrylamide, 19:1) containing 6M urea and 1 X TBE buffer (Tris-HCl, EDTA, Boric acid). 15µl of each sample (10µl PCR product: 10µl loading buffer [98% formamide, 10mM $Na_2EDTA$ (pH 8), 0.05% (w/v) bromophenolblue, 0.05% (w/v) xylene cyanol FF] were loaded after 5 min of denaturing at 95°C and immediate quenching on ice. Samples along with a 50bp or 100bp ladder were separated by electrophoresis in a 1 X TBE buffer at 60W for 4 hours. Band fragments were revealed by following silver staining procedures described by [30]; (1) 20 min in fixing solution (0.5% (v/v) acetic acid, 10% (v/v) ethanol) (2) two 5 min rinses in distilled water (3) 20 min in 0.1% (w/v) silver nitrate (4) 10 s rinse in distilled water (5) 10-15 min in developing solution (1.5% (w/v) sodium hydroxide, 0.16% (v/v) formaldehyde). Gels were then rinsed in distilled water and digitally recorded using a Nikon digital camera and a light-table.

### 2.6. Data analysis

The resulting band fragments were scored as alleles and their sizes determined using the 50bp or 100bp ladder (Promega, Madison, Wisconsin, USA). Repeats of SSRs were scored as haplotypes with each locus representing the genetic information of each SSR marker. The scored band sizes were used to compute the frequency-based distances between cultivars using the CSChord [31] method, Weir's [32] gene diversity and its alternative polymorphism information content (PIC) were calculated using PowerMarker v3.25 [33] to indicate the ability of each marker to detect polymorphisms in the study population. The scored fragments were also used in a cluster analysis using both the unweighted pair group method with arithmetic average (UPGMA) [34] and Neighbour-joining (NJ) [35] trees to show the relationships among the cultivars.

## 3. RESULTS AND DISCUSSION

### 3.1. A and B genome markers

Of the 57 wheat microsatellite primer pairs tested, 48 markers (35 located on the A genome chromosomes and 13 located on the B genome chromosomes) produced amplicons in the nine accessions (eight spring triticale and one wheat). The 48 markers were chosen after excluding those markers that showed ambiguity and no cross-amplification. Overall 84% of the wheat markers amplified microsatellite loci in triticale, suggesting a good transferability of these wheat markers to triticale.

**Table. 1A:** A statistical summary generated in PowerMarker v.3.25 for the A genome.

| Genome | Marker | M.A.F | Allele No | Gene diversity | PIC |
|---|---|---|---|---|---|
| A-genome | cfe2-6A,6B,6D | 0.11 | 9 | 0.89 | 0.88 |
| | cfe22-4A | 0.67 | 3 | 0.49 | 0.44 |
| | cfe34-4BL,5A,5B | 0.78 | 3 | 0.37 | 0.34 |
| | cfe37-5A,5B | 0.78 | 3 | 0.37 | 0.34 |
| | cfe53-2AL | 0.22 | 8 | 0.86 | 0.85 |
| | cfe110-4A,4BL | 0.44 | 4 | 0.67 | 0.61 |
| | cfe125-6A,6B | 0.44 | 6 | 0.74 | 0.71 |
| | cfe132-6A | 0.56 | 5 | 0.64 | 0.61 |
| | cfe166-7A,7B,7D | 0.56 | 3 | 0.57 | 0.49 |
| | cfe260-7A,7B,7D | 0.78 | 3 | 0.37 | 0.34 |
| | cfe270-4D,5A,5B | 0.56 | 4 | 0.62 | 0.57 |
| | cfe284-7A | 0.78 | 3 | 0.37 | 0.34 |
| | Xbarc12-3A | 0.44 | 3 | 0.64 | 0.57 |
| | Xbarc19-3A | 0.89 | 2 | 0.20 | 0.18 |
| | Xbarc37-6A | 0.89 | 2 | 0.20 | 0.18 |
| | Xbarc206-6A | 0.78 | 3 | 0.37 | 0.34 |
| | Xbarc263-1AS | 0.56 | 4 | 0.62 | 0.57 |
| | Xbarc1021-3A | 1.00 | 1 | 0.00 | 0.00 |
| | Xcfa2019-7AL | 0.56 | 4 | 0.62 | 0.57 |
| | Xcfa2153-1A | 0.78 | 3 | 0.37 | 0.34 |
| | Xcfd58-1A,1D | 0.22 | 5 | 0.79 | 0.76 |
| | Xcfd79-3A | 0.22 | 6 | 0.81 | 0.79 |
| | Xgwm2-3A | 0.44 | 5 | 0.72 | 0.68 |
| | Xgwm5-3A | 0.89 | 2 | 0.20 | 0.18 |
| | Xgwm135-1A | 0.33 | 5 | 0.77 | 0.73 |
| | Xgwm136-1A | 0.67 | 3 | 0.49 | 0.44 |
| | Xgwm160-4A | 0.44 | 6 | 0.74 | 0.71 |
| | Xgwm357-1A | 0.67 | 2 | 0.44 | 0.35 |
| | Xgwm369-3A | 0.22 | 7 | 0.84 | 0.82 |
| | Xgwm666-1A | 0.67 | 3 | 0.49 | 0.44 |
| | Xwmc59-1A | 0.56 | 3 | 0.57 | 0.49 |
| | Xwmc167-2A | 0.56 | 4 | 0.62 | 0.57 |
| | Xwmc169-3A | 0.89 | 2 | 0.20 | 0.18 |
| | Xwmc254-1A | 0.67 | 3 | 0.49 | 0.44 |
| | Xwmc256-6A | 0.56 | 4 | 0.62 | 0.57 |

**M.A.F** Major allele frequency, No**.** Number, **PIC** Polymorphism information content.

A total of 197 alleles were detected with these primer sets, located on 12 chromosomes with six chromosomes from each of the genomes as is shown in Table 1. The average number of alleles per locus was four with the minimum number detected at microsatellite loci Xbarc1021, Xgwm285 and Xwmc25. In the A genome, 136 alleles were detected that varied from the 61 alleles detected in the B genome. This shows a difference in the amount of allelic variation contributed by each genome in triticale.

However, the average alleles detected per locus between the two genomes did not vary significantly; five in the B genome and four in the A genome. This expresses a slight difference in the allelic variation detected in these two genomes in triticale. PIC and gene diversity values were calculated to better comprehend the genetic differentiation among the screened lines. These values are listed in Table 1. Not much difference was noted in the PIC values between the A genome (0.5) and B genome (0.59) implying that both genomes express relatively the same genetic variation in triticale. Primers highlighted in grey were chosen for the construction of the UPGMA dendogram. A subset of five markers (cfe2, cfe53, Xcfd79, Xgwm108, Xgwm369) covering chromosomes 2AL, 3A, 6A, 3B and 6B were selected to calculate the frequency-based distances (Table 2) amongst the cultivars, and consequently generate UPGMA and NJ dendograms. These markers were chosen based on the following criteria;

- High PIC value (≥0.8);
- Repeatable and distinct banding pattern; and
- Ability to detect differences among cultivars compared to other SSRs.

**Table. 1B**: A statistical summary generated in PowerMarker v.3.25 for the B genome.

| Genome | Marker | M.A.F | Allele No | Gene diversity | PIC |
|---|---|---|---|---|---|
| B-genome | cfe274-4BL | 0.56 | 5 | 0.64 | 0.61 |
| | Xcn13-6BS | 0.33 | 6 | 0.79 | 0.76 |
| | Xgwm18-1B,4B | 0.44 | 5 | 0.72 | 0.68 |
| | Xgwm46-7BS | 0.22 | 7 | 0.84 | 0.82 |
| | Xgwm108-3B | 0.22 | 6 | 0.81 | 0.79 |
| | Xgwm165-4BL/S | 0.78 | 3 | 0.37 | 0.34 |
| | Xgwm285-3B | 0.89 | 2 | 0.20 | 0.18 |
| | Xgwm340-3B | 0.44 | 5 | 0.72 | 0.68 |
| | Xgwm369-3A | 0.22 | 7 | 0.84 | 0.82 |
| | Xgwm389-3B | 0.56 | 4 | 0.62 | 0.57 |
| | Xgwm429-2B | 0.33 | 6 | 0.79 | 0.76 |
| | Xgwm493-3B | 0.44 | 4 | 0.67 | 0.61 |
| | Xgwm550-1B | 0.33 | 6 | 0.79 | 0.76 |
| | Xwmc25-2B | 0.89 | 2 | 0.20 | 0.18 |

**M.A.F** Major allele frequency, No**.** Number, **PIC** Polymorphism information content

### 3.2. R genome markers

Out of the 29 rye microsatellite markers tested, 23 (9 ESTs and 14 genomic) cross-amplified showing a 79.3% transferability of these markers to triticale. This percentage far surpasses the 39% transferability of 28 rye microsatellites among 80 global accession of triticale assessed by Kuleung *et al* [36]. Approximately 20.7% of the microsatellites were excluded due to their inability to amplify within the rye genome or due to the ambiguity of the band fragments produced on the PAGE gels. According to Leonova *et al.* [37], such band fragments may not even contain any SSR sequences. A sum total of 102 alleles were detected spanning over chromosomes 1R, 2R, 3R, 5R, 6R and 7R. Primer set SCM80 detected only one allele while SCM120 detected eight alleles and an average of four alleles was noted. These results are summarised in Table 3. Furthermore, the computed major allele frequencies (M.A.F) ranged from 0.2 to 1 and exhibited an indirectly proportional relationship to the PIC values expressed by each primer set. For instance, SCM80 had the highest M.A.F (1) but the lowest PIC (0) while Xrems1266 that had a low M.A.F (0.3), had a higher PIC value (0.8). An average PIC value of 0.55 was detected for all loci conveying moderate variability among the rye and triticale accessions. This is similar to the 0.54 average variability detected among 80 hexaploid triticale accessions sourced from 17 countries (South Africa included) by Kuleung *et al.* [36].

### 3.3. Dendogram construction

When comparing the genetic diversity detected by the EST primers and by the genomic primers, it was found that the nine ESTs exhibited an average PIC of 0.58 while the 14 genomic microsatellites exhibited an average of 0.55. This hints at the detection of moderate polymorphisms in the rye genome with fewer EST SSRs than genomic SSRs. Primers highlighted in grey were chosen for the construction of the UPGMA dendogram.UPGMA and NJ dendograms (figures 1 and 2) were generated for the wheat genome SSRs. Both trees were bootstrapped to 1000 permutations and a consensus value of 70% was computed. Both trees generated bootstrap values below 70% indicative of unsupported clustering. The SBL (shortest branch length) values were 3.84 and 3.71 for the UPGMA and NJ dendograms respectively. The wheat cultivar "SST88" is an out-group in this study cohort and was used as such when generating both trees.

This cultivar was also used to root both trees. Five SSRs were chosen based on their distinct and repeatable banding patterns, PIC values (≥0.8) and the ability to distinguish among the closely related rye and triticale cultivars. Frequency-based distances (Table 4) computed for primer sets SCM38 and Xrems1266 spanning chromosomes 3R and 5R respectively were used to generate UPGMA and NJ dendograms (figures 3 and 4). Both trees were bootstrapped to 1000 permutations. An SBL value of 2.70 was generated for both trees. Rye cultivars, "Henoch" and "Duiker" were used as out-groups in the generated trees. Again, both dendograms exhibited the same relationships among the cultivars.

Therefore, a consensus tree was generated to show the overall relationships among the spring triticale while using data generated by all seven markers developed in wheat and rye. This NJ tree is shown in figure 5. This tree manages to cluster cultivars "US2008" and "US2010" which according to the available pedigree, are sister lines. Also shown, are the clusterings of "Bacchus" and "US2007" and "Tobie" and "US2009" which share parents with similar pedigrees. Cultivar "Rex" is also more closely related to the other triticale cultivars unlike cultivar "Ibis" which is closer to durum wheat than triticale and this is shown in the NJ dendogram generated. Therefore, the panel of seven SSR markers developed in wheat and rye, managed to show differences among the commercial triticale cultivars. These markers also generated phylogenetic fingerprints that exhibited relationships that are consistent with the available pedigree data and ultimately managed to accurately identify the commercial cultivars.

**Table 2:** Frequency-based distances computed for wheat genome SSRs using the CS Chord distance method.

| OTU | Bacchus | Ibis | Rex | SST | Tobie | US2007 | US2008 | US2009 | US2010 |
|---|---|---|---|---|---|---|---|---|---|
| **Bacchus** | 0.00 | 0.72 | 0.90 | 0.90 | 0.72 | 0.72 | 0.90 | 0.72 | 0.90 |
| **Ibis** | 0.72 | 0.00 | 0.72 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| **Rex** | 0.90 | 0.72 | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| **SST** | 0.90 | 0.90 | 0.90 | 0.00 | 0.72 | 0.90 | 0.90 | 0.90 | 0.72 |
| **Tobie** | 0.72 | 0.90 | 0.90 | 0.72 | 0.00 | 0.72 | 0.90 | 0.72 | 0.90 |
| **US2007** | 0.72 | 0.90 | 0.90 | 0.90 | 0.72 | 0.00 | 0.90 | 0.90 | 0.90 |
| **US2008** | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.00 | 0.90 | 0.90 |
| **US2009** | 0.72 | 0.90 | 0.90 | 0.90 | 0.72 | 0.90 | 0.90 | 0.00 | 0.90 |
| **US2010** | 0.90 | 0.90 | 0.90 | 0.72 | 0.90 | 0.90 | 0.90 | 0.90 | 0.00 |

O.T.U Observed taxonomic unit.

**Table. 3**: A statistical summary generated in PowerMarker v.3.25 for the R genome.

| Marker | M.A.F | Allele No. | Gene Diversity | PIC |
|---|---|---|---|---|
| SCM2-6RL | 0.70 | 2 | 0.42 | 0.33 |
| SCM4-1R | 0.50 | 4 | 0.66 | 0.61 |
| SCM5-3RL | 0.60 | 4 | 0.58 | 0.54 |
| SCM9-1RS | 0.80 | 3 | 0.34 | 0.31 |
| SCM36-1R | 0.40 | 6 | 0.76 | 0.73 |
| SCM38-2R | 0.30 | 6 | 0.80 | 0.77 |
| SCM39-1R | 0.80 | 3 | 0.34 | 0.31 |
| SCM40-7R | 0.40 | 7 | 0.78 | 0.76 |
| SCM43-2R | 0.60 | 4 | 0.58 | 0.54 |
| SCM66-3R | 0.80 | 2 | 0.32 | 0.27 |
| SCM69-2R | 0.30 | 7 | 0.82 | 0.80 |
| SCM75-2RL | 0.80 | 3 | 0.34 | 0.31 |
| SCM80-1R | 1.00 | 1 | 0.00 | 0.00 |
| SCM83-2R | 0.60 | 2 | 0.48 | 0.36 |
| SCM86-7R | 0.40 | 3 | 0.64 | 0.56 |
| SCM109-5RL | 0.40 | 5 | 0.72 | 0.68 |
| SCM112-3R | 0.40 | 6 | 0.76 | 0.73 |
| SCM120-5RL | 0.20 | 8 | 0.86 | 0.84 |
| SCM206-3RS | 0.70 | 4 | 0.48 | 0.45 |
| Xrems1135-7R | 0.30 | 5 | 0.76 | 0.72 |
| Xrems1162-7R | 0.40 | 5 | 0.72 | 0.68 |
| Xrems1197-7R | 0.50 | 5 | 0.68 | 0.64 |
| Xrems1266-5R | 0.30 | 7 | 0.82 | 0.80 |

**M.A.F** Major allele frequency, **No.** Number, **PIC** Polymorphism information content

**Table. 4**: Frequency-based distances computed for rye genome SSRs using the CS Chord 1967distance method

| OTU | Bacchus | Duiker | Henoch | Ibis | Rex | Tobie | US2007 | US2008 | US2009 | US2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Bacchus** | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.00 | 0.45 | 0.90 | 0.90 | 0.45 |
| **Duiker** | 0.90 | 0.00 | 0.90 | 0.90 | 0.45 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| **Henoch** | 0.90 | 0.90 | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| **Ibis** | 0.90 | 0.90 | 0.90 | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.45 | 0.90 |
| **Rex** | 0.90 | 0.45 | 0.90 | 0.90 | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| **Tobie** | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.00 | 0.45 | 0.90 | 0.90 | 0.45 |
| **US2007** | 0.45 | 0.90 | 0.90 | 0.90 | 0.90 | 0.45 | 0.00 | 0.90 | 0.90 | 0.90 |
| **US2008** | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.00 | 0.45 | 0.90 |
| **US2009** | 0.90 | 0.90 | 0.90 | 0.45 | 0.90 | 0.90 | 0.90 | 0.45 | 0.00 | 0.90 |
| **US2010** | 0.45 | 0.90 | 0.90 | 0.90 | 0.90 | 0.45 | 0.90 | 0.90 | 0.90 | 0.00 |

**O.T.U** Observed taxonomic unit.

### 3.4. Polymorphisms detected in the A, B and R genomes

Interestingly, only a few B genome SSRs are needed to detect the moderate genetic variation among the triticale which is not the case for the A genome SSRs. This could be due to the fact that eight of the 35 A genome markers that produced amplicons did not show exclusive clustering to the A genome. These eight SSRs also show clustering in the D and B genomes. This could explain why the scale gets a little swayed to the B genome side in terms of genetic variability detected.

Of note, is the actual amplification of some of the D genome SSR fragments in hexaploid triticale which essentially lacks this genome. Tams *et al.* [17] attained amplification of hexaploid triticale DNA fragments while using D genome SSRs mapping to all but chromosomes 1D and 2D. Leonova *et al.* [37] suggests that this may be associated to wheat-rye or wheat-wheat translocations within triticale. Similarly, it has been shown that the change in triticale genomic composition and gene expression results from the effect of rye chromosomes and not the wheat-rye translocation composition. Both ways, confirmatory studies need to be done using *in situ* hybridization and C-banding [26]. The detected overall average PIC value (0.52) in this study resembles the 0.54 value obtained by Kuleung *et al.* [36] when screening 80 hexaploid triticale accessions of global origin, using 43 wheat developed SSRs. This implies that the Elite and commercially

available triticale cultivars bred in the SU-PBL exhibit a genetic diversity equivalent to that of the global triticale germplasm screened by Kuleung *et al* [36]. Comparatively in a study assessing the genetic diversity among 54 Brazilian triticale accessions using 42 wheat SSRs, Da Costa *et al.* [38] recorded an average PIC value of 0.36 which is much lower than that obtained in this study. This was not surprising as most of the accessions were of Mexican origin. This implies that the South African spring triticale screened have a higher level of genetic variability than the triticale accessions that were of Mexican origin. An average PIC value of 0.55 was detected in the R-genome conveying moderate variability among the rye and triticale accessions. This is similar to the 0.54 average variability detected among 80 hexaploid triticale accessions sourced from 17 countries (South Africa included) by Kuleung *et al.* [18]. In comparing the genetic diversity detected by the EST primers as opposed to the genomic primers, it was found that the nine ESTs exhibited an average PIC of 0.58 while the 14 genomic microsatellites exhibited an average of 0.55. This hints at the detection of moderate polymorphisms in the rye genome with fewer EST SSRs than genomic SSRs. Most EST SSRs either have a gene function or a close linkage to coding regions, which is not established in most genomic SSRs. For instance, Hackauf and Wehling [6] did a BLASTX search of the 70 rye ESTs and found 44.5% of the SCM loci could be associated with proteins of known or unknown function, and primer set SCM80 was found to be associated with a putative RNA-binding protein the *Oryza sativa* species. This implies that EST SSRs have the potential to increase the efficiency of SSR marker implementation in plant breeding as they express a higher transferability among related species, and have an ability to detect variation in the transcribed regions genomes such as rye [6, 14].

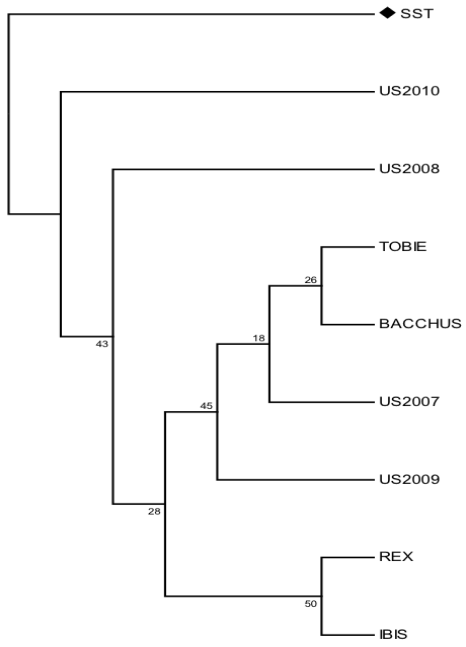


**Fig. 1:** A UPGMA dendogram showing the relationship of eight spring triticale cultivars based on CS Chord frequency-based distances generated by five SSRs developed in wheat. "SST" is used as the out-group and is used to root the tree.

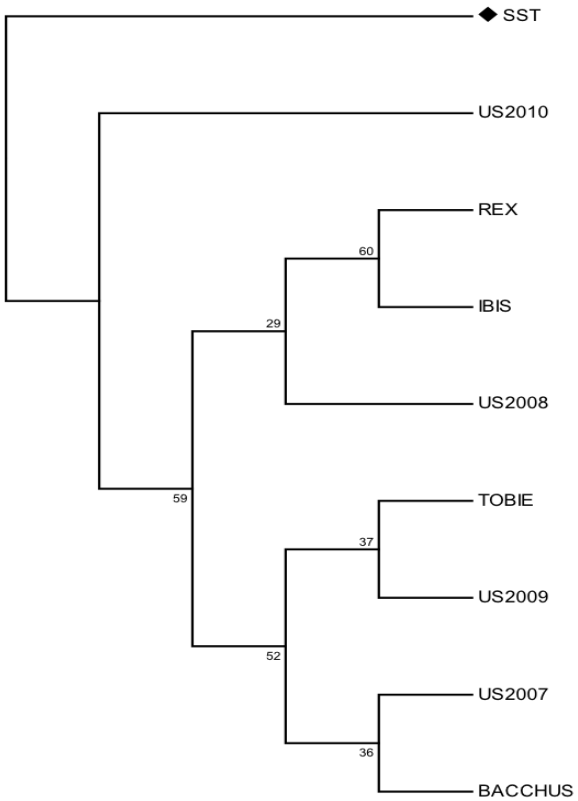


**Fig. 2:** A NJ dendogram showing the relationship of eight spring triticale cultivars based on CSChord frequency-based distances generated by five SSRs developed in wheat. "SST" is the out-group and is used to root the tree.

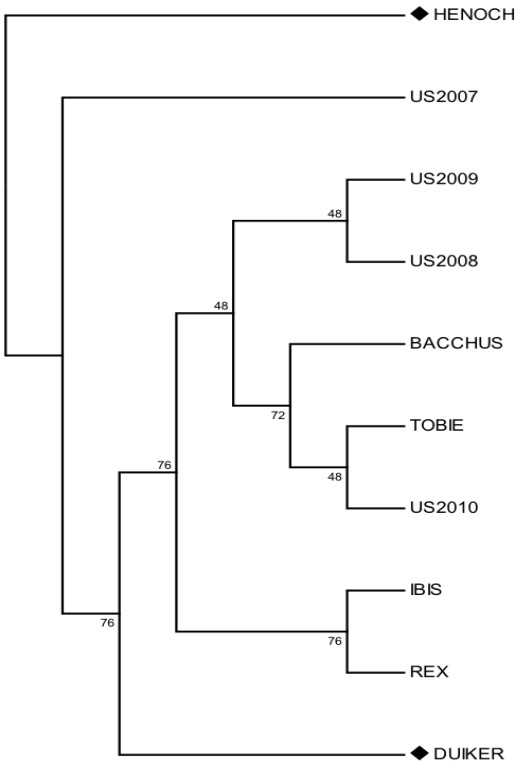


**Fig. 3:** A UPGMA dendogram showing the relationship of eight spring triticale cultivars based on CSChord frequency-based distances generated by two SSRs developed in rye. Rye cultivars "Henoch" and "Duiker" are out-groups and were used to root the tree.

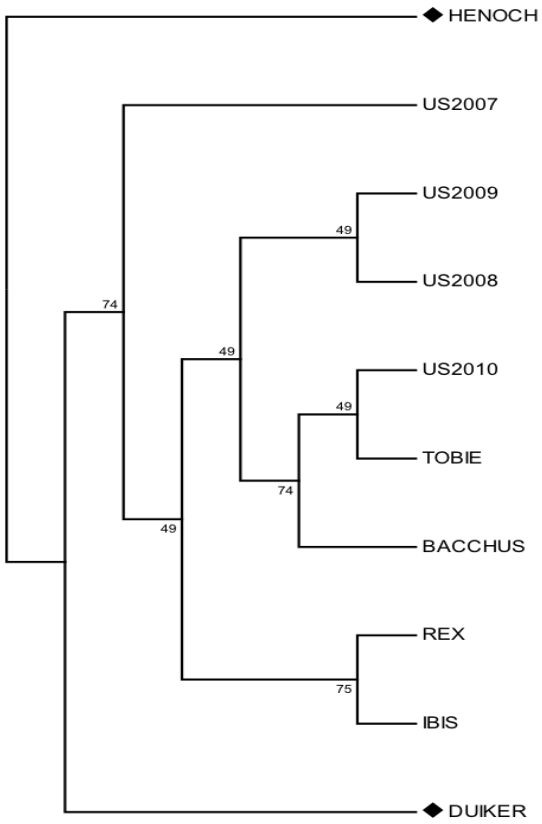


**Fig. 4:** A NJ dendogram showing the relationship of eight spring triticale cultivars based on CSChord frequency-based distances generated by two SSRs developed in rye. Rye cultivars "Henoch" and "Duiker" are out-groups and were used to root the tree.

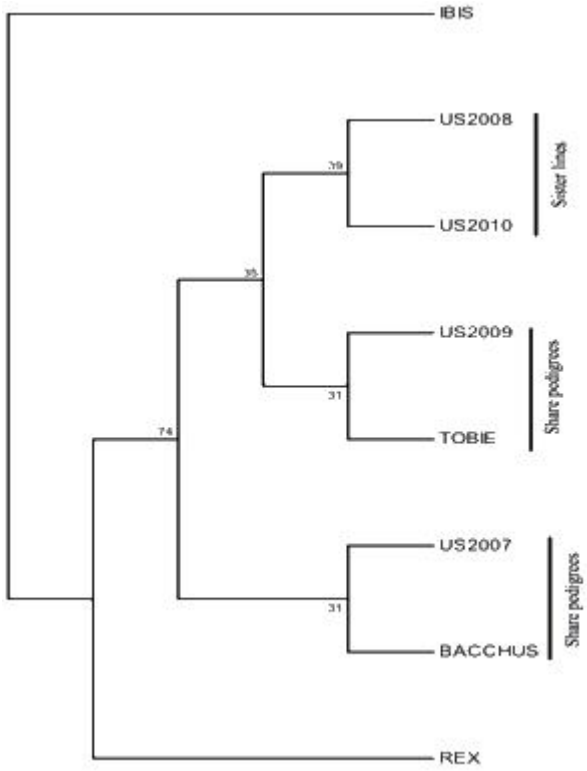


**Fig. 5:** A NJ dendogram showing the relationships among eight spring triticale cultivars using frequency-based distances generated by seven SSRs developed in wheat and rye.

## 4. CONCLUSIONS

SSR markers that are specific to the wheat genome were successfully optimised and transferred to the Elite and commercially thriving triticale cultivars in the SU-PBL. These markers showed an overall transferability of 80% and managed to detect a moderate genetic variability (PIC 0.52) among the genotypes. SSR markers specific to the rye genome were also optimised and showed 79.3% transferability across the triticale cultivars. These markers also detected moderate levels of genetic variability among the genotypes (PIC 0.55).

Data generated by a panel of seven SSR markers specific to both the wheat and rye genomes, was used to generate phylogenetic fingerprints for the cultivars. These markers were chosen because of their ability to discern the differences among the triticale cultivars and their ability to exhibit high levels of polymorphisms. The generated fingerprints managed to group those cultivars that shared parents with the same pedigrees and effectively managed to accurately identify the cultivars. Therefore, this panel of seven SSR markers was chosen for the implementation of a high throughput analysis.

## 5. REFERENCES

1. Gupta, P, K, Varshney, R, K, Sharma, P, C and Ramesh, B. Molecular markers and their applications in wheat breeding. Plant Breeding. 1999; 118:369-390.
2. Röder MS, Lapitan NLV, Sorrells ME, Tanksley SD. Genetic and physical mapping of barley telomeres. Molecular and General Genetics. 1993; 238:294-303.
3. Röder, MS, Korzun, V, Wendehake, K, Plaschke, J, Tixier, MH, Leroy, P and Ganal, MW. A microsatellite map of wheat. Genetics, 1998; 149:2007-2023.
4. Röder MS, Plaschke J, Konig SU,Sorrells MS. Abundance, variability and chromosomal location of microsatellite in wheat.Molecular and General Genetics. 1995; 246:327-333.
5. Singh NK, Shepherd KW, McIntosh RA. Linkage mapping of genes for resistance to leaf, stem and stripe rusts and ω-secalins on the short arm of rye chromosome 1R. Theoretical and Applied Genetics. 1990; 80:609-616.
6. Hackauf B, Wehling P. Identification of microsatellite polymorphisms in an expressed portion of the rye genome.Plant Breeding. 2002; 121:17-25.
7. Khlestkina EK, Than MHM, Pestsova EG, Röder MS, Malyshev SV, Korzun V,Börner A. Mapping of 99 new microsatellite-derived loci in rye (*Secale cereale* L.) including 39 expressed sequence tags. Theoretical and Applied Genetics. 2004; 109: 725-732.
8. Devos KM, Atkinson MD, Chinoy CN, Francis HA, Hartcourt RL, Koebner RMD, Liu CJ, Masojc P, Xie DX, Gale MD. Chromosomal rearrangements in the rye genome relative to that of wheat. Theoretical and Applied Genetics. 1993; 85:673-680.
9. Senft P, Wricke G. An extended genetic map of rye (*Secale cereale*L.). Plant Breeding. 1996; 115:508-510.
10. Korzun V, Malyshev S, Voylokov AV, Börner A. A genetic map of rye (*Secale cereale* L.) combining RFLP, isozyme, protein, microsatellite and gene loci. Theoretical and Applied Genetics. 2001; 102: 709-717.
11. Schubert AC. Quantitative trait loci analysis in spring wheat comparing two advanced backcross populations derived from an exotic wheat accession. University of Bonn (PhD dissertation). 2009.
12. Song QJ, Fickus EW,Cregan PB. Characteristics of trinucleotide markers in wheat. Theoretical and Applied Genetics. 2002; 104:286-293.

13. Song Q, Shi J, Singh S, Fickus E, Costa J, Lewis J, Gill B, Ward R,Cregan P. Development and mapping of microsatellite (SSR) markers in wheat. Theoretical and Applied Genetics. 2005; 110:550-560.
14. Zhang LY, Bernard M, Leroy P, Feuillet C,Sourdille P. High transferability of bread wheat EST-derived SSRs to other cereals. Theoretical and Applied Genetics. 2005; 111:677-687.
15. Varshney RK, Thiel T, Stein N, Langridge P,Graner A. In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cellular and Molecular Biology Letters. 2002; 7:537-546.
16. Tams SH, Bauer E, Oettler G, Melchinger AE. Genetic diversity in European winter triticale determined with SSR markers and coancestry coefficient. Theoretical and Applied Genetics. 2004; 108:1385-1391.
17. Tams SH, Melchinger AE, Bauer E. Genetic similarity among European winter triticale Elite germplasms assessed with AFLP and comparisons with SSR and pedigree data. Plant Breeding. 2005; 124:154-160.
18. Kuleung C, Baezinger PS, Kachman SD,Dweikat I. Evaluating the genetic diversity of triticale with wheat and rye SSR markers. Crop Science. 2006; 46:1692-1700.
19. Honing, J. Evaluation and implementation of DNA-based diagnostic methodology to distinguish wheat genotypes. University of Stellenbosch (Master's thesis). 2007.
20. Ma ZQ, Röder M, Sorrells, ME. Frequencies and sequence characteristics of di-, tri-, and tetra-nucleotide microsatellites in wheat. Genome. 1996; 39:123-130.
21. Rampling LR, Harker N, Shariflou MR, Morrel MK. Detection and analysis systems for microsatellite markers in wheat. Australian Journal of Agricultural Research. 2001; 52:1131-1141.
22. Gupta PK, Rustgi S, Sharma S, Singh R, Kumar N,Balyan HS. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat.MolecularGenetics and Genomics. 2003; 270:315-323.
23. Sourdille P, Tavaud M, Charmet G,Bernard M. Transferability of wheat microsatellites to diploid *Triticeae* species carrying the A, B and D genomes. Theoretical and Applied Genetetics. 2001; 103:346-352.
24. Somers D, Isaac P, Edwards, K. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum*L.). Theoretical and Applied Genetics. 2004; 109:1105-1114.
25. Kong L, Ohm HW, Cambron SE,Williams CE. Molecular mapping determines that Hessian fly resistance gene H9 is located on chromosome 1 A of wheat. Plant Breeding. 2005; 124:525-531.
26. Zhang LY, Bernard M, Raveli C, Balfourier F, Leroy P, Feuillet C, Sourdille, P. Wheat EST-SSRs for tracing chromosome segments from a wide range of grass species.Plant Breeding. 2007; 126:251-258.
27. Rözen L, Skaletsky G. Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology. 2000; 132:365-86.
28. Saal B, Wricke G. Development of simple sequence repeat markers in rye (*Secale cereale*L.).Genome. 1999; 42:964-972.
29. Rychlik W, Rhoads RE. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Research. 1989; 17: 8543-8551.
30. Tixier MH, Sourdille P, Röder M, Leroy P, Bernard M. Detection of wheat microsatellites using a non radioactive silver-nitrate staining method. Journal of Genetics Breeding. 1997; 51:175-177.
31. Cavalli-Sforza LL, Edwards AWF. Phylogenetic analysis: models and estimation procedures. American Journal of Human Genetics. 1967; 19:233-257.
32. Weir BS, Genetic data analysis II. Sinauer Associates, Incorporated, Sunderland. 1996.
33. Liu, K, Muse SV. PowerMarker: Intergrated environment for genetic marker data. Bioinformatics. 2005;21(9):2128-2129.
34. Sneath PHA,Sokal RR. Numerical Taxonomy, W, H, Freeman, San Franciso, 1973.
35. Saitou N, Nei M. The neighbour-joining method; a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 1987; 4:406-425.
36. Kuleung C, Baezinger PS, Dweikat I. Transferability of SSR markers among wheat, rye and triticale. Theoretical and Applied Genetics. 2004; 108: 1147-1150.
37. Leonova IN, Dobrovolskaya OB, Kaminskaya LN, Adonina IG, Koren LV, Khotyljova LV, Salina EA. Molecular analysis of the triticale lines with different *vrn*gene systems using microsatellite markers and hybridisation *in situ*. Russian Journal of Genetics. 2005; 41:1014-1020.
38. Da Costa CT, Albuquerque ACS, Junior AN, Marcelino FC, Pereira JF. Genetic diversity of Brazilian triticale evaluated with genomic wheat microsatellites. Pesquisa Agropecuária Brasília. 2007; 42(11): 1577-1586.



**How to cite this article:**

W.C. Botes and D. Bitalo., Identification, evaluation and optimization of a minimum simple sequence repeat marker set for triticale breeding. J App Biol Biotech. 2013; 1 (04):016-023