© Indian Academy of Sciences

RESEARCH NOTE

Genetic diversity among old Portuguese bread wheat cultivars and botanical varieties evaluated by ITS rDNA PCR-RFLP markers

A. CARVALHO*, H. GUEDES-PINTO and J. LIMA-BRITO

Institute for Biotechnology and Bioengineering (IBB), Centre of Genetics and Biotechnology (CGB) of University of Trás-os-Montes and Alto Douro (IBB/CGB-UTAD), P. O. Box 1013, 5001-801 Vila Real, Portugal

Introduction

In contrast to the highly conserved 18S-5.8S-25S ribosomal RNA (rRNA) genes, the ribosomal DNA (rDNA) spacers, such as the internal transcribed spacer (ITS) and the intergenic spacer (IGS) present high variability and faster rates of evolution (Baldwin *et al*. 1995; Barker *et al*. 1988). ITS length variants have been reported for several plant species, including hexaploid wheat (Baldwin *et al*. 1995; Álvarez and Wendel 2003; Nalini *et al*. 2007; Saini *et al*. 2008). The ITS polymorphism might occur at the genus, species or individual levels, making it useful for phylogenetic, evolutionary and biogeographical studies.

The present study undertakes a genetic diversity analysis of an old Portuguese bread wheat collection, comprising of 48 cultivars belonging to nine botanical varieties, based on the ITS rDNA variation. This collection constitutes an excellent repository of germplasm in Portugal. Additionally, we intend to test the reliability of the ITS rDNA PCR-RFLP markers for defining phylogenies among bread wheat botanical varieties which could be helpful for designing wheat breeding strategies.

Materials and methods

Plant material

The old Portuguese bread wheat collection was kindly provided by the Instituto Nacional de Investigação Agrária e das Pescas (INIAP), Elvas, Portugal (see table 1 in electronic supplementary material at http://www.ias.ac.in/jgenet/). Homonym cultivars were denoted by (a) and (b). According to the passport data, the 48 cultivars belong to nine different botanical varieties (see table 1 in electronic supplementary material). Regarding their geographical origin/distribution areas, these cultivars might be considered sympatric across the centre and south of Portugal.

ITS rDNA amplification and PCR-RFLP

Genomic DNA was extracted from young leaves by the CTAB based protocol. For the amplification of the rDNA ITS1-5.8S-ITS2 region as a single molecule, we used the reverse primer ITS-4 (TCCTCCGCTTATTGATATGC) (White *et al*. 1990) and a modified forward primer (GTC-CACTGAACCTTATCATTTAG) reported by Urbatsch *et al*. (2000).

Amplification reactions (50 μL) were constituted by: 50 ng of template DNA, 1× KCl reaction buffer (Fermentas, Burlington, USA), 5 mM $MgCl_2$, 0.25 mM dNTPs mix, 62 μM of each primer and 4 units of *Taq* DNA polymerase (Fermentas, Burlington, USA). The ITS region was amplified under the following conditions: an initial denaturation for 5 min at 94°C, 34 thermal cycles of 1 min at 94°C, 1 min at 53°C and 1 min at 72°C and a final extension step at 72°C for 5 min. Ten μL of each PCR product were analysed after electrophoresis on a 1.5% agarose gel and staining with ethidium bromide. The remaining PCR product (40 μL) was digested with 10 units of each restriction enzyme (*Alu*I, *Hpa*II, *Rsa*I and *Taq*I) in a 30 μL PCR-RFLP reaction, following manufacturer's instructions. The PCR-RFLP products were separated after electrophoresis on 2% agarose gels at a constant voltage of 90 V. Bands were visualized after staining with ethidium bromide and photographed with a CCD video imager (Vilber Lourmat, Eberhardzell, Germany). The size of restricted products was estimated with the molecular weight marker GeneRuler^{TM} 100-bp DNA Ladder Plus (Fermentas, Burlington, USA). We only scored fragments that were easily visualized.

*For correspondence. E-mail: anacarvalho@mail.com.

Keywords. internal transcribed spacer variation; PCR-RFLP; ribosomal DNA; *Triticum aestivum* L. em. Thell.

Statistical analysis

The ITS PCR-RFLP data were analysed with the software POPGENE 1.32 (Yeh *et al.* 1999). Diversity in the frequency of fragment-size patterns from digests was apportioned within and among botanical varieties using Shannon's information index - I (Shannon and Weaver 1949). Nei's (1973) gene diversity index ($h = 1 - \sum pi^2$), where p_i is the frequency of the ith allele of the locus, H_T the total genetic diversity, H_S the genetic diversity within varieties, G_{ST} the relative magnitude of differentiation among varieties, ($D_{ST} = H_T - H_S$) the inter-variety genetic diversity as well as the estimate of gene flow (Nm) from G_{ST} (Slatkin and Barton 1989), $Nm = 0.5\times(1-G_{ST})/G_{ST}$ (McDermott and McDonald 1993), were also calculated.

The phylogenetic relationships among the bread wheat botanical varieties were estimated after the construction of a cladogram with the TreeView software (Page 1996), based on the genetic distance (Nei 1987) and neighbour-joining method (Saitou and Nei 1987).

The genetic similarity among the 48 bread wheat cultivars was estimated by the Jaccard's coefficient between pairs. The phenogram was based on the sequential agglomerative hierarchical nested (SAHN) module and the unweighted pair group method with arithmetical averages (UPGMA), and it was constructed with the NTSYSpc 2.02 version software (Rohlf 1998).

Results and discussion

ITS variants detected by PCR-RFLP

All bread wheat cultivars presented a 700-bp PCR product of invariant length. All enzymes, except for *Rsa*I, revealed variation in the ITS region (figure 1). Each enzyme produced three ITS PCR-RFLP patterns among the 48 cultivars (table 1), with differences characterized by the presence of one or two extra bands (table 1). Except for the *Hpa*II enzyme, we considered the patterns composed by only four monomorphic bands as a homozygous genotype as well as those patterns constituted by the four monomorphic and one extra band. In this sense, we suggest the occurrence of a mix of the two homozygous genotypes in 46 and 47 cultivars, regarding the *Alu*I and *Taq*I results, respectively. However, for the *Hpa*II patterns, we could not achieve a confident identification of the heterozygous genotype (pattern e or f; table 1). Since this endonuclease is sensitive to methylation, we suppose that one of these patterns has one methylated *Hpa*II recognition site, avoiding the detection of a second extra band.

According to Flavell (1980), the number of chromosomal nucleolar organizer (NOR) regions correspond to the number of multiple ITS sequences found in the same taxon. We support this assumption, since the nine patterns varied from four to six fragments, resembling the maximum numbers of Ag-NORs that we previously scored per cultivar (data not shown).

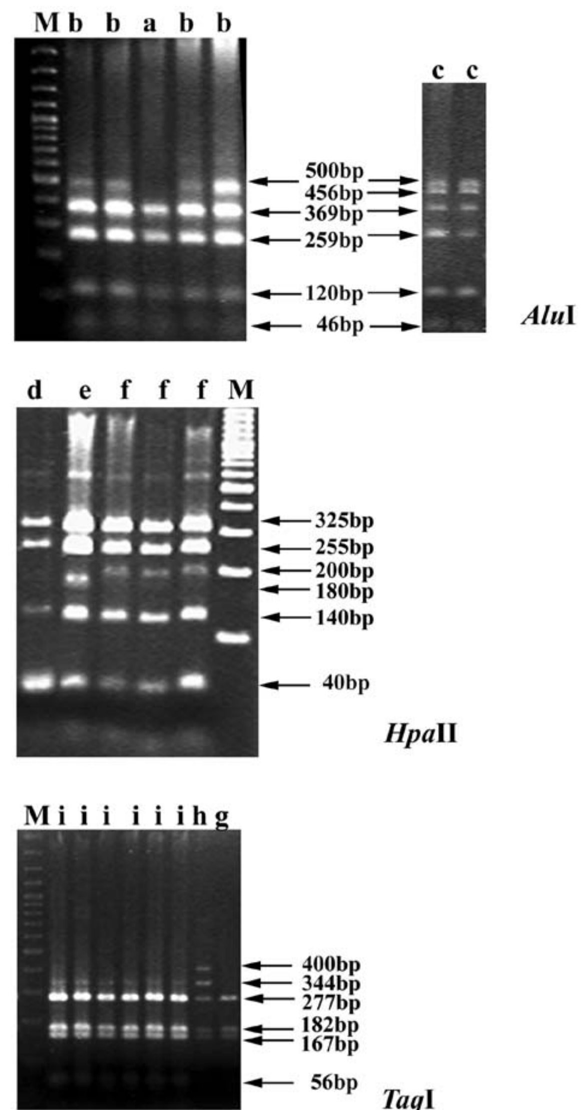

Figure 1. ITS PCR-RFLP patterns (indicated above each lane) produced by the *Alu*I (upper); *Hpa*II (middle) and *Taq*I (lower) in the 48 bread wheat cultivars. Arrows indicate both monomorphic and polymorphic bands. Lane M GeneRuler™ 100-bp DNA Ladder Plus (Fermentas, Burlington, USA).

The patterns b, d and i showed highest frequencies (table 1). The sum of the four monomorphic bands provided by the *Alu*I and *Hpa*II enzymes exceeded 700 bp, indicating the presence of different ITS types per individual genome. As reported by Nalini *et al.* (2007), our results confirmed that the PCR-RFLP technique was useful for the detection of ITS variants in hexaploid wheat.

The pool of the ITS PCR-RFLP data obtained with *Alu*I, *Hpa*II and *Taq*I allowed the detection of nine polymorphic loci out of 18, resulting in 50% polymorphism.

The diversity analysis yielded mean values (± s.d.) of Shannon's index (I) and Nei's gene diversity (h)

Table 1. ITS PCR-RFLP patterns achieved in the 48 bread wheat cultivars by *Alu*I, *Hpa*II and *Taq*I.

Enzyme	Patterns	Monomorphic bands (bp)	Polymorphic bands (bp)	Cultivars N°
	a	369, 259, 120, 46	None	16
*Alu*I	b	369, 259, 120, 46	500	30
	c	369, 259, 120, 46	500 and 456	2 (93 and 101)
	d	325, 255, 140, 40	None	20
*Hpa*II	e	325, 255, 140, 40	180	10
	f	325, 255, 140, 40	200	18
	g	277, 182, 167, 56	None	1 (code 50)
*Taq*I	h	277, 182, 167, 56	400 and 344	1 (code 48)
	i	277, 182, 167, 56	344	46

for all cultivars of 0.1980 (±0.2466) and 0.1246 (±0.1639), respectively, which were quite low. Additionally, the mean (± s.d.) total genetic diversity (H_T) was 0.0818 (±0.0170), while the mean (± s.d.) intra-variety genetic diversity (H_S) was 0.0361 (±0.0031). These two values were used for the determination of the inter-variety genetic diversity ($D_{ST} = H_T - H_S$), yielding a value of 0.0457, which is somewhat high. When we observed the morphological characteristics of the analysed cultivars in the field, we noticed differences among botanical varieties and, according to Nalini *et al.* (2007), some ITS polymorphic bands could be related to QTLs such as the spike size. This feature could be in the origin of the higher genetic diversity among rather than within varieties. Besides, the G_{ST} parameter (relative magnitude of differentiation among varieties) was reasonable (0.5587), explaining the high value of estimate of gene flow (*Nm*) which was 0.3949.

The UPGMA phenogram constructed with the pool of the ITS PCR-RFLP data revealed that most cultivars (47) were clustered on a major group and just one constituted a branch – the cultivar 'Mocho Cabeçudo', passport code 48 (figure 2a). Within the major group, the cultivars were clustered based on their ITS variant patterns (presented in table 1) and/or regarding the combination of homozygotic ITS patterns. Generally, the cultivars belonging to the same botanical variety were clustered together and most of them showed the highest value of genetic similarity (1.00), reflecting the reduced intra-variety genetic diversity (figure 2,a).

To estimate the phylogenetic relationships among the nine botanical varieties based on the ITS PCR-RFLP data, a cladogram (figure 2, b) and a phylogram (figure 2, c) were constructed with the Tree View software. A cladogram might provide the evolutionary paths of the taxonomic units, and it could be taken as an inferred phylogenetic tree, despite the unknowable temporal sequence of the events. The root of each cladogram is considered the most recent common ancestor. In figure 2b, seven clades or events originated the nine botanical varieties, the var. *aestivum* being the oldest one since it diverged directly from the ancestor. More recently, four clades gave rise to the varieties *plenoerythrospermum*, *ferrugineum*, *hostianum* and *nigroerythrospermum* (figure 2b). One clade diverged into the varieties *milturum* and *lutescens* suggesting that both had a common recent ancestor and explaining the clustering of some *lutescens* cultivars with the four belonging to var. *milturum* in the phenogram (figure 2a). The latest event seems to be in the origin of the *creticum* (code 49) and *graecum* (code 36) varieties which shared a more recent common ancestor and could be considered as sister groups, as well as *milturum* and *lutescens*. The cultivars *creticum* and *graecum* were placed together in the same cluster of the cultivars representative of the *hostianum* (code 100) and *nigroerythrospermum* (code 87) varieties (figure 2a). These four cultivars shared the same pattern with each restriction endonuclease; namely, the patterns a, d and g, characterized by the absence of extra bands (homozygotic genotypes; table 1). The closest genetic relationship among these four varieties was evident in the phylogram (figure 2c). This result could be explained by a higher rate of concerted evolution relatively to the remaining cultivars. Species with multiple rDNA loci, such as the allopolyploid wheat could undergo a slower rate of concerted evolution, but in these four cultivars belonging to the botanical varieties *creticum*, *graecum*, *hostianum* and *nigroerythrospemum*, it could have been accelerated by mechanisms of unequal crossing over and recombination or biased gene conversion. Both the cladogram (figure 2b) and the phylogram (figure 2c) resembled the UPGMA clustering of the 48 bread wheat cultivars (figure 2a).

ITS sequences have been widely used in several species (Hsiao *et al.* 1994; De Bustos and Jouve 2002; Sharma *et al.* 2002) for phylogenetics inferences. We found that the ITS PCR-RFLP markers were less polymorphic than other DNA based markers, but constitute feasible and specific tools for genetic variation and diversity assessment among bread wheat cultivars and higher taxa, such as botanical varieties, and are useful for the estimation of phylogenies. The knowledge of genetic relationships and phylogenies might contribute for the designing of intraspecific crosses between cultivars of this collection with potential interest for wheat breeding programmes.

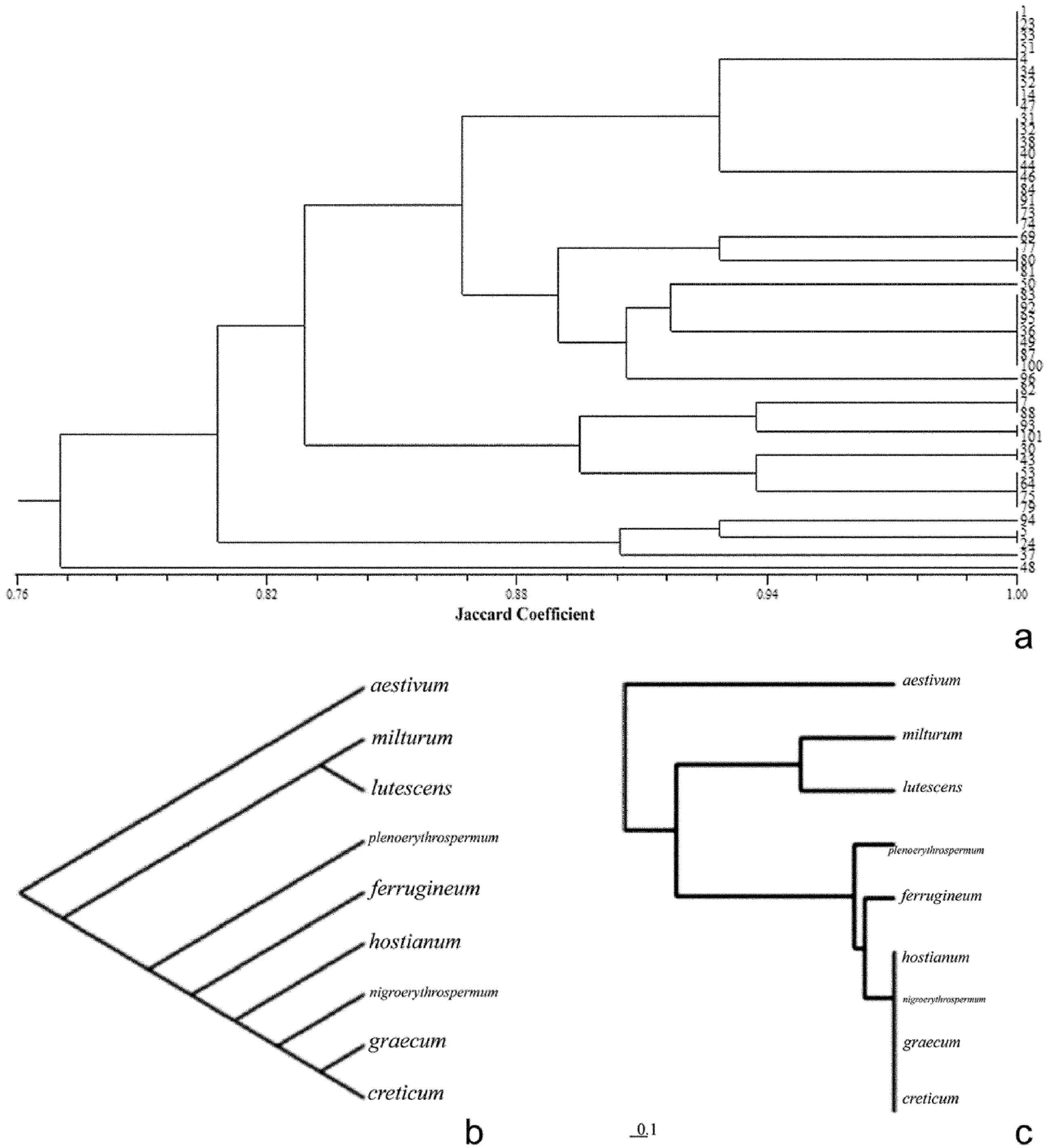

Figure 2. (a) Phenogram generated by the UPGMA clustering of the ITS PCR-RFLP data obtained in the 48 bread wheat cultivars, (b) cladogram, and (c) phylogram of the Portuguese bread wheat botanical varieties constructed based on the pool of the ITS PCR-RFLP data.

Acknowledgements

This work was supported by the PTDC/AGR-GPL/65876/2006 project and the Ph.D. grant SFRH/BD/17348/2004 financed by the Portuguese Foundation for Science and the Technology (FCT).

References

Álvarez I. and Wendel J. F. 2003 Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogenet. Evol.* **29**, 435–455.

Baldwin B. G., Sanderson M. J., Porter J. M., Wojeiechowski M. F.,

Capbell C. S. and Donoghe M. J. 1995 The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* **82**, 247–277.

Barker R. F., Harberd N. P., Jarvist M. G. and Flavell R. B. 1988 Structure and evolution of the intergenic reion in a ribosomal DNA repeat unit of wheat. *J. Mol. Biol.* **201**, 1–17.

De Bustos A. and Jouve N. 2002 Phylogenetic relationships of the genus *Secale* based on the characterisation of rDNA ITS sequences. *Plant Syst. Evol.* **235**, 147–154.

Flavell R. B. 1980 The molecular characterization and organization of plant chromosomal DNA sequences. *Annu. Rev. Plant Physiol.* **31**, 569-596.

Hsiao C., Chatterton N. J., Asay K. H. and Jensen K. B. 1994 Phylogenetic relationships of 10 grass species: an assessment of phylogenetic utility of the internal transcribed spacer region in nuclear ribosomal DNA in monocots. *Genome* **37**, 112–120.

McDermott J. M. and McDonald B. A. 1993 Gene flow in plant pathosystems. *Annu. Rev. Phytopathol.* **31**, 353–373.

Nalini E., Bhagwat S. G. and Jawali N. 2007 Identification and characterization of some ITS variants from hexaploid wheat (*Triticum aestivum* L.). *Plant Sci.* **173**, 262–268.

Nei M. 1973 Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70**, 3321–3323.

Nei M. 1987 *Molecular evolutionary genetics*. Columbia University Press, New York, USA.

Page R. D. M. 1996 TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biol. Sci.* **12**, 357–358.

Rohlf F. J. 1998 *NTSYS-pc ver. 2.02. Numerical taxonomy and multivariate analysis system*. Exeter Publishing, Setauket, USA.

Saini A., Reddy S. K. and Jawali N. 2008 Intra-individual and intra-species heterogeneity in nuclear rDNA ITS region of *Vigna* species from subgenus *Ceratotropis. Genet. Res.* **90**, 299–316.

Saitou N. and Nei M. 1987 The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

Shannon C. and Weaver W. 1949 *The mathematical theory of communication*. University of Illinois Press, Urbana, USA.

Sharma S., Rustgi S., Balyan H. S. and Gupta P. K. 2002 Internal transcribed spacer (ITS) sequences of ribosomal DNA of wild barley and their comparison with ITS sequences in common wheat. *Barley Genet. Newslett.* **32**, 38–45.

Slatkin M. and Barton N. H. 1989 A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43**, 1349–1368.

Urbatsch L. E., Baldwin B. and Donoghue M. J. 2000 Phylogeny of coneflowers and relatives (*Heliantheae: Asteraceae*) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. *Syst. Bot.* **25**, 539–565.

White T. J., Bruns T., Lee S. and Taylor J. 1990 Amplification and direct sequencing of fungal ribosomal genes for phylogenetics. In *PCR protocols: a guide to methods and applications* (ed. M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White), pp. 315–322. Academic Press, Orlando, USA.

Yeh F. C., Yang R. C., Boyle T. B. J., Ye Z. H. and Mao J. X. 1999 POPGENE (the user-friendly shareware for population genetic analysis), version 1.32. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Canada.

Received 17 March 2009, in revised form 22 June 2009; accepted on 20 July 2009
Published on the Web: 16 October 2009