Contrasting Patterns of Evolution of 45S and 5S rDNA Families Uncover New Aspects in the Genome Constitution of the Agronomically Important Grass *Thinopyrum intermedium* (Triticeae)

Václav Mahelka,*[1] David Kopecký,[2] and Bernard R. Baum[3]

[1]Institute of Botany, Academy of Sciences of the Czech Republic, Průhonice, Czech Republic

[2]Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Academy of Sciences of the Czech Republic, Olomouc-Holice, Czech Republic

[3]Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada

***Corresponding author:** E-mail: vaclav.mahelka@ibot.cas.cz.

Associate editor: Brandon Gaut

Abstract

We employed sequencing of clones and in situ hybridization (genomic and fluorescent in situ hybridization [GISH and rDNA-FISH]) to characterize both the sequence variation and genomic organization of 45S (herein ITS1-5.8S-ITS2 region) and 5S (5S gene + nontranscribed spacer) ribosomal DNA (rDNA) families in the allohexaploid grass *Thinopyrum intermedium*. Both rDNA families are organized within several rDNA loci within all three subgenomes of the allohexaploid species. Both families have undergone different patterns of evolution. The 45S rDNA family has evolved in a concerted manner: internal transcribed spacer (ITS) sequences residing within the arrays of two subgenomes out of three got homogenized toward one major ribotype, whereas the third subgenome contained a minor proportion of distinct unhomogenized copies. Homogenization mechanisms such as unequal crossover and/or gene conversion were coupled with the loss of certain 45S rDNA loci. Unlike in the 45S family, the data suggest that neither interlocus homogenization among homeologous chromosomes nor locus loss occurred in 5S rDNA. Consistently with other Triticeae, the 5S rDNA family in intermediate wheatgrass comprised two distinct array types—the long- and short-spacer unit classes. Within the long and short units, we distinguished five and three different types, respectively, likely representing homeologous unit classes donated by putative parental species. Although the major ITS ribotype corresponds in our phylogenetic analysis to the E-genome species, the minor ribotype corresponds to *Dasypyrum*. 5S sequences suggested the contributions from *Pseudoroegneria*, *Dasypyrum*, and *Aegilops*. The contribution from *Aegilops* to the intermediate wheatgrass' genome is a new finding with implications in wheat improvement. We discuss rDNA evolution and potential origin of intermediate wheatgrass.

Key words: **Triticeae, polyploidy, hybridization, ribosomal DNA, concerted evolution, in situ hybridization.**

Introduction

Ribosomal DNA (rDNA) has belonged to one of the most widely used markers in plant biosystematics since the routine application of DNA sequencing. Its wide use is due to the fact that ribosomal RNA (rRNA) represents a highly conserved gene product present in all cells across all kingdoms (Smit et al. 2007). rRNA is an essential structural component of ribosomes. They are complex machines playing a central role in protein synthesis. Ribosomes consist of two subunits, each of which is, apart from various proteins, composed of one or more rRNA molecules. In eukaryotes, the large 60S ribosomal subunit (LSU) consists of three rRNA molecules (25–26S, 5.8S, and 5S) and 46 proteins, whereas the small 40S subunit includes one rRNA (18S) molecule and 33 proteins (Ben-Shem et al. 2011). In most organisms, three of the RNA molecules—the so-called large ribosomal genes 18S, 5.8S, and 26S—are arrayed in a conserved cluster, in plants referred to as 45S. The genes are separated by two internal transcribed spacers (ITS1 and ITS2) and together constitute a single transcription unit. These rDNA units are organized in large tandem arrays, the rDNA loci, on one or a small number of chromosomes. Diploid species from within the Triticeae such as diploid wheats usually carry two (or more) major 45S rDNA loci per haploid genome and may additionally carry a variable number of minor loci (Dubcovsky and Dvořák 1995). In seed plants, 5S rDNA genes are usually localized in separate clusters (Wicke et al. 2011). The 5S rDNA units are composed of a gene region, which varies little from 120 bp and a nontranscribed spacer (NTS) of highly variable length, both arranged in tandem arrays (Long and Dawid 1980). As huge numbers of ribosomes are necessary for protein synthesis during periods of rapid growth, there are hundreds to thousands of copies of transcription units of both 45S and 5S families within a haploid genome (Long and Dawid 1980).

One of the most fascinating features of rDNA as an exemplar member of a multigene family is its ability to maintain

Article

sequence homogeneity. The ability of all rDNA units to change their sequence in a highly synchronized manner has been described as concerted evolution (Arnheim et al. 1980). Although it was long thought that concerted evolution is a general model of evolution of multigene families leading to the homogenization of repeats, it is worth noting that most findings regarding concerted evolution concern major ribosomal genes. In 5S rRNA genes and nonribosomal genes, this model of evolution seems to be less satisfactory (Nei and Rooney 2005). The mechanisms of concerted evolution are not entirely clear. It is supposed, however, that unequal crossover and gene conversion are the major drivers behind this process (Eickbush TH and Eickbush DG 2007). Besides, purifying selection is also likely responsible for the high conservation of coding regions (Nei and Rooney 2005). The number and location of rDNA loci within a genome are likely playing a key role in the process of concerted evolution at the molecular level (Wendel 2000; Eickbush TH and Eickbush DG 2007): A large number of rDNA loci may slow down concerted evolution because homogenization occurs more readily within loci than between them. Furthermore, concerted evolution is facilitated by the near-telomeric location of rDNA arrays, whereas rDNA arrays more distant from telomeres may conserve multiple rDNA repeat types. When rDNA loci are dispersed throughout a genome, which is unusual in eukaryotes, such arrays may even escape from concerted evolution (Rooney 2004).

The process of sequence homogenization is usually very effective, so even ITS can be treated as a single gene, and if some variability in ITS is found, it is usually confined to organisms that are hybrids (but see Blattner 2004), be it diploid or polyploid (for a review, see Álvarez and Wendel 2003). The comparably faster evolutionary rates of nongenic regions of rDNA—the ITS of 45S and the NTS of 5S—make these regions phylogenetically useful and thus widely applicable in phylogenetics and other biosystematic disciplines (Hsiao et al. 1995; Scoles et al. 1988). Especially, the nrITS region of 45S rDNA has been among the most widely used markers in plant biosystematics. Its domination stems from its advantageous properties for phylogenetic inference: universality, biparental inheritance, intragenomic uniformity, and relative simplicity (Álvarez and Wendel 2003). Although these properties are largely valid for diploid nonhybrid organisms, in (allo)polyploids the situation is less straightforward. In polyploids, besides general pitfalls potentially complicating the interpretation of rDNA—for example, the presence of divergent paralogs (Buckler et al. 1997; Muir et al. 2001) or pseudogenes (Mayol and Rosselló 2001)—there are others that are typical of allopolyploids. It has been shown that in early generations of newly arisen polyploids rDNA arrays may undergo complex evolution and reorganization that reflect whole-genome rearrangements, which are sometimes unpredictable (Skalická et al. 2005; Baum and Feldman 2010; Renny-Byfield et al. 2011). In hybrids/allopolyploids, partly or completely unhomogenized copies are usually maintained (Fehrer et al. 2009; Mahelka and Kopecký 2010). When such intra-individual polymorphism is due to homeology (i.e., hybridization) and not paralogy (i.e., duplication), ITS represents a useful marker for inferring the origin of polyploid species (Soltis PS and Soltis DE 1991; Liu et al. 2006). However, there are examples in which different copies do not reflect real contributions from different parents, as ratios of respective parental sequences are skewed toward one ribotype or another via interlocus homogenization (Wendel et al. 1995), rDNA loci of either parent were lost (Kotseruba et al. 2010) or both processes took place (Volkov et al. 1999).

Although the paradigm of concerted evolution is largely based on studies dealing with large ribosomal genes, a growing body of studies questions whether concerted evolution is the prominent mechanism involved in the evolution of 5S RNA (Pinhal et al. 2011; Vizoso et al. 2011). Numerous studies have revealed the existence of remarkable 5S rDNA variants in such diverse groups of organisms as fungi (Rooney and Ward 2005), herbaceous plants (Kellogg and Appels 1995), trees (Negi et al. 2002), fishes (Pinhal et al. 2011) or molluscs (Vizoso et al. 2011). In some cases, this variation was due to the presence of distinguishable NTS classes residing in different chromosomal regions (Sastri et al. 1992); in others different 5S rDNA classes occupied one rDNA array (Falistocco et al. 2007), or distinct copies were dispersed throughout the genome (Rooney and Ward 2005). In the 5S rDNA family, the rate of concerted evolution is apparently dependent on its genomic organization, and if concerted evolution occurs, it operates within separate arrays, while little, if any, exchange occurs between arrays (Kellogg and Appels 1995; Cronn et al. 1996; Scoles et al. 1988). Thanks to this feature, the NTS has been successfully employed in identifying different haplomes of polyploid species (Baum and Johnson 2008). On the contrary, dispersion of arrays through a genome may lead to between-species gene clustering in phylogenies (Rooney and Ward 2005), a phenomenon that may potentially complicate the inference of species histories. As the observed pattern of 5S rDNA evolution could not be explained by concerted evolution alone, a new model called birth-and-death evolution was introduced (for a review, see Nei and Rooney 2005). In this model of evolution, new gene variants are produced by gene duplications, and some of the duplicated genes diverge functionally and are maintained in the genome, whereas others become pseudogenes because of deleterious mutations or get deleted from the genome. Unlike in concerted evolution, purifying selection is the major force that maintains high conservation of genes (Nei and Rooney 2005).

The focus of this study is the allohexaploid ($2n = 6x = 42$) grass *Thinopyrum intermedium* (intermediate wheatgrass). In a previous study, Mahelka et al. (2007) demonstrated extensive but incomplete homogenization of ITS sequences in the *T. intermedium* accessions analyzed. These results contrasted with a preliminary analysis of the same accessions carried out using 5S rDNA sequences, in which extensive variation was found and which is part of this study. Intermediate wheatgrass possesses many desirable agronomic traits that make it an invaluable source of genetic material useful in wheat improvement (Li and Wang 2009). However, its genome constitution has not been satisfactorily resolved so far (Mahelka et al. 2011; Tang et al. 2011—see Discussion), and every

piece of variation contained in the ribosomal genes might help clarify the species' origin. In this study, we therefore report on the analysis of rDNA in the allohexaploid grass *T. intermedium* at the sequence and cytogenetic levels. In particular, we analyzed and characterized sequence diversity of both 45S (ITS region) and 5S rDNA families. By employing genomic and fluorescent in situ hybridization (GISH and rDNA-FISH), we examined physical disposition of both rDNA families and discussed the evolutionary dynamics of both families with respect to the concerted or birth-and-death models of evolution. Finally, we compare the two markers as to their utility for haplome identification in this allohexaploid and present new aspects of its genomic constitution.

Results

Sequence Comparisons and Inference of Functionality

To examine the rates of homogenization and sequence functionality of 45S and 5S rDNA, we analyzed and characterized 67 ITS and 89 5S clones polymerase chain reaction (PCR) amplified from four accessions of *T. intermedium*.

Sequence Characteristics and Nucleotide Diversity

ITS region. Nucleotide diversity and distribution of substitutions in the ITS region of major and minor ribotypes are summarized in table 1. There were 29 different haplotypes within the 38 sequences of the major ribotype. The length of ITS1 ranged from 220 to 221 nucleotides; the lengths of ITS2 and the 5.8S gene were identical in all sequences of the major ribotype (217 and 164 nucleotides, respectively). Singleton variable sites clearly prevailed over parsimony-informative sites (39 and 9 in the entire region). Nucleotide diversities π of the complete ITS1-5.8S-ITS2 region and both spacers of the major ribotype were identical (0.007), nucleotide diversity of the 5.8S gene was 0.006. Among 29 sequences of minor ribotype, there were 25 different haplotypes. ITS1 in all sequences of the minor ribotype contained only 52 nucleotides because of restriction digestion (see Materials and Methods). The length of the 5.8S gene was identical in all sequences (164 nt), the length of ITS2 ranged from 215 to 217 nucleotides. Out of 102 polymorphic sites, 69 were singletons and 33 were parsimony-informative. Sequences of the minor ribotype were markedly more diverse than those of the major ribotype, with nucleotide diversity of the whole region being 0.027. Nucleotide diversity of the 5.8S gene ($\pi = 0.020$) was somewhat lower than those of both ITS1 (0.028, only 52 sites) and ITS2 (0.032). The mean distance (K2P) between sequences of the major and minor ribotype was 0.058.

5S region. We distinguished 8 different unit classes within the 89 5S sequences (table 2). Among the 89 sequences, there were 85 different haplotypes. Short and long units differed in the lengths of the spacer region, which varied from 283 to 331 nucleotides in the short units and from 349 to 371 nucleotides in the long units. The length of the gene region varied from 120 to 121 nucleotides in all sequences, except for one case (124 nucleotides). Like in the ITS region, singleton variable sites dominated over parsimony-informative sites. Nucleotide diversity π of the whole region of the short units ranged from 0.018 (short I) to 0.024 (short III). Nucleotide diversity of the gene region ranged from 0.014 (short II) to 0.031 (short III), and from 0.017 (short I) to 0.022 (short III) in the spacer. Average values were 0.020 for the whole region, 0.021 for the gene and 0.020 for the spacer. Average values for the gene and spacer regions did not differ significantly in the short (t test, $P = 0.861$) as well as long ($P = 0.411$) unit classes. Nucleotide diversity π of the whole region of the long units ranged from 0.017 (long V) to 0.031 (long III). Nucleotide diversity of the gene region ranged from 0.008 (long IV) to 0.030 (long II), and from 0.014 (long V) to 0.033 (long III) in the spacer. Average values were as follows: 0.024 (whole region), 0.020 (gene) and 0.025 (spacer). Mean

Table 1. Nucleotide Diversity and Distribution of Substitutions in the ITS Region of the Major and the Minor ITS Ribotypes of *Thinopyrum intermedium*.

		ITS1	5.8S	ITS2	Region
Major ribotype					
Number of sequences/haplotypes	38/29				
Number of sites		220–221	164	217	601–602
Number of polymorphic sites		21	11	16	48
Singletons/parsimony informative		16/5	10/1	13/3	39/9
Number of mutations		21	11	17	49
Nucleotide diversity π (K2-P)		0.007	0.006	0.007	0.007
Standard error of π		0.002	0.003	0.003	0.001
Minor ribotype (from alignment position 170)					
Number of sequences/haplotypes	29/25				
Number of sites		52	164	215–217	431–433
Number of polymorphic sites		10	34	58	102
Singletons/parsimony informative		6/4	28/6	35/23	69/33
Number of mutations		12	35	65	112
Nucleotide diversity π (K2-P)		0.028	0.020	0.032	0.027
Standard error of π		0.010	0.004	0.004	0.003

Table 2. Nucleotide Diversity and Distribution of Substitutions within the Different 5S Unit Classes of *Thinopyrum intermedium*.

Number of Sequences/Haplotypes	Short I 27/25			Short II 18/18			Short III 6/6			Long I 17/15		
	Gene	Spacer	Region	Gene	Spacer	Region	Gene	Spacer	Region	Gene	Spacer	Region
Number of sites	120–121	283–301	403–422	120–121	320–330	440–450	120–124	331	451–455	120–121	368–371	488–491
Number of polymorphic sites	18	42	60	11	45	56	10	21	31	9	49	58
Singletons/parsimony informative	14/4	33/9	47/13	9/2	40/5	49/7	8/2	20/1	28/3	6/3	27/22	33/25
Number of mutations	18	44	62	11	47	58	10	21	31	9	53	62
Nucleotide diversity π (p-distance)	0.018	0.017	0.018	0.014	0.021	0.019	0.031	0.022	0.024	0.013	0.029	0.025
Standard error of π	0.005	0.003	0.002	0.005	0.003	0.003	0.009	0.004	0.004	0.005	0.005	0.003
Number of Sequences/Haplotypes	**Long II 11/11**			**Long III 4/4**			**Long IV 4/4**			**Long V 2/2**		
	Gene	Spacer	Region	Gene	Spacer	Region	Gene	Spacer	Region	Gene	Spacer	Region
Number of sites	120–121	363–368	483–489	120	363,365	483,485	120–121	349,350	469–471	120	360,366	480,486
Number of polymorphic sites	17	45	62	6	22	28	2	14	16	3	5	8
Singletons/parsimony informative	16/1	39/6	55/7	6/0	16/6	22/6	2/0	13/1	15/1	3/0	5/0	8/0
Number of mutations	19	47	66	6	22	28	2	15	17	3	5	8
Nucleotide diversity π (p-distance)	0.030	0.027	0.028	0.025	0.033	0.031	0.008	0.021	0.018	0.025	0.014	0.017
Standard error of π	0.007	0.004	0.003	0.010	0.006	0.005	0.006	0.006	0.004	0.014	0.006	0.006

between-group distances are presented in table 3. Within the long units, long III and long V were the most similar with mean distance of 0.072. Long I and long II were the most diverged unit classes (0.131). Within the short units, short II and short III were the most similar pair (0.049), and short I and short III were the most diverged pair (0.099). We did not compare long units with short units, as we could not assure sequence homology of the NTS in the alignment.

To infer the rate of concerted evolution (i.e., to compare the levels of sequence homogenization within and among unit classes), we compared the values of within-group nucleotide diversities and between-group mean distances (tables 1–3). In this case, both characteristics are equivalent to the mean number of nucleotide differences per site among all pairs of respective sequences, that is, either within groups or among them. In ITS region, mean distance between sequences of major and minor ribotypes (0.058) exceeded within-group diversities (0.007 and 0.027 in major and minor ribotype, respectively). Similarly, between-group mean distances of the whole region of the short as well as long units exceeded within-group diversities of short and long units (0.049–0.099 and 0.072–0.131 vs. 0.018–0.024 and 0.017–0.031, respectively), suggesting that concerted evolution in intermediate wheatgrass works well within arrays (unit classes) but apparently not so well among them.

Inference of Functionality

Substitutions at Conserved Sites and GC Content

ITS Region. Among the sequences of the major ribotype, there were only two sequences—both of accession *T. intermedium-4*—not matching the five conserved motifs at all sites (supplementary table S1, Supplementary Material online). Clone *Thin4-05* contained a substitution in the motif M4 and clone *Thin4-11* in the motif M1. Sequences of the minor ribotype were more variable with nine substitutions distributed across the conserved motifs M1–M4 of the 5.8S gene and one substitution located within the M5 of the ITS2. Conserved motif M2 was the most variable one with five sequences not matching it perfectly. Clone *Thin2-11* contained substitutions within two motifs (M1 and M4). GC content within the 5.8S gene region was similar in both ribotypes, with mean values 59.4% and 59.1% in the major and minor ribotype, respectively. Within the ITS2 region, clones of major ribotype reached a higher mean value (63.2%) than those of the minor ribotype (61.7%).

5S Region. We encountered substitutions within all five motifs examined (supplementary table S2, Supplementary Material online). Regarding the internal control region (ICR), intermediate element (IE) was the most conserved motif with four substitutions found in three clones. Within the A-box, nine clones contained a single substitution. The least conserved (also the longest one) was the C-box, in which 34 clones contained one or two substitutions. Most frequently (in 21 clones), a G → A transition occurred at the sixth position of the motif (85. (86.) position within the gene). Otherwise, diverse substitutions occurred throughout the motif. Notably, in clone *Thin2-19*, a GGAT insertion occurred

Table 3. Mean Distances between Different 5S Unit Classes in *Thinopyrum intermedium*.

	Long I	Long III	Long V	Long II	Long IV	Short II	Short III	Short I
Long I		0.010	0.011	0.013	0.013	—	—	—
Long III	0.085		0.010	0.013	0.012	—	—	—
Long V	0.084	0.072		0.013	0.013	—	—	—
Long II	0.131	0.123	0.110		0.012	—	—	—
Long IV	0.116	0.102	0.108	0.094		—	—	—
Short II	—	—	—	—	—		0.008	0.012
Short III	—	—	—	—	—	0.049		0.013
Short I	—	—	—	—	—	0.089	0.099	

NOTE.—The number of base differences per site from averaging over all sequence pairs between groups are shown. Below diagonal—distances (*p*-distance, gaps treated with the pairwise deletion option); above diagonal—standard errors.

at position 2 of this motif. A C-residue, commonly found at position 1 upstream of the transcription-starting site (TSS), was present in all clones but four. In three cases, a cytosine was replaced by a thymine and in one case by an adenine. Finally, a TTTTT track at position + 121 downstream TSS was disrupted in three clones. A TATA-box with its typical motif –TATAAAA– (or its variation) was present in none of the clones.

Overall mean GC content was 49.3% in the NTS and 53.6% in the gene. Although the NTS of short units had a higher GC content than those of long units (51.0% vs. 47.1%), the values were identical in the gene (53.6%).

Secondary Structure Analyses

ITS Region. All 5.8S-RNA transcripts of the major ribotype but one were able to build up the constrained secondary structure (fig. 1*b*; supplementary table S1, Supplementary Material online). Free energy levels of the secondary structures ranged from −13.7 (also in the reference transcript) to −11.5 kcal mol^{-1} (mean value −12.6 kcal mol^{-1}). Among the 29 sequences of the minor ribotype, the inability of gene transcript to build the secondary structure was recorded in four cases. Free energy values of those transcripts enabling formation of proper secondary structure ranged from −13.7 (also recorded for the reference transcript) to −9.1 kcal mol^{-1} (mean value −13.1 kcal mol^{-1}). All the ITS2 transcripts of the major ribotype folded into the four-helix secondary structure (fig. 1*c*); however, there were eight cases in which the homology with the reference model was not perfect. In no case was the homology below 90%. Free energy levels of the secondary structures ranged from −80.8 to −71.6 kcal mol^{-1} (mean value −78.4 kcal mol^{-1}). ITS2 transcripts of the minor ribotype showed decreased homology overall with the reference model. In no case was the homology of all four helices perfect. In two cases, the homology of one helix of the ITS2 transcript with its reference counterpart was below 75% (clone *Thin2-13*, helix I and clone *Thin2-14*, helix IV).

5S Gene. When we applied rigid constrictions to comply with the secondary structure models of *Oryza* and *Triticum*, 44 transcripts out of 89 were able to build the requested structure (fig. 1*a*; supplementary table S2, Supplementary Material online). Among the rest, 39 transcripts did not conform the model structure because of one substitution. Most often, in 21 cases, it was an adenine at position 85 (86) that could not pair with a cytosine 92 (93) within the helix IV. Five transcripts did not conform to the structure because of two couples of unpaired bases and one transcript because of three couples of unpaired bases. Free energy levels of the secondary structures ranged from −38.1 to −30.9 kcal mol^{-1} (mean value −36.7 kcal mol^{-1}).

Without the application of constraints, 62 transcripts formed a three-domain Y-shaped secondary structure with five helices and loops. In these clones, mostly single substitutions did not have a deleterious effect on the folding of the molecule. The effect on its proper function is unclear, however. Twenty-seven transcripts did not fold into the requested structure. In eight cases, this was due to two Cs at positions 5 and 6, which occurred in 10 clones and which disrupted the structure of loop A, while retaining the Y-shaped form. In the other clones, transcripts did not form the three-domain, Y-shaped structure. Unconstrained secondary structures had lower ΔG values (ranging from −44.3 to −31.6 kcal mol^{-1}, mean −39.7 kcal mol^{-1}) than constrained structures, suggesting a higher stability of the molecules folded without the application of constraints.

Overall, 5S-RNA transcripts of 41 clones folded into both the constrained and unconstrained secondary structures and out of these, 33 clones also did not contain any substitution within the examined conserved motifs.

Relative Rate Tests

The two-cluster relative rate test (Takezaki et al. 1995) for ITS sequences revealed rate constancy between the major and the minor ribotype of *T. intermedium* in both regions tested. Results of the test were as follows: 5.8S: L_a (major ribotype) = 0.005459, L_b (minor ribotype) = 0.009534, and $L_a - L_b = -0.0040 \pm 0.0039$ ($Z = 1.039$); ITS2: L_a (major ribotype) = 0.0303, L_b (minor ribotype) = 0.0451, and $L_a - L_b = -0.0147 \pm 0.0178$ ($Z = 0.827$).

Pairwise relative rate tests of Tajima (1993) performed separately for sequences of major (38) and minor (29) ribotypes (5.8S and ITS2 separately) revealed an increased evolutionary rate in five cases. In the 5.8S region, we found two clones of one accession (*Thin2-11* and *-12*), and in ITS2 region three clones of two accessions (*Thin1-09* and *-13* and *Thin2-14*), all corresponding to the minor ribotype, to be potential

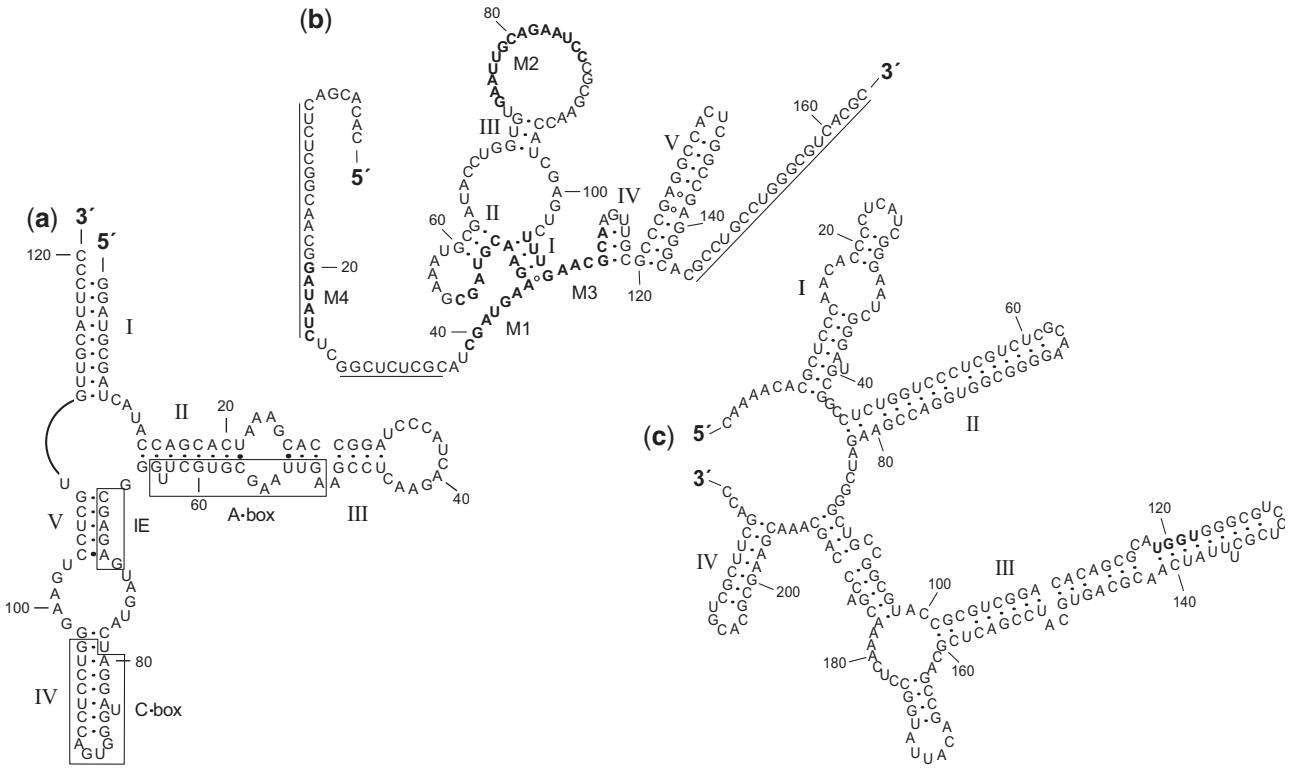

FIG. 1. Examples of functional secondary structures of the 5S and 5.8S rRNA genes and ITS2 of *Thinopyrum intermedium*. Every 20th nucleotide is numbered. Helices in (*a*–*c*) are numbered using Roman numerals. (*a*) 5S-RNA transcript of clone *Thin4-10* mapped on the three-domain secondary structure of the 5S rRNA gene of *Triticum monococcum* (for details, see Materials and Methods). Conserved regions of the A-box, the IE and the C-box are boxed. (*b*) 5.8S-RNA transcript of clone *Thin1-01* mapped on the secondary structure of the 5.8S rRNA gene of *Oryza sativa*. Four conserved motifs M1–M4 are in boldface. Lines depict sites of interaction with 26S. (*c*) Common four-helix secondary structure of ITS2, representing the major ITS ribotype of *T. intermedium* (clone *Thin1-01*). The common TGGT motif in helix III is in boldface.

pseudogenes (supplementary table S1, Supplementary Material online). Clones *Thin1-09* and *-13* are almost identical to each other and both are clearly different from the rest of sequences of the minor ribotype. We therefore suppose that their divergence is not caused by pseudogenization but accumulation of parsimony-informative sites. Out of the five earlier-mentioned cases, no sequence identified as a pseudogene based either on the 5.8S or the ITS region was also identified as a pseudogene based on the other region.

The two-cluster relative rate test for 5S rDNA sequences revealed rate constancy between all pairs of short-unit classes but not so in long-unit classes. The pairs long II/long IV in the gene region and long II/long III and long II/long IV in the NTS did not comply with the rate constancy (table 4). The results suggest that perhaps the sequences of long II accumulated a higher number of substitutions than clones of other long unit classes.

Phylogenetic Analyses

ITS Region. Final alignment of 622 nucleotides contained 163 parsimony-informative characters. Both phylogenetic analyses produced congruent trees as to the placement of our *T. intermedium* sequences within the Triticeae. Otherwise, there was no striking conflict in the grouping of Triticeae taxa; only values of clade robustness differed. The maximum parsimony (MP) analysis resulted in 800 most parsimonious trees with a length of 628 steps (consistency index [CI] = 0.600, retention index [RI] = 0.783). Herein, we present a Bayesian phylogenetic tree in which branches found in both Bayesian and MP strict consensus trees are indicated in bold lines (fig. 2). In the tree, homeologous ITS sequences of our *T. intermedium* accessions fell into two clades. Sequences representing the major ribotype fell into a large clade composed mostly of polyploid species *Thinopyrum* and *Trichopyrum*, all of which contain haplomes **E** or **St**, or their combination (supplementary table S3, Supplementary Material online). The only monogenomic diploid species in this group is *T. elongatum* (haplome $\mathbf{E^e}$). In this clade, all *T. intermedium* accessions were represented by two sequences, except for accession *T. intermedium-3* with only one sequence. The second clade in which our *T. intermedium* sequences appeared also comprised sequences of *D. villosum* and one sequence of *T. intermedium* from a previous study (Li et al. 2004). *Thinopyrum intermedium* sequences of this clade represent the minor ribotype of our accessions. Each accession is represented by one sequence, except for accession *T. intermedium-1*, which is represented by two sequences.

As to the overall phylogenetic relationships within the Triticeae, there is a remarkable polyphyly in *Pseudoroegneria* and *Thinopyrum*. There are at least two groups of *Pseudoroegneria* sequences—*spicata*, *libanotica*, *strigosa* ssp. *aegilopoides*, and *tauri* in one, and *stipifolia*, *cognata*, *gracillima*, and *strigosa* in the other. Besides, there are sequences of

Table 4. Two-Cluster Relative-Rate Tests for Evolutionary Rate Constancy of 5S rDNA Unit Classes in *Thinopyrum intermedium*.

	Cluster[a]		L_a	L_b	δ^b	Z Score
	A	B				
Short units, gene	Short I	Short II	0.0092	0.0075	0.0017	0.5570
	Short I	Short III	0.0089	0.0160	−0.0071	1.4087
	Short II	Short III	0.0071	0.0159	−0.0088	1.7063
Short units, NTS	Short I	Short II	0.0067	0.0513	0.0152	1.1867
	Short I	Short III	0.0682	0.0610	0.0072	0.5338
	Short II	Short III	0.0253	0.0332	−0.0080	0.7893
Long units, gene	Long I	Long II	0.0070	0.0148	−0.0078	1.8979
	Long I	Long III	0.0070	0.0121	−0.0051	0.9709
	Long I	Long IV	0.0070	0.0038	0.0032	0.9660
	Long I	Long V	0.0066	0.0118	−0.0051	0.7428
	Long II	Long III	0.0147	0.0120	0.0027	0.4729
	Long II	Long IV	0.0147	0.0037	0.0110	3.0319*
	Long II	Long V	0.0144	0.0117	0.0027	0.3598
	Long III	Long IV	0.0120	0.0036	0.0050	1.6515
	Long III	Long V	0.0115	0.0115	0.0000	—
	Long IV	Long V	0.0031	0.0115	−0.0083	1.2734
Long units, NTS	Long I	Long II	0.0731	0.0936	−0.0205	1.0716
	Long I	Long III	0.0614	0.0450	0.0164	1.0997
	Long I	Long IV	0.0817	0.0705	0.0112	0.6124
	Long I	Long V	0.0555	0.0506	0.0049	0.3125
	Long II	Long III	0.0958	0.0588	0.0370	2.0509*
	Long II	Long IV	0.07593	0.0442	0.0317	2.2366*
	Long II	Long V	0.0818	0.0565	0.0254	1.5379
	Long III	Long IV	0.0634	0.0687	−0.0052	0.3197
	Long III	Long V	0.0383	0.0499	−0.0116	0.8413
	Long IV	Long V	0.0669	0.0733	−0.0064	0.3650

[a]The two-cluster test compares the averages of observed numbers of substitutions per site (branch lengths, L_a and L_b) from the common ancestor (outgroup) of clusters A and B. For tests of short and long units, *Psathyrostachys juncea* and *Hordeum brevisubulatum* ssp. *violaceum* were used as the outgroups.

[b]$\delta = L_a - L_b$.

*Significant at the 5% level.

related polyploids in both groups. The second genus with a conflicting placement within the Triticeae is *Thinopyrum*. In particular, *T. elongatum* sequences were polyphyletic in the tree, the sequences of accessions 1 and 2 being placed in different part of the tree than sequences of accession 3 (sequences 3a and 3b). Furthermore, the sequence of the closely related *T. bessarabicum* did not group together with either group of *T. elongatum* sequences. Otherwise in the tree, there were well-defined monophyletic groups of *Hordeum*, *Psathyrostachys*, *Agropyron*/*Australopyrum*, and *Secale* but a nonmonophyletic group comprising *Aegilops* s. l. and *Triticum*.

5S Region. Long units—the final alignment of 43 sequences contained 607 sites of which 229 were parsimony-informative. The MP analysis resulted in 900 most parsimonious trees with a length of 767 steps (CI = 0.701, RI = 0.760). As in the ITS tree herein, we present the Bayesian tree with branches found in both Bayesian and MP consensus trees displayed in bold (fig. 3). Both the BI and MP analyses yielded similar topologies within the gene tree, with two remarkable differences. First, in the MP consensus tree, the long A1 unit class of *Triticum* formed a clade with *Heteranthelium piliferum*, and the clade was sister to the whole large *Aegilops* + *Triticum* + *T. elongatum* assemblage (bootstrap support 1.00/81). However, the *Triticum* A1 + *H. piliferum* clade had less than 50% bootstrap support in the MP tree. Second, in the MP tree, *Agr. cristatum* + *A. mongolicum* clade was not sister to *Pseudoroegneria* + *T. intermedium* long V as in the BI tree but appeared together within the unresolved polytomy with *Triticum* A1, *Australopyrum*, or *Heteranthelium*. In both BI and MP trees, neither of the genera *Aegilops*, *Triticum*, and *Pseudoroegneria* was monophyletic. The placement of consensus sequences representing the different unit classes of *T. intermedium* accessions was congruent in both kinds of analyses. The consensus sequence of *T. intermedium* long I fell within the *Aegilops* + *Triticum* + *T. elongatum* assemblage, of which *A. tauschii* and *Triticum* D1 were the most closely related sequences to *T. intermedium* long I. *T. intermedium* long II and IV grouped with *D. villosum* sequences; however, they did not group tightly with any particular sequence. The last two consensus sequences of *T. intermedium* grouped with *Pseudoroegneria*. Long III was sister to the *P. tauri* + *P. tauri* ssp. *libanotica* clade, and long V grouped with *P. strigosa*, which was sister to the rest of *Pseudoroegneria* sequences.

Short Units. The alignment of 36 sequences contained 523 characters of which 132 were parsimony-informative. Both Bayesian and MP analyses resulted in virtually identical tree topologies (fig. 4) with the following minor differences. First, in the MP consensus tree *H. piliferum* was sister to the *P. tauri* + *P. tauri* ssp. *libanotica* clade, but with low bootstrap

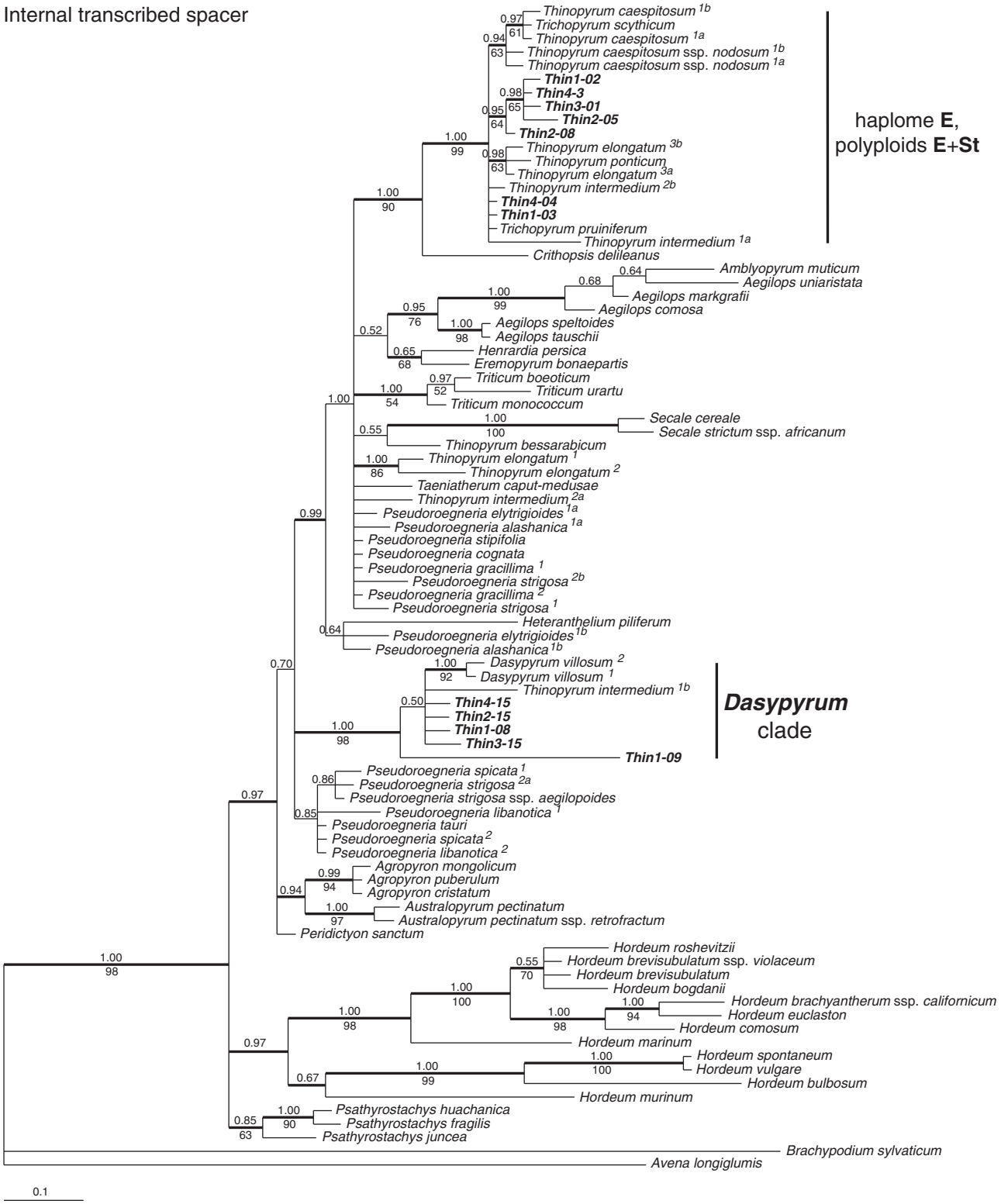

Fig. 2. Bayesian phylogenetic tree based on ITS sequences. Branches found in both Bayesian- and MP-strict consensus trees are indicated by bold lines. *Thinopyrum intermedium* sequences amplified in this study are in bold. Clone designations refer to individual plants analyzed and individual clones of each plant (after dashes). The numbers above and below the branches are Bayesian posterior probabilities and bootstrap values for MP, respectively.

support (<50%). Second, in the MP consensus tree, the *Secale vavilovii* + *S. strictum* ssp. *africanum* clade was basal to the rest of the ingroup taxa (if we omit *Psathyrostachys*, which grouped with the outgroup *Hordeum brevisubulatum* ssp. *violaceum*). As to the overall pattern within the Triticeae, *Triticum* + *Aegilops* was monophyletic. *Pseudoroegneria* was monophyletic, but with low support (0.54 in the Bayesian tree and <50% in the MP). Within the poorly supported *Pseudoroegneria* clade, *P. tauri* + *P. tauri* ssp. *libanotica* was sister to the rest of *Pseudoroegneria* taxa. The three consensus

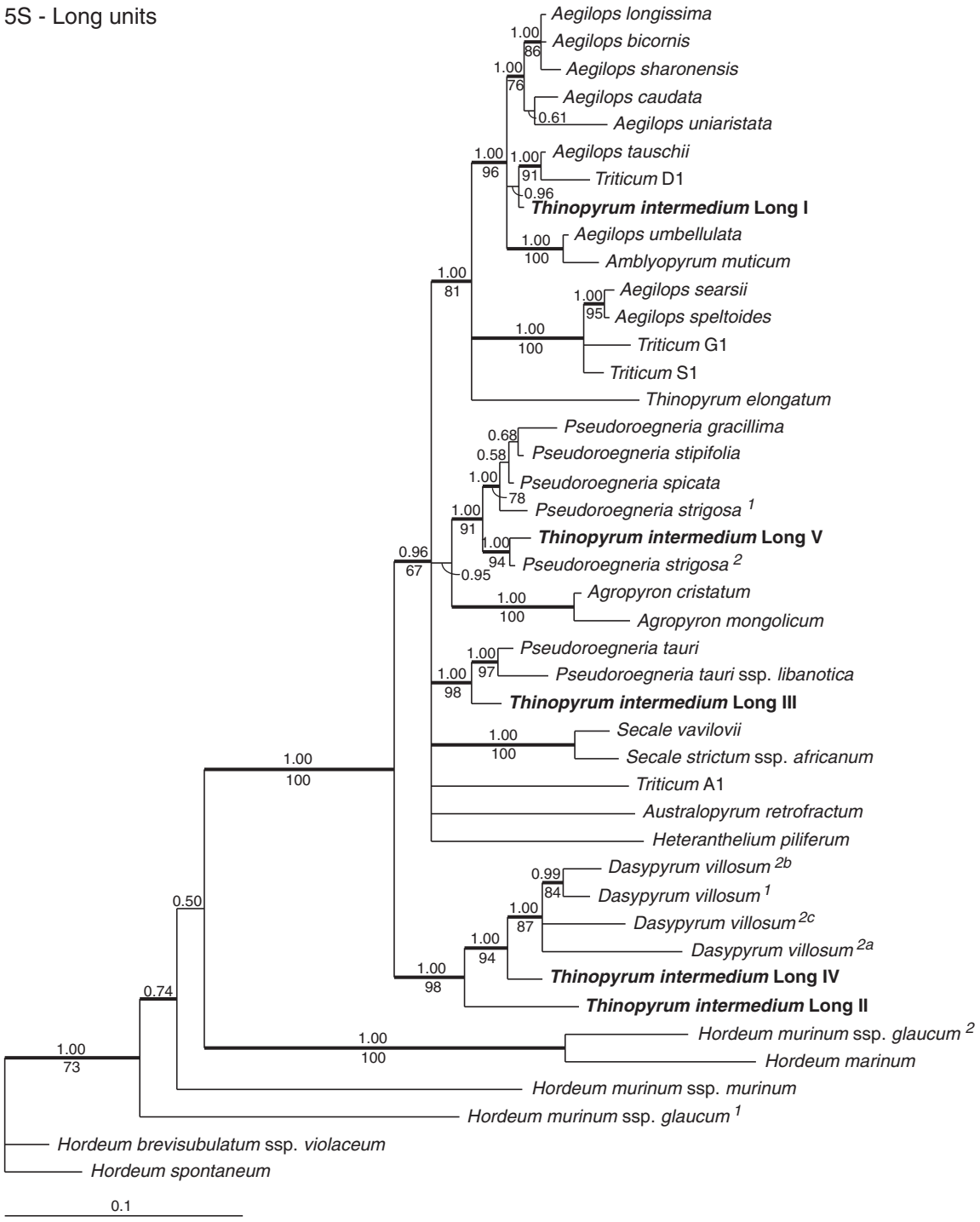

Fig. 3. Bayesian phylogenetic tree based on long-spacer 5S sequences. Branches found in both Bayesian and MP strict consensus trees are indicated by bold lines. Most of the sequences are consensus sequences (see Materials and Methods). Consensus sequences of different unit classes amplified in four *T. intermedium* accessions are in bold. The numbers above and below branches are Bayesian posterior probabilities and bootstrap values for MP, respectively.

sequences of *T. intermedium* fell into two clades. First, *T. intermedium* short I fell into the strongly supported *Aegilops* clade. Within this clade, *T. intermedium* short I is most similar to *A. searsii*, *A. tauschii*, and *A. uniaristata* sequences. Second, both *T. intermedium* short II and III grouped with *Pseudoroegneria*. *Thinopyrum intermedium* short II formed a clade with *P. stipifolia*, *strigosa*, *spicata*, and *gracillima*, whereas *T. intermedium* short III was sister to this clade.

In Situ Hybridization

GISH with each of the genomic DNA of *P. spicata*, *D. villosum*, and *A. tauschii* produced signal on separate 14 chromosomes

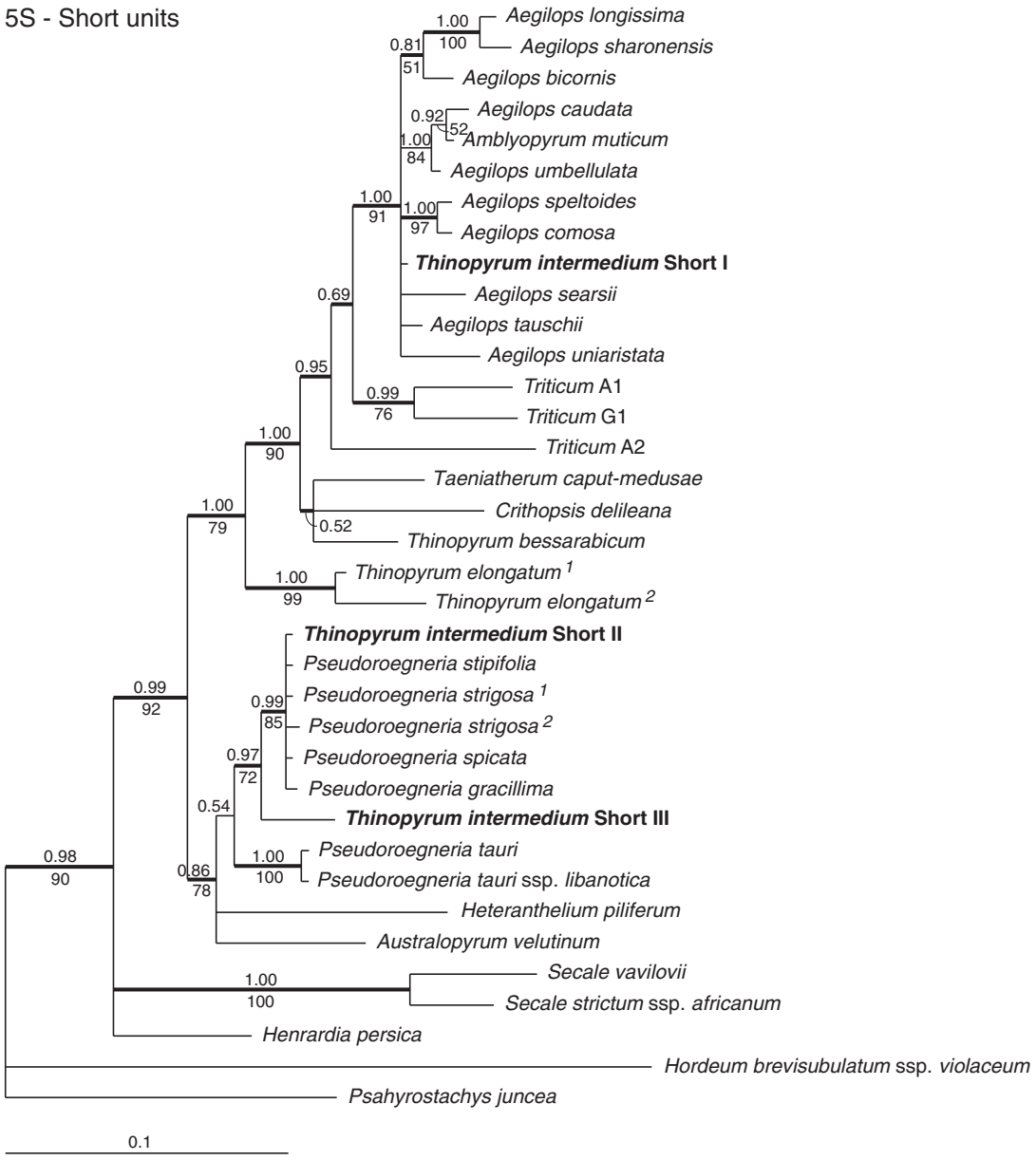

Fig. 4. Bayesian phylogenetic tree based on short-spacer 5S sequences. Branches found in both Bayesian and MP strict consensus trees are indicated by bold lines. Most of the sequences are consensus sequences (see Materials and Methods). Consensus sequences of different unit classes amplified in four *T. intermedium* accessions are in bold. The numbers above and below the branches are Bayesian posterior probabilities and bootstrap values for MP, respectively.

representing individual subgenomes of *T. intermedium* (fig. 5). FISH with 45S rDNA and 5S rDNA probes exhibited variation in the number of these loci among the accessions analyzed (fig. 5 and table 5). Four chromosomes with both 45S and 5S rDNA loci belonging to the *Aegilops* subgenome were found in all but one accession (*T. intermedium-2*, in which a 45S rDNA locus in the subtelomeric region was missing from one chromosome). Furthermore, accession *T. intermedium-1* carried an additional 5S rDNA locus. Two chromosomes of *Dasypyrum*-like subgenome carried a 5S rDNA locus in all accessions. However, we detected 45S rDNA locus on two (*T. intermedium-1* and -3), one (*T. intermedium-4*), or none (*T. intermedium-2*) of the chromosome(s) belonging to the *Dasypyrum* subgenome. Both 5S rDNA and 45S rDNA loci were found on one (*T. intermedium-1* and -3) or two (*T. intermedium-2* and -4) chromosomes (presumably homologous) of the *Pseudoroegneria* subgenome. Additional loci of 5S rDNA were detected on two chromosomes belonging to the *Pseudoroegneria* subgenome in all accessions.

Discussion

Genomic Organization of 45S and 5S rDNA Families in the Allohexaploid *T. intermedium*

Both 45S and 5S rDNA families in our *T. intermedium* accessions are organized in separate loci on a varying number of chromosomes. The number of 45S loci in our accessions varied between five in accession 2 and seven in

Table 5. Numbers of rDNA Loci Residing on Chromosomes Belonging to Different Subgenomes as Detected in Four Accessions of Allohexaploid *Thinopyrum intermedium*.

Subgenome	*Pseudoroegneria*		*Dasypyrum*		*Aegilops*	
Accession	45S	5S	45S	5S	45S	5S
T. intermedium-1	•	•••	••	••	••••	•••••
T. intermedium-2	••	••••	–	••	•••	••••
T. intermedium-3	•	•••	••	••	••••	••••
T. intermedium-4	••	••••	•	••	••••	••••

Note.—Bullets indicate the number of rDNA loci.

the others (table 5). One or two loci belonged to the *Psudoroegneria*-like chromosomes, three or four loci to the *Aegilops*-like chromosomes, whereas *Dasypyrum*-like chromosomes contained one, two, or no loci (fig. 5). A majority of loci were located in terminal or subterminal regions of chromosomes. Only those loci belonging to *Dasypyrum* chromosomes were located more interstitially (fig. 5). A similar pattern was observed by Li et al. (2004), who in three *T. intermedium* accessions detected 6 or 8 major loci within the 42 chromosomes, with two or three pairs located in terminal parts. One pair of loci was located interstitially. Additionally, depending on stringency conditions, they observed a high polymorphism in the number of minor loci. We performed in situ hybridization (FISH and GISH) under conditions of 77% stringency. While such conditions most likely ensured the capture of all major loci (in terms of relative size, regardless of their nucleolus organizing activity), they may be too stringent to detect additional minor loci (Mahelka and Kopecký 2010). In accession 2, we did not detect any 45S locus residing in the *Dasypyrum* subgenome. We detected, however, *Dasypyrum*-like ITS sequences in this accession. This suggests that the number of detected loci is rather an underestimation of the real state, unless we consider interlocus recombination between homeologous chromosomes.

Theoretically, the number of rDNA loci in an allopolyploid should equal the sum of loci of its progenitors. However, rDNA loci in allopolyploid species often experience a dynamic evolution, and contributions from progenitor species are often unpredictable (Baum and Feldman 2010; Malinská et al. 2010). Following polyploid species formation, polyploids may exhibit complete or nearly complete additivity of rDNA loci with respect to their progenitors (e.g., *T. ponticum*—Li and Zhang 2002; Brasileiro-Vidal et al. 2003; *Nicotiana*—Kovařík et al. 2004), as well as loss of some loci (*Zingeria*—Kotseruba et al. 2010). Rarely, the number of rDNA loci may increase (*Triticum*—Jiang and Gill 1994). The organization of rDNA loci in Triticeae diploids suggests that elimination of some loci during the intermediate wheatgrass' history had likely occurred. Diploid Triticeae, including those likely involved in the formation of *T. intermedium* (except for *Dasypyrum*, i.e., species from the genera *Pseudoroegneria*, *Aegilops* and *Thinopyrum*—see Mahelka et al. 2011), usually carry two (or more) major rDNA loci per haploid genome plus a variable number of minor loci (Dubcovsky and Dvořák 1995; Li and Zhang 2002). In diploid *D. villosum*, a single 45S rDNA locus per haploid genome was localized on the short arm of chromosome 1V (Galasso et al. 1997; Uslu et al. 1999). These data led us to the conclusion that in our accessions, elimination of 45S loci mostly occurred in the *Pseudoroegneria* and *Dasypyrum* subgenomes.

In contrast to the 45S loci, it seems that the 5S rDNA in our accessions was less heavily, if at all, affected by locus loss (table 5). Our observations as to the number of 5S rDNA loci are consistent with other studies on the Triticeae. In a majority of the Triticeae, major loci have been found on chromosomes 1 and 5 (Sastri et al. 1992). Although the long units are located on chromosome 5 in the Triticeae, short units are located on chromosome 1. Furthermore, some species lack long and/or short units, and this feature is likely to be linked with the loss of respective loci. For example, *Agropyron* species, *D. villosum* and *A. speltoides* lack short-spacer units (Dvořák et al. 1989; Kellogg and Appels 1995; Baum et al. 2008). In *D. villosum*, Galasso et al. (1997) localized two 5S sites in close proximity to each other at a subterminal site on the short arm of a chromosome pair different from that carrying 45S loci. Uslu et al. (1999) only localized a single 5S site on the short arm of 5V in a haploid chromosomal set. Thus, consistent with our results, a complete, diploid chromosomal set of *D. villosum* contained two (or four) 5S loci located on different chromosomes. Thus, in our accessions analyzed, it seems that only a single 5S locus was lost (or remained undetected) from the *Pseudoroegneria* subgenome in two accessions (1 and 3), and a single locus was lost (or remained undetected) from the *Aegilops* subgenome in accession 2.

In the *Pseudoroegneria* and *Aegilops* subgenomes, the 45S loci always co-localized on the same chromosomes with the 5S loci, and the 45S loci were always located distally in relation to the 5S loci. Besides, these subgenomes contained extra 5S loci on separate chromosomes. Such a pattern may explain different rates of interlocus homogenization of 45S and 5S arrays (discussed later). In decaploid *T. ponticum*, the same relative position of 45S and 5S loci was observed (Brasileiro-Vidal et al. 2003).

Sequence Polymorphism and Functionality Tests

ITS Region. Sequences of the ITS region of our accessions were largely homogenized toward one parental ribotype, suggesting effective concerted evolution in the 45S rDNA family. In another study, Li et al. (2004) also detected two distinct sequence types in the allohexaploid *T. intermedium*, providing further support for our results. Nucleotide diversity (π) of sequences of the minor ribotype was approximately three to four-times higher than in the major ribotype (table 1). Such a ratio is similar to the ratio between the nucleotide diversity value of pseudogenes and functional genes in *Mammillaria* (Harpke and Peterson 2006). However, our values of π of the entire ITS region of major ribotype are still eight times lower than the values of functional genes in *Mammillaria* and two times lower in the case of the minor ribotype. The question arises whether both ribotypes represent samples of a functional rDNA family. The inference of rRNA gene functionality is a complex issue, and it is less straightforward compared with protein-coding genes. It has

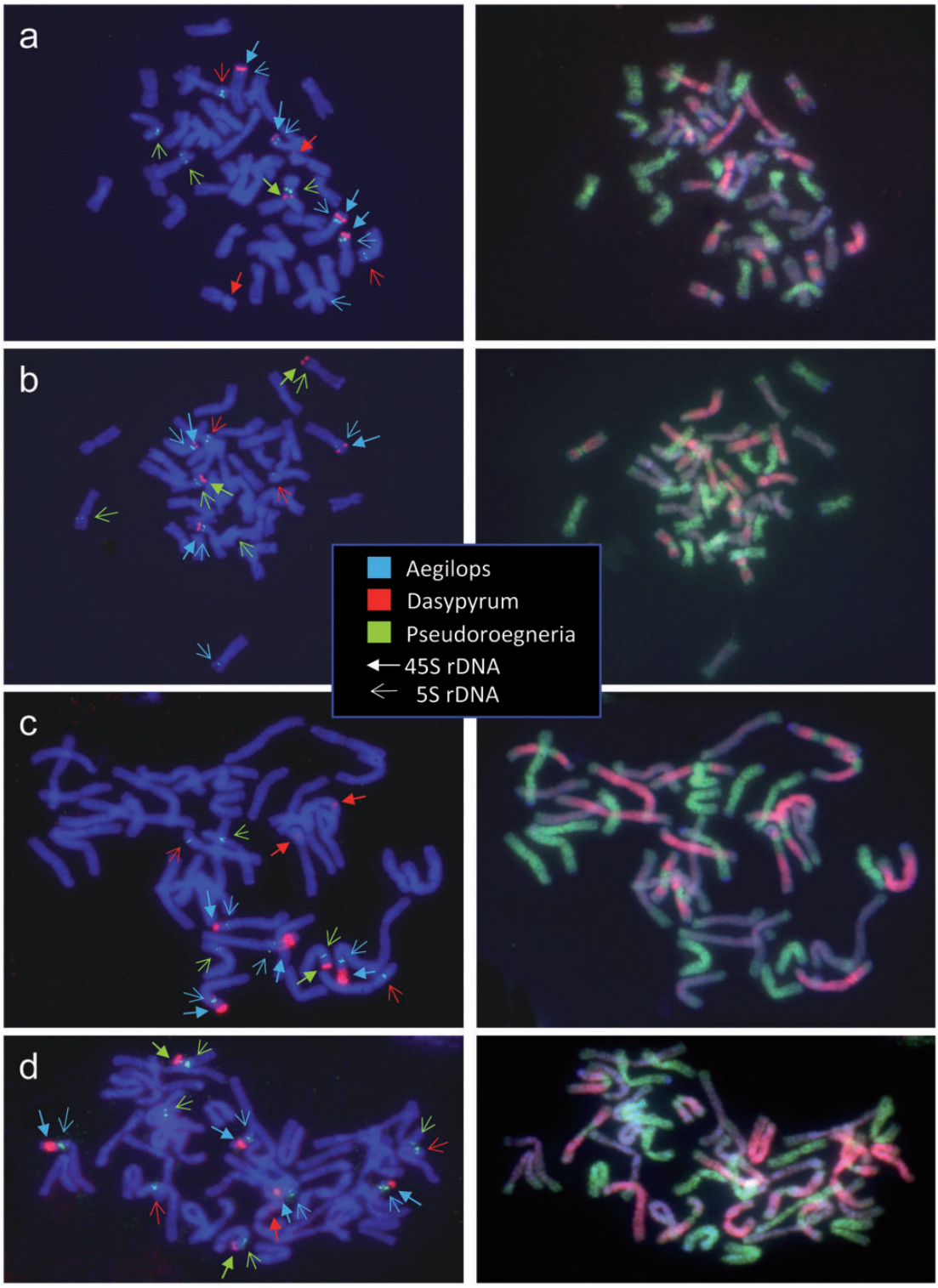

Fig. 5. Molecular cytogenetic analysis of *Thinopyrum intermedium*. FISH with 5S and 45S rDNA probes (left side) and genomic in situ hybridization (right side) of *T. intermedium*-1–4 (*a*–*d*). Fluorescent signals of 45S-rDNA labeled with biotin (red color) and 5S-rDNA labeled with digoxigenin (green color). After the microscopic evaluation of FISH, reprobing of slides was performed using total genomic DNA of *Pseudoroegneria spicata* labeled with digoxigenin (green color), total genomic DNA of *Dasypyrum villosum* labeled with biotin (red color) and sheared total genomic DNA of *Aegilops tauschii* used as blocking DNA (blue color of the DAPI counterstain). FISH with 45S rDNA and 5S rDNA probes showed variation in the number of these loci among the four accessions. Four chromosomes with both 45S and 5S rDNA loci belonging to the *Aegilops* subgenome (indicated by blue arrows) were found in all accessions but one (*b*), where the 45S rDNA locus in the subtelomeric region is missing from one chromosome. Two chromosomes of the *Dasypyrum* subgenome carry a 5S rDNA locus (red arrows) in all accessions. However, we detected a 45S rDNA locus on two (*a*, *c*), one (*d*) or none (*b*) of the chromosome(s) belonging to the *Dasypyrum* subgenome. Both 5S rDNA and 45S rDNA loci were found on one (*a*, *c*) or two (presumably homologous) chromosomes (*b*, *d*) of the *Pseudoroegneria* subgenome (green arrows). Additional loci of 5S rDNA were detected on two chromosomes belonging to the *Pseudoroegneria* subgenome in all accessions.

been postulated that silenced genes are less vulnerable to sequence homogenization and thus accumulate mutations more often, whereas the active units are homogenized more readily whereby eliminating substitutions (Kovařík et al. 2008). Harpke and Peterson (2008) demonstrated that the inference of pseudogenes based on the substitutions within the 5.8S conserved motifs yielded similar results as other tests such as relative rate and bootstrap hypothesis tests. We based our inference of functionality mainly on substitution patterns, since substitutions at conserved motifs, the ability to build secondary structures, free energy values, and relative rate tests are interrelated (supplementary table S1, Supplementary Material online). These data suggest that a majority of clones of the major ribotype represent functional copies. By contrast, almost one-third of clones of the minor ribotype showed violations of functionality rules in one of the characteristics examined, suggesting some degeneration toward nonfunctionality. It seems, however, that most of the minor-ribotype clones likely represent a functional rDNA family. Based upon these results, we conclude that concerted evolution operating on the major-ribotype arrays is strong when homogenizing copies effectively but rather relaxed in the cases of minor ribotype arrays whose copies show a certain degree of degeneration.

5S Region. 5S sequences comprised a higher diversity than sequences of the ITS region. Among the 89 clones sequenced, we distinguished eight different sequence types (unit classes). The presence of two distinct array types—the long- and the short-spacer units—is consistent with other studies on Triticeae (Sastri et al. 1992; Kellogg and Appels 1995) and other plant groups (Fulneček et al. 2002). The biological or evolutionary significance of this phenomenon is unknown, but its presence in various organismal groups has been explained as a possible remnant of gene duplication (Pinhal et al. 2011). Out of the 89 sequences of four accessions, there were 85 different haplotypes, which means that there were just a few identical clones (table 2). The polymorphism in 5S sequences found in our accessions fits well within the range of nucleotide diversities found in the Triticeae (Kellogg and Appels 1995). In the work of Kellogg and Appels (1995), the spacer was more diverse than the gene only in short units if gaps were included in calculations, otherwise the gene and the spacer were similarly diverse. In our accessions, the diversities of the gene region and the spacer did not differ significantly.

Between-group mean distances of the whole region of the short and long units exceeded within-group diversities, suggesting that concerted evolution in intermediate wheatgrass works well within arrays. It, however, does not work so well between them, apparently. In other species, for example, in *Gossypium*, intraindividual polymorphisms in diploid species (with a single array in each) in some cases exceeded interspecific comparisons (Cronn et al. 1996). However, with increasing evolutionary divergence of species, fixed interspecific differences (namely in the spacer) were increasingly overwhelming polymorphisms within arrays (Cronn et al. 1996). In the Triticeae, the amount of variation among species was found to be equal to or greater than that of within individuals only in the spacer region (Kellogg and Appels 1995).

In an attempt to decide whether the clones of different unit classes represent functional rDNA arrays, we characterized the data set using various criteria. Based upon the examination of all 89 clones (supplementary table S2, Supplementary Material online), we concluded that within each unit class, there were functional as well as pseudogenized clones. Unlike in the conserved motifs in the 5.8S gene, the functionality of conserved motifs linked with 5S RNA gene is relatively well understood. The ICR with the three boxes A, IE and C is considered to play a key role in the transcription of the gene through the interaction with the transcription factors of Pol III (Pieler et al. 1987; Orioli et al. 2012). The presence and conservation within the described motifs is therefore used to discriminate functional genes from potential pseudogenes (Sajdak et al. 1998; Garcia et al. 2009). Within the conserved motifs of our clones, the C-box was the least conserved motif in which 34 clones contained one or two substitutions, when compared with the consensus sequence computed across all 89 clones. Interestingly, the conservation of the C-box was recently questioned, because it was found that even genes severely deviating in the C-box sequence from the plant consensus were expressed (Garcia et al. 2012). It is thus possible, that some of the substitutions in the C-box found in our clones do not have any deleterious effect on the gene functionality.

Pol III does not necessarily require upstream promoters, such as the TATA-box, which is usually located 25–30 bp upstream of the TSS. In some species, the TATA-box was present, however (Garcia et al. 2009; Vizoso et al. 2011). The TATA box with its conserved sequence 5′ TATAAAA 3′ was present in none of the 89 clones examined in this study. Instead, we identified a highly conserved motif 5′ GG CAAGCATAAGGG 3′ at the 3′ end of the NTS, with the part CATAAGG starting at position −30 (according to clone *Thin1-19*) and being completely free of substitutions, suggesting that this region could be somehow involved in the 5S RNA gene regulation. The presence of such additional conserved sequence motifs, apart from the TATA-box, is not unusual (Orioli et al. 2012).

Regardless of the functionality tests, it is likely that the proportion of pseudogenes in our clones is underestimated, because we only sampled a fraction of repetitive rDNA copies that lacked substitutions within the primer sites, which targeted the coding region. Furthermore, all clones must have retained the BamHI site necessary for cloning of the PCR product (Appels et al. 1992).

45S and 5S rDNA Families as Phylogenetic Tools Reflect Different Patterns of rDNA Evolution at the Genome Level

We assume that the different ITS and 5S ribotypes (unit classes) detected in our accessions represent homeologous copies contributed by the progenitor species that were secondarily combined in the allohexaploid genome. With regard to the 5S rDNA data, based upon our summarized results coupled with

available data on the Triticeae, we suggest that it is possible to infer the origin of different unit classes. We suppose that short I + long I unit classes were donated by *Aegilops* (close to the D genome of *A. tauschii*), short II + long V by *Pseudoroegneria* (close to *P. strigosa* and allied species), short III + long III by *Pseudoroegneria* (perhaps different from *P. strigosa*) and long II + long IV unit classes by *Dasypyrum*. The long V unit class is typical of accession *T. intermedium-4*. As sequences of this unit class were rarely amplified even in this accession, we assume that their absence in other accessions is caused by PCR bias rather than by their true absence from the genomes. Unlike in 45S, the evolution of 5S rDNA according to the birth-and-death model of evolution has been suggested (Pinhal et al. 2011; Vizoso et al. 2011). If 5S sequences of individual unit classes really represent orthologous loci inherited from respective progenitors, which is the most parsimonious explanation, then the within-species clustering pattern is consistent with concerted evolution rather than the birth-and-death model of evolution (Rooney and Ward 2005; Pinhal et al. 2011). The pattern described earlier is consistent with the presence of three distinct subgenomes of the allohexaploid genome, only it would imply the contributions from two different *Pseudoroegneria* species (the presence of paralogs is less likely due to the presence of both long and short units of *Pseudoroegneria* origin). On the contrary, the presence of two different long units corresponding to *Dasypyrum* may not necessarily imply the contributions from different *Dasypyrum* species, as we detected in diploid *D. villosum* three distinct sequence types (designated *2a*–*2c* in fig. 2) among 10 clones sequenced, suggesting the presence of some variation in this diploid. Recently, a surprising diversity of GBSSI copies corresponding to a *Dasypyrum*-like progenitor was revealed in the same accessions of *T. intermedium* (Mahelka et al. 2011). In this respect, further examination of *Dasypyrum* accessions as to the variation of 5S sequences would be desirable.

In the case of ITS sequences, the major ribotype corresponded in phylogenetic analyses with the **E** genome of *T. elongatum* and also with the ribotypes of **ESt** polyploid grasses. Because of the polyphyly in the ITS tree of the **E**-genome diploids, the proposed origin of the major ribotype should be regarded as provisional, and further studies are required. Sequences of the minor ribotype corresponded with *Dasypyrum*. *Dasypyrum*-like copies likely remained conserved within the *Dasypyrum* subgenome at rDNA loci more distal from telomeres. These *Dasypyrum*-like rDNA arrays possibly represent epigenetically silenced, inactive rDNA arrays that escaped selection and thus accumulate mutations (Kovařík et al. 2008). In the work of Li et al. (2004), the authors concluded that the two ribotypes of allohexaploid *T. intermedium* corresponded with *Pseudoroegneria* and the **E**-genome diploid species. We disagree with the conclusions of Li et al. (2004) because their phylogenetic analysis was not conclusive, as they only included *Pseudoroegneria* and *Thinopyrum* (**E** genome) diploids in the analysis. We included respective sequences (1a and 1b, see supplementary table S3, Supplementary Material online) in our analyses, and the two different ribotypes clearly corresponded to our results: one grouped with the **E**-genome species, and the other fell into the *Dasypyrum* clade.

It has been repeatedly shown that the 5S rDNA region may provide a more suitable marker for reconstructing histories of allopolyploid species than ITS. This is mainly due to the different behavior of both families following allopolyploid species formation and during their evolutionary histories, when 45S rDNA often undergoes concerted evolution, whereas 5S rDNA is less vulnerable to this phenomenon (Wendel et al. 1995; Cronn et al. 1996; Fulneček et al. 2002; Kovařík et al. 2004; Baum and Feldman 2010). It is therefore reasonable to expect that 5S rDNA better reflects genome contributions from parental species to the allohexaploid *T. intermedium*. If this is correct, then the three subgenomes were donated by *Pseudoroegneria*, *Dasypyrum*, and *Aegilops*. Taking this scenario into account, the pattern of 45S rDNA suggests strong concerted evolution of 45S rDNA arrays operating within the *Pseudoroegneria* and *Aegilops* subgenomes. Because of the absence of *Aegilops*-like ITS sequences in *T. intermedium*, we suggest that 45S rDNA arrays residing on the *Aegilops*-like subgenome had converged to one of the ribotypes, perhaps the **E**-like ribotype, whereas 5S arrays retained *Aegilops*-like copies. Such a pattern is not exceptional in plants and has, for example, been observed in *Gossypium* (Wendel et al. 1995; Cronn et al. 1996), *Nicotiana* (Fulneček et al. 2002; Kovařík et al. 2004), or *Byblis* (Fukushima et al. 2011), in which 45S families evolved in a concerted manner through interlocus homogenization, whereas no interlocus homogenization in 5S sequences occurred. Possible explanations of this phenomenon have been described elsewhere (Cronn et al. 1996; Fulneček et al. 2002; Kovařík et al. 2004). Our data only suggest that the relative positions of 45S and 5S loci on *T. intermedium* chromosomes may play a role; 5S loci were always located more proximally.

On the Origin of Allohexaploid *T. intermedium*: A Two-Step Gene Conversion of 45S Arrays?

The genome constitution of the allohexaploid intermediate wheatgrass has not yet been satisfactorily resolved (Mahelka et al. 2011, Tang et al. 2011). It is obviously a complex species with surprising genomic diversity as to the amount of genetic material secondarily combined in one genome. A recent cytogenetic study showed that intermediate wheatgrass contains three distinct subgenomes, two of which have likely been contributed by *Pseudoroegneria* (which also represents the maternal lineage) and *Dasypyrum*. The identity of the third subgenome remained unclear, with possible contributions from *Aegilops* and diploid *Thinopyrum* (Mahelka et al. 2011). Other studies have suggested the contribution from *Dasypyrum* to the origin of intermediate wheatgrass (Kishii et al. 2005; Liu et al. 2009). Furthermore, *T. intermedium* has a close relationship with *Secale cereale* (Kishii et al. 2005; Tang et al. 2011). The results of this study provide further support for the contributions from *Dasypyrum* and *Aegilops*. Based upon the average nrITS substitution rate of 4.13×10^{-9} in plants (Kay et al. 2006), we may roughly date the divergence time of *Dasypyrum*-like ITS sequences

(ITS2 only in presumably functional clones) amplified in *T. intermedium* accessions and those downloaded from GeneBank (sequences 1 and 2 in supplementary table S3, Supplementary Material online) to be approximately 2 Ma ($T = 1/2\ [0.016/4.13] \times 10^9 = 1{,}937{,}046$ years). The contribution from *Aegilops* to the origin of intermediate wheatgrass provides new and interesting aspect namely with respect to its utility in wheat improvement. The potential of intermediate wheatgrass lies in the transfer of desirable traits such as resistance to various diseases and pests or drought tolerance into the wheat genome (Li and Wang 2009). In this respect, natural populations of intermediate wheatgrass represent an invaluable source of genetic material that is potentially useful in crop improvement.

Liu and Wang (1993) suggested that allohexaploid *T. intermedium* probably originated from one of these tetraploids: *Elytrigia caespitosa*, *Lophopyrum nodosum*, and *P. geniculata* ssp. *scythica*. These taxa appear under the names *T. caespitosum*, *T. caespitosum* ssp. *nodosum*, and *Tri. scythicum* in supplementary table S3, Supplementary Material online, and figure 2, and they all have basically the same genome formulas combining the haplomes **E** (*T. elongatum*) and **St** (*Pseudoroegneria*) (Liu and Wang 1993; Yu et al. 2010). This is a likely scenario because all these allopolyploids share almost identical ITS sequences (the **E**-type in fig. 2), suggesting that they evolved from the same precursor. Apparently, the **E**-genome ribotype already dominated in these allotetraploids, suggesting effective concerted evolution. The origin of *T. intermedium* from one of the **ESt** tetraploids would likely require the following steps. First, only the **St** genome carrying the **E**-like ribotypes contributed to the formation of the allohexaploid. Yet, the process of how *Pseudoroegneria*, *Dasypyrum* and perhaps *Aegilops* genomes merged is unknown. Second, at the hexaploid stage, the 45S rDNA arrays residing on *Pseudoroegneria* and *Aegilops* chromosomes evolved in a concerted manner, leading to the complete conversion of *Aegilops*-like arrays toward the **E**-type, whereas *Dasypyrum*-like arrays remained intact. 5S rDNA arrays would not have been involved in any of the interlocus homogenizing or converging processes and would have retained relatively intact copies with respect to the three (sub)genomes.

Conclusions

Hybridization and polyploidy have played a major role in the evolution of the wheat tribe Triticeae, giving rise to many allopolyploid species with combined genomes (Löve 1984). By combining different genomes together into new organisms, these phenomena create patchwork genomes that represent interesting objects for the study of evolutionary dynamics of gene families. Ribosomal 45S and 5S DNA families in the allohexaploid *T. intermedium* are organized within several rDNA loci localized within all three subgenomes. Despite having similar architecture at the genome level, both gene families have undergone different patterns of evolution. Although the 45S rDNA family has largely evolved in a concerted manner, when effective homogenization mechanisms operating within as well as among majority of rDNA loci have been accompanied by loss of several loci, the data obtained suggest that in 5S rDNA family effective concerted evolution only operates within loci. Thanks to the markedly less effective concerted evolution, the 5S rDNA family retained more phylogenetic signal useful for reconstructing the allopolyploid history of intermediate wheatgrass.

Materials and Methods

Plant Material

Four accessions of hexaploid *T. intermedium* (Host) Barkworth et D. R. Dewey (intermediate wheatgrass; syn. *Elytrigia intermedia* [Host] Nevski, *Agr. intermedium* [Host] P. Beauv.) were used for analyses. The identification and choice of accessions were based on morphological, flow cytometric, cpDNA, and ITS diagnostic markers (Mahelka et al. 2005, 2007). For further details on the accessions, see Mahelka et al. (2011).

For genomic in situ hybridization, accessions of *P. spicata* (Pursh) Á. Löve (USDA accession identifier PI563869, origin in the USA), *D. villosum* (L.) P. Candargy (PI639751, Greece), and *A. tauschii* Coss (PI542278, Turkey) were used as DNA probes. 5S rDNA sequences were obtained from the accessions of *D. villosum* (same as discussed earlier), *H. piliferum* Hochst. ex Jaub. and Spach (W67257, Turkey), *Taeniatherum caput-medusae* (L.) Nevski (PI598389, Turkey), and *T. elongatum* (Host) D. R. Dewey (PI531718, Turkey). Seeds of the accessions were provided by the Germplasm Resources Information Network of the United States Department of Agriculture (USDA). All accessions were confirmed to be diploids ($2n = 14$) by chromosome counts.

DNA Extraction, Amplification, and Cloning

Genomic DNA of all plant material was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

ITS Region. PCRs of ITS1-5.8S-ITS2 region were done in triplicates using the ITS-Poa-f and ITS4 primers as described in Mahelka et al. (2007). Mahelka et al. (2007) showed that ITS in *T. intermedium* were largely homogenized but that they contained a minor proportion (~1%) of unhomogenized copies that comprised additional ITS diversity. The minor ITS variants could be pre-selected from the pooled ITS amplicon by SmaI-RFLPs (Mahelka et al. 2007). Pooled PCR products of each *T. intermedium* accession were used for cloning to analyze the "major" ribotype. Cloning was performed using the TOPO TA cloning kit (Invitrogen, Carlsbad, CA) following the manufacturer's instructions but downscaled to half reactions. Positive clones were denatured and re-amplified using the same primer pair and conditions as in the original PCR. Between 11 and 13 clones corresponding to the major ribotype were sequenced using the ITS4 primer (via GATC Biotech, Konstanz, Germany) in each accession. Amplification and sequencing of some minor ribotypes is possible in this case, as the pooled PCR products should contain all the ribotypes, albeit some of them in minor proportions. The minor ribotype was retrieved from the pooled PCR product with SmaI-RFLP as described in Mahelka et al. (2007).

SmaI cuts the PCR fragments at alignment position 178, so we cloned and sequenced only the longer fragments of approximately 430 nucleotides. The digested fragments containing the minor ribotypes were cut from gels, purified (Zymo Research, Orange, CA) and cloned using the Blunt-ended PCR Cloning Kit (GE Healthcare, Little Chalfont, UK) following the manufacturer's instructions. Positive strains were denatured and reamplified with the T7 promoter and U-19 primers. Between 13 and 15 clones corresponding to minor ribotypes were sequenced using the T7 primer in each accession. In total, we sequenced 102 ITS clones (47 clones of the major ribotype and 55 clones of the minor ribotype).

5S rDNA. PCR amplicons were obtained using the 5S DNA-L and -R primers of Appels et al. (1992). Reaction volumes of 25 µl contained 2.5 µl of Mg^{2+}-free reaction buffer, 2.5 mM $MgCl_2$, 200 µM of each dNTP, 0.15 µM of each primer, 20–50 ng of genomic DNA and 0.5 unit of *Taq* DNA polymerase. The thermocycling profile was as follows: 95 °C/3 min, 35× (95 °C/30 s, 58 °C/30 s, and 72 °C/2 min), 72 °C/30 min. PCR reactions were done in triplicates and mixed to avoid potential amplification bias. The primers amplify DNA in tandem arrays, each containing the 3′-end of the gene, the NTS and the 5′-end of the adjacent copy of the gene (Appels et al. 1992). To facilitate cloning and sequencing, the PCR products were digested with BamHI (16 h at 37 °C in a reaction volume of 50 µl). Following digestion, the fragments falling within the range of approximately 350–500 bp size were excised from the gel, purified, and cloned using the Blunt-ended PCR Cloning Kit following the manufacturer's instructions. Reamplifications were done with the T7 promoter and U-19 primers. Between 28 and 30 clones per accession were sequenced using the T7 primer, 115 clones in total.

For the purpose of phylogenetic analyses (discussed later), we obtained sequences of the following Triticeae diploids: *D. villosum*, *H. piliferum*, *T. caput-medusae*, and *T. elongatum*. The procedures of PCR and cloning were the same as in *T. intermedium*, only we sequenced 10 clones in each of the diploids.

Alignments and Choice of Representative Sequences for Phylogenetic Analysis

ITS Region. At first, 11 unspecific clones identified using BLASTn were excluded from the data set. Afterwards, true ITS sequences were aligned separately for each accession using ClustalX (Thompson et al. 1997) and then refined manually in BioEdit version 7.1.3.0 (Hall 1999) as necessary. The alignments were straightforward along their entire lengths, clearly differentiating sequences of major and minor ribotypes. Sequences displaying a mosaic pattern (i.e., combining different parts of the different ribotypes) were considered as recombinant and excluded from the analyses. Out of the 102 ITS clones of four *T. intermedium* accessions, 35 clones were excluded and 67 were kept for further analyses. Sequences of the 67 ITS clones are available in GenBank under the accession numbers KC843788–KC843854 (in the order as they appear in supplementary table S1, Supplementary Material online). Thirty-eight sequences corresponded to the major ribotype, whereas 29 sequences to the minor ribotype. Representative sequences covering the variation amplified from each accession (12 sequences in total, see supplementary table S1, Supplementary Material online) were then selected and used for phylogenetic analysis. The 12 representative sequences of all accessions were aligned together with the sequences of selected species from across the Triticeae (supplementary table S3, Supplementary Material online). In this respect, we included—apart from diploid Triticeae species representing different haplomes—also selected polyploids containing the **E** and **St** haplomes and their combination. The data set thus provided a reasonable representation of Triticeae taxa, which enabled us to analyze the *T. intermedium* sequences with respect to presumed diploid parental species (*Pseudoroegneria*, *Dasypyrum*, *Aegilops*, and/or *Thinopyrum* genera) as well as related polyploids. The data set of 84 sequences was augmented with *Brachypodium sylvaticum* (L.) P. Beauv. (Pooideae, Brachypodieae) and *Avena longiglumis* Durieu (Pooideae, Aveneae) used as outgroups. The final alignment thus consisted of 86 sequences and 622 sites.

5S Region. Prior to further analyses, all the 5S sequences were re-arranged in such a way that in each clone the sequence of NTS was followed by the sequence of the gene. Consequently, sequences of each accession were aligned separately using ClustalX and checked for recombinants as described in ITS. Supposed recombinants were excluded from the data sets. Furthermore, sequences containing deletions adjacent to the BamHI restriction site were considered as cloning artefacts and were also excluded from further analyses. From the total number of 115 clones, 26 clones were excluded (13 recombinants and 13 for the presence of deletions in the gene region). Eighty-nine clones were kept for further analyses. Sequences were uploaded to GenBank (accessions KC843855–KC843943 in the order as they appear in supplementary table S2, Supplementary Material online). Sequences of all accessions were then aligned altogether so that different orthologous groups could be defined (Baum et al. 2001). Putative orthologous groups, termed as unit classes, were first identified based upon visual inspection of the alignment, when sequences were grouped based upon overall similarity in size and sequence. Further, as the nucleotide sequence pattern rather than length variation in sequences per se is the key factor in determining unit classes, we confirmed the assignment into unit classes by a phylogenetic analysis. Using PAUP* 4b10 (Swofford 2003), we ran MP analyses using a heuristic search with 10 random addition replicates, keeping no more than 100 trees of length greater than or equal to 1 in each replicate, and with tree bisection-reconnection (TBR) branch swapping. *Hordeum brevisubulatum* ssp. *violaceum* (DQ776979) was used as the outgroup. Gaps were treated as missing data. Clade robustness was assessed by bootstrapping with 1,000 replicates with the same settings.

Alignment inspection and the MP led to the same conclusions. When we identified eight different unit classes in the

four *T. intermedium* accessions analyzed, three of them corresponded to the short units (51 sequences, designated short I–III), and five to the long units (38 sequences, long I–V) (supplementary fig. S1, Supplementary Material online). In the MP analysis, all the unit classes formed clades with bootstrap support of 100%, except for unit class short II with 90%. Units short I and II, and long I, II, and III contained sequences of all four accessions. Short III contained sequences of accessions *T. intermedium-1*, *-2*, and *-4*, long IV contained sequences of *T. intermedium-1* and *-3*, and long V were typical of accession *T. intermedium-4*. In the case of 5S sequences, we used a different approach for the choice of representative sequences than in the ITS data set. Instead of selecting individual clones representing the variation within the data sets, we used consensus sequences. This approach was adopted because of two reasons. First, nucleotide diversity within particular regions in some unit classes even exceeded that of the minor ITS ribotypes (discussed later), so the consensus sequences better represented the variation present within each unit class. Second, consensus sequences were also used to characterize different haplomes within the Triticeae (Baum et al. 2009), facilitating the subsequent analysis of different unit classes in a phylogenetic context within the Triticeae. Hence, within each unit class, a consensus sequence was computed using BioEdit. The threshold frequency for inclusion in the consensus sequence was set to 80%, except for two unit classes (short III and short V) that contained only 4 sequences and in which 75% was used (in such a case, if a substitution occurred in one sequence, the consensus sequence would still include a character state present in the majority of sequences). It is important to note that the consensus sequences do not represent variation present within the individual accessions but polymorphisms present within individual unit classes (i.e., found across all four *T. intermedium* accessions).

The different unit classes of *T. intermedium* were assigned to haplomes in a phylogenetic analysis with the representative haplomes of Triticeae. Representative sequences of different Triticeae haplomes were retrieved from different sources. Consensus sequences were either directly adopted from relevant sources if available or computed based on published sequences (supplementary table S3, Supplementary Material online). The threshold frequency for inclusion in the consensus was set to 80% to ensure consistency with the *T. intermedium* sequences. In some genera, only a single sequence appeared in either GenBank or published literature. In such cases, these single sequences were included in the analyses. Data sets containing sequences of short and long unit classes were analyzed separately. The alignments contained 36 sequences and 523 sites in the short units and 43 sequences and 607 sites in the long units, both including outgroups.

Sequence Comparisons and Inference of Functionality

Sequence Nucleotide Diversity

To quantify the diversity of ITS and 5S rDNA sequences, we calculated the following characteristics: 1) number of haplotypes; 2) number of polymorphic sites (singleton variable sites vs. parsimony informative); 3) number of mutations; 4) mean within-group nucleotide diversity (π); and 5) between-group mean distance. Descriptors 1–3 were calculated using DnaSP software (Rozas et al. 2003). Descriptors 1–4 were calculated separately for each sequence group, that is, for both ITS ribotypes and for the different 5S unit classes. Nucleotide diversity π as an average number of nucleotide differences per site between two sequences (eqs. 10.5 or 10.6 in Nei 1987) and between-group mean distances were calculated using MEGA5 (Tamura et al. 2011). Based on the transition/transversion bias (>2.5 in ITS and ~1 in 5S sequences), we used a Kimura 2-parameter substitution model and *p*-distance for ITS and 5S data sets, respectively. In both ITS and 5S data sets, gaps were treated using the pairwise deletion option. In ITS and 5S data sets, the characteristics were calculated for the complete regions as well as for the genes and spacers separately. As the ITS sequences of minor ribotypes were amplified after restriction digestion, the ITS1 region in these sequences contained only 52 nucleotides.

Inference of Functionality

In an attempt to distinguish sequences representing functional genes from those of pseudogenes, we examined both ITS and 5S data sets and characterized individual sequences as to the following three major aspects: substitutions at conserved sites, the ability to form secondary structures and relative rate tests.

Substitutions at Conserved Sites and GC Content

ITS Region. As all but three ITS sequences of the minor ribotype contain incomplete ITS1, we only focused on the 5.8S gene and the ITS2 to enable a comparison between the full-length sequences of both ribotypes. Within the ITS region, the 5.8S gene is the most valuable indicator of the functionality of the region (Harpke and Peterson 2008). We thus checked 5.8S sequences for the presence of four conserved motifs (M1–M4) and the ITS2 for the presence of one conserved motif (M5): 1) the Spermatophyta 16-bp motif M1 from Harpke and Peterson (2008): M1—CGATGAAGAACGTAGC; 2) the Angiosperm 14-bp motif M2 from Jobes and Thien (1997): M2—GAATTGCAGTCC; 3) the Viridiplantae 10-bp motif M3 from Harpke and Peterson (2008): M3—TTTGAA (C/T)GCA; 4) EcoRV site near the 5′ of the 5.8S gene of Viridiplantae (Liston et al. 1996): M4—GATATC; and 5) the Angiosperm motif present in ITS2 (Coleman 2007): M5—TGGT. Substitutions within these conserved motifs were considered as indications of pseudogenization. Additionally, GC content (%) of each sequence was calculated in BioEdit.

5S Region. In 5S genes, the ICR is considered to play a key role in the transcription of the gene. Transcription of the 5S gene is operated by RNA polymerase III (Pol III), whose interaction with the DNA strand is mediated by transcription factors TFIIIA and TFIIIC. These proteins bind to the DNA strand at three places, the so-called A-box (TFIIIC), the IE and the C-box (both TFIIIA, Pieler et al. 1987; Orioli et al. 2012). The positions of these gene-internal binding sites are 50–64 (A-box), 67–72 (IE), and 80–97 (C-box) downstream of the

TSS (Pieler et al. 1987). Further, a poly-T track (with a minimum of 4–6 Ts depending on the organism) downstream of the gene at position +121 is involved in RNA Pol III transcription termination (Fulneček and Kovařík 2007; Orioli et al. 2012). Finally, all the sequences were checked for the presence of gene-external upstream elements—the TATA box, usually located 25–30 bp upstream of the TSS and a C residue at position −1—both involved in the transcription initiation. GC content (%), which is never less than 50% in functional 5S molecules (Smirnov et al. 2008), was calculated for each sequence (5S gene and the NTS separately) using BioEdit.

Secondary Structure Analyses

5.8S Gene. Secondary structures of the 5.8S gene transcripts were predicted using the mFold web server (Zuker 2003). As both the 5′- and the 3′-terminal sequences of 5.8S rRNA do not pair with other parts of the same molecule, but are involved in the association with the 26/28S rRNA (Veldman et al. 1981), we used a constrained folding according to a secondary structure model. As there was no model of LSU of any Triticeae available at The Comparative RNA Web (CRW) Site (Cannone et al. 2002), the ability to build up the secondary structure reported for *Oryza sativa* (accession numbers M16845 and M11585) was examined in all 5.8S sequences. *Oryza sativa* is a suitable model for the Triticeae, as the 5.8S sequence of *O. sativa* differs from the consensus 5.8S sequence of the Triticeae (computed using the ITS data set used in this study with the threshold frequency 90%) at only three positions with no effect on the folding of the molecule. Settings for base pairing in mFold were as follows: F 46 106 3, F 49 62 3, F 70 97 3, F 111 119 3, F 120 143 4, and F 126 137 4. All other nucleotides were left unpaired: P 1 0 44, P 52 0 8, P 63 0 7, P 73 0 22, and P 144 0 20. As a means of assessing the stability of the structures, free energy levels (ΔG in kcal mol^{-1} at 37 °C) were calculated for each 5.8S secondary structure using the mFold server.

ITS2. Assessing the functionality of ITS2 secondary structure was based on the ability of ITS2 transcripts to form the common four-helix model (Coleman 2003, 2007). Secondary structures of ITS2 transcripts were analyzed using the ITS2 database version 3.0.13 (Selig et al. 2008; Koetschan et al. 2010) using the homology-modeling approach. We compared secondary structures of each ITS2 transcript with an a priori defined model by examining the homology (%) of each of the four helices with those of the model. We used a different model for each ITS ribotype. As the reference for sequences of the major ribotype, we used a direct sequence of the ITS region of *T. intermedium* accession 2 (DQ859052), as the reference for the minor ribotype we used *D. villosum* (AJ608150).

5S Gene. 5S RNA is a highly conserved molecule with a three-domain Y-shaped organization (Smirnov et al. 2008). Although the general shape and organization of the molecule is well known, potential deleterious effects of individual substitutions and/or deletions on its function are relatively unknown. Secondary structures were therefore examined in two ways. First, 5S RNA transcripts were analyzed as to their ability to form a secondary structure commonly found in the models available. We used models of *O. sativa* (M18171) and *Tri. monococcum* (X66383) available at the CRW site (Cannone et al. 2002), whose 5S RNA transcripts form virtually identical, three-domain structures with five helices and five loops. Using the mFold web server, the following constrictions were applied for pairing: helix I, F 1 118 9; helix II, F 14 65 2; F 16 62 5; helix III, F 27 52 2; F 29 48 4; helix IV, F 78 98 6, F 85 92 2; and helix V, F 67 108 4. All other nucleotides were left unpaired to allow the loops to form: P 33 0 12; P 66 0 1; P 72 0 6. Second, 5S RNA transcripts were directly folded without any constraints, and a molecule was considered functional if it formed a three-domain, *Y*-shaped secondary structure with all helices and loops, whereas minor disorders such as single substitutions were omitted. The stability of the secondary structures was assessed using free energy levels at default temperature (ΔG in kcal mol^{-1} at 37 °C). Visualizations of all secondary structures were done using RNAviz2 software (De Rijk et al. 2003).

Relative Rate Tests

ITS Region. Relative rate tests were performed in two ways. First, a two-cluster relative-rate test (Takezaki et al. 1995) implemented in the software package PHYLTEST version 2.0 (Kumar 1996), was used to examine the evolutionary rate constancy of rDNA sequences among the two different ribotypes (major and minor). In this test, the constancy of the molecular clock is examined for two lineages A and B, each represented by multiple sequences, when an outgroup lineage is given (C). The test compares the averages of observed numbers of substitutions per site (branch lengths, L_a and L_b) from the common ancestor of clusters A and B. Therefore, under the constancy of the molecular clock, the null hypothesis is that $L_a = L_b$, that is, $\delta = L_a - L_b = 0$. The deviation of δ from 0 is tested by a two-tailed normal deviate test *Z*. Rate constancy was examined using the K2-P distance method, and the test was done separately for the 5.8S and ITS2 regions with *Bromus tectorum* (AJ608154) used as outgroup.

Second, the relative rate test of Tajima (1993) was performed in a pairwise fashion using the software MEGA5. The test compares two sequences with an outgroup sequence by counting unique substitutions in both sequences. The molecular clock hypothesis can be rejected if one of the sequences accumulates a significantly larger number of unique substitutions. Individual sequences of both major and minor ribotypes were compared with sequences representing potentially functional rDNA variants corresponding to each ribotype. As no cDNA sequence of neither *T. intermedium* nor other Triticeae is available in GenBank, we used rDNA sequences of *T. intermedium* (DQ859052; direct sequence amplified from accession 2 of this study) and *D. villosum* (AJ608150), which best represent mutation-free sequences of major and minor ribotypes, respectively. In both cases, *B. tectorum* was used as the outgroup. The test was done separately for the 5.8S and ITS2 regions as earlier.

5S Region. In 5S sequences, we only performed the two-cluster relative-rate test, in particular due to the lack of reference sequences that would represent functional gene variants

corresponding to individual unit classes. The two-cluster relative-rate test was performed for all pairs of short and long unit classes and for the gene and the NTS separately. *Psathyrostachys juncea* (Z11437) and *Hor. brevisubulatum* ssp. *violaceum* (DQ776957) were used as the outgroups for the short and long unit classes, respectively.

Phylogenetic Analyses

To place both the ITS and 5S sequences obtained from *T. intermedium* in a phylogenetic context within the Triticeae, two phylogenetic reconstruction methods were employed: Bayesian inference (BI) and MP analysis.

ITS Region. BI was carried out as follows: 1) the GTR + G was determined as the best-fitting model of molecular evolution by hierarchical likelihood ratio tests (hLRTs) with MrModeltest version 2.3 (Nylander 2004); 2) six substitution rates and gamma distribution were specified as settings (MrBayes 3.1.2, Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003); 3) *Bra. sylvaticum* (L36511) was used as the outgroup; 4) two simultaneous Metropolis coupled Markov chain Monte Carlo analyses (Geyer 1991) with four chains each were run, incrementally heated by a temperature of 0.1 for 10.5 million generations, at the sampling frequency of 100; 5) standard deviation of split frequencies (0.01) was used as a convergence diagnostic; 6) after stationarity was reached, the first 25% trees were discarded as burn-in, and a consensus tree with branch lengths and posterior probabilities was computed. MP analysis was run in PAUP* 4b10 using the heuristic search, 10 random addition replicates, keeping no more than 100 trees of length greater than or equal to 1 in each replicate, TBR branch swapping and gaps treated as missing data. *Brachypodium sylvaticum* and *A. longiglumis* (Z11758) were used as outgroup. A strict consensus tree was constructed. Bootstrapping as a measure of topological robustness was performed with 1.000 replicates using the same settings.

5S Region. Phylogenetic analyses with 5S sequences were carried out using the same approach and settings as in the ITS data, only with the following modifications: BI, long units—1) the GTR + I + G was determined as the best-fitting model of molecular evolution; 2) six substitution rates and gamma distribution with a proportion of invariable sites were specified as settings; 3) *Hor. brevisubulatum* ssp. *violaceum* (Boiss. and Huet) Tzvelev (a long unit, DQ776957) was used as outgroup; 4) the analysis was run for 3 million generations. BI, short units—1) *Hor. brevisubulatum* ssp. *violaceum* (a consensus sequence—see supplementary table S3, Supplementary Material online) was used as the outgroup; 2) 2 million generations were run. The MP analyses of both long and short units were done as described for ITS, only *Hor. brevisubulatum* ssp. *violaceum* was used as the outgroup in both cases as discussed earlier.

In Situ Hybridization

FISH was performed as described in Mahelka and Kopecký (2010). DNA clone pTa71 (Gerlach and Bedbrook 1979) containing a 9 kb EcoRI fragment of wheat rDNA, which carries the 18S-5.8S-26S cluster of rRNA genes (here referred to as 45S rDNA) was labeled with biotin. A digoxigenin-labeled probe for 5S rDNA was prepared using PCR with a pair of specific primers (RICRGAC1 and RICRGAC2), which amplify 303 bp in rice (Fukui et al. 1994), using rice genomic DNA as a template. Sites of probe hybridization were detected by the Anti-DIG-FITC conjugate (Roche) and the streptavidin-Cy3 conjugate (Amersham). Chromosomes were counterstained with 1.5 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) made in a Vectashield antifade solution (Vector Laboratories). Slides were evaluated under an Olympus AX70 microscope equipped with epi-fluorescence and a SensiCam B/W camera. After the microscopic evaluation of FISH, reprobing of slides was performed using total genomic DNA of *P. spicata* labeled with digoxigenin, total genomic DNA of *D. villosum* labeled with biotin and sheared total genomic DNA of *A. tauschii* used as blocking DNA. ScionImage and Adobe Photoshop software were used for processing of color pictures.

Supplementary Material

Supplementary tables S1–S3 and figure S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by the Czech Science Foundation grant number 13-04454S to V.M., the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01) to D.K., and by the long-term research development project no. RVO 67985939. The authors thank Petra Caklová for excellent technical assistance in DNA lab and Ladislava Paštová for counting the chromosomes of diploid Triticeae. They also thank Jaroslav Fulneček for providing advice on assessing the 5S rRNA functionality.

References

Álvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol.* 29:417–434.

Appels R, Baum BR, Clarke BC. 1992. The 5S DNA units of bread wheat (*Triticum aestivum* L.). *Plant Syst Evol.* 183:183–194.

Arnheim N, Krystal M, Schmickel R, Wilson G, Ryder O, Zimmer E. 1980. Molecular evidence for genetic exchanges among ribosomal genes on non-homologous chromosomes in man and apes. *Proc Natl Acad Sci U S A.* 77:7323–7327.

Baum BR, Edwards T, Johnson DA. 2008. Loss of 5S rDNA units in the evolution of *Agropyron*, *Pseudoroegneria*, and *Douglasdeweya*. *Genome* 51:589–598.

Baum BR, Edwards T, Johnson DA. 2009. Phylogenetic relationships among diploid *Aegilops* species inferred from 5S rDNA units. *Mol Phylogenet Evol.* 53:34–44.

Baum BR, Feldman M. 2010. Elimination of 5S DNA unit classes in newly formed allopolyploids of the genera *Aegilops* and *Triticum*. *Genome* 53:430–438.

Baum BR, Johnson DA. 2008. Molecular confirmation of the genomic constitution of *Douglasdeweya* (Triticeae: Poaceae): demonstration of the utility of the 5S rDNA sequence as a tool for haplome identification. *Mol Genet Genomics.* 279:621–628.

Baum BR, Johnson DA, Bailey LG. 2001. Defining orthologous groups among multicopy genes prior to inferring phylogeny, with special emphasis on the Triticeae (Poaceae). *Hereditas* 135: 123–138.

Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M. 2011. The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* 334:1524–1529.

Blattner FR. 2004. Phylogenetic analysis of *Hordeum* (Poaceae) as inferred by nuclear rDNA ITS sequences. *Mol Phylogenet Evol.* 33: 289–299.

Brasileiro-Vidal AC, Cuadrado A, Brammer SP, Zanatta ACA, Prestes AM, Moraes-Fernandes MIB, Guerra M. 2003. Chromosome characterization in *Thinopyrum ponticum* (Triticeae, Poaceae) using in situ hybridization with different DNA sequences. *Genet Mol Biol.* 26: 505–510.

Buckler ES, Ippolito A, Holtsford TP. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* 145:821–832.

Cannone J, Subramanian S, Schnare M, et al. (14 co-authors). 2002. The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* 3:2.

Coleman AW. 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends Genet.* 19:370–375.

Coleman AW. 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Res.* 35:3322–3329.

Cronn RC, Zhao X, Paterson AH, Wendel JF. 1996. Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. *J Mol Evol.* 42: 685–705.

De Rijk P, Wuyts J, De Wachter R. 2003. RnaViz2: an improved representation of RNA secondary structure. *Bioinformatics* 19:299–300.

Dubcovsky J, Dvořák J. 1995. Ribosomal RNA multigene loci: nomads of the Triticeae genomes. *Genetics* 140:1367–1377.

Dvořák J, Zhang H-B, Kota RS, Lassner M. 1989. Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. *Genome* 32:1003–1016.

Eickbush TH, Eickbush DG. 2007. Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 175:477–485.

Falistocco E, Passeri V, Marconi G. 2007. Investigations of 5S rDNA of *Vitis vinifera* L.: sequence analysis and physical mapping. *Genome* 50: 927–938.

Fehrer J, Krak K, Chrtek J Jr. 2009. Intra-individual polymorphism in diploid and apomictic polyploid hawkweeds (*Hieracium*, Lactuceae, Asteraceae): disentangling phylogenetic signal, reticulation, and noise. *BMC Evol Biol.* 9:239.

Fukui K, Kamisugi Y, Sakai F. 1994. Physical mapping of 5S rDNA loci by direct-cloned biotinylated probes in barley chromosomes. *Genome* 37:105–111.

Fukushima K, Imamura K, Nagano K, Hoshi Y. 2011. Contrasting patterns of the 5S and 45S rDNA evolutions in the *Byblis liniflora* complex (Byblidaceae). *J Plant Res.* 124:231–244.

Fulneček J, Kovařík A. 2007. Low abundant spacer 5S rRNA transcripts are frequently polyadenylated in *Nicotiana*. *Mol Genet Genomics.* 278:565–573.

Fulneček J, Lim KY, Leitch AR, Kovařík A, Matyášek R. 2002. Evolution and structure of 5S rDNA loci in allotetraploid *Nicotiana tabacum* and its putative parental species. *Heredity* 88:19–25.

Galasso I, Blanco A, Katsiotis A, Pignone D, Heslop-Harrison JS. 1997. Genomic organization and phylogenetic relationships in the genus *Dasypyrum* analysed by Southern and in situ hybridization of total genomic and cloned DNA probes. *Chromosoma* 106:53–61.

Garcia S, Crhák Khaitová L, Kovařík A. 2012. Expression of 5S rRNA genes linked to 35S rDNA in plants, their epigenetic modification and regulatory element divergence. *BMC Plant Biol.* 12:95.

Garcia S, Lim KY, Chester M, Garnatje T, Pellicer J, Valles J, Leitch AR, Kovařík A. 2009. Linkage of 35S and 5S rRNA genes in *Artemisia* (family Asteraceae): first evidence from angiosperms. *Chromosoma* 118:85–97.

Gerlach WL, Bedbrook JR. 1979. Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res.* 8: 4851–4855.

Geyer CJ. 1991. Markov chain Monte Carlo maximum likelihood. In: Keramidas EM, editor. Computing science and statistics: proceedings of the 23rd Symposium on the Interface. Fairfax Station (VA): Interface Foundation. p. 156–163.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 41:95–98.

Harpke D, Peterson A. 2006. Non-concerted ITS evolution in *Mammillaria* (Cactaceae). *Mol Phylogenet Evol.* 41:579–593.

Harpke D, Peterson A. 2008. 5.8S motifs for the identification of pseudogenic ITS regions. *Botany* 86:300–305.

Hsiao C, Chatterton NJ, Asay KH, Jensen KB. 1995. Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences. *Genome* 38:211–223.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.

Jiang J, Gill BS. 1994. New 18S-26S ribosomal RNA gene loci: chromosomal landmarks for the evolution of polyploid wheats. *Chromosoma* 103:179–185.

Jobes DV, Thien LB. 1997. A conserved motif in the 5.8S ribosomal RNA (rRNA) gene is a useful diagnostic marker for plant internal transcribed spacer (ITS) sequences. *Plant Mol Biol Rep.* 15: 326–334.

Kay KM, Whittall JB, Hodges SA. 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evol Biol.* 6:36.

Kellogg EA, Appels R. 1995. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. *Genetics* 140: 325–343.

Kishii M, Wang RRC, Tsujimoto H. 2005. GISH analysis revealed new aspect of genomic constitution of *Thinopyrum intermedium*. *Czech J Genet Plant.* 41(Special issue):91–95.

Koetschan C, Förster F, Keller A, Schleicher T, Ruderisch B, Schwarz R, Müller T, Wolf M, Schultz J. 2010. The ITS2 database III-sequences and structures for phylogeny. *Nucleic Acids Res.* 38:D275–D279.

Kotseruba V, Pistrick K, Blattner FR, et al. (11 co-authors). 2010. The evolution of the hexaploid grass *Zingeria kochii* (Mez) Tzvel. (2n = 12) was accompanied by complex hybridization and uniparental loss of ribosomal DNA. *Mol Phylogenet Evol.* 56:146–155.

Kovařík A, Dadejová M, Lim YK, Chase MW, Clarkson JJ, Knapp S, Leitch AR. 2008. Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. *Ann Bot.* 101: 815–823.

Kovařík A, Matyášek R, Lim KY, Skalická K, Koukalová B, Knapp S, Chase M, Leitch AR. 2004. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. *Biol J Linn Soc.* 82:615–625.

Kumar S. 1996. PHYLTEST: phylogenetic hypothesis testing software. Version 2.0. Pennsylvania (PA): Pennsylvania State University, University Park.

Li DY, Ru YY, Zhang XY. 2004. Chromosomal distribution of the 18S-5.8S-26S rDNA loci and heterogeneity of nuclear ITS regions in *Thinopyrum intermedium* (Poaceae: Triticeae). *Acta Bot Sin.* 46: 1234–1241.

Li DY, Zhang XY. 2002. Physical localization of the 18S-5.8S-26S rDNA and sequence analysis of ITS regions in *Thinopyrum ponticum* (Poaceae: Triticeae): implications for concerted evolution. *Ann Bot.* 90:445–452.

Li HJ, Wang XM. 2009. *Thinopyrum ponticum* and *Th. intermedium*: the promising source of resistance to fungal and viral diseases of wheat. *J Genet Genomics.* 36:557–565.

Liston A, Robinson WA, Oliphant JM, Alvarez-Buylla ER. 1996. Length variation in the nuclear ribosomal DNA internal transcribed spacer region of non-flowering seed plants. *Syst Bot.* 21:109–120.

Liu C, Yang ZJ, Jia JQ, Li GR, Zhou JP, Ren ZL. 2009. Genomic distribution of a long terminal repeat (LTR) Sabrina-like retrotransposon in Triticeae species. *Cereal Res Commun.* 37:363–372.

Liu Q, Ge S, Tang H, Zhang X, Zhu G, Lu BR. 2006. Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast trnL-F sequences. *New Phytol.* 170:411–420.

Liu ZW, Wang RRC. 1993. Genome analysis of *Elytrigia caespitosa*, *Lophopyrum nodosum*, *Pseudoroegneria geniculata* ssp. *scythica*, and *Thinopyrum intermedium* (Triticeae: Gramineae). *Genome* 36: 102–111.

Long EO, Dawid IB. 1980. Repeated genes in eukaryotes. *Annu Rev Biochem.* 49:727–764.

Löve Á. 1984. Conspectus of the Triticeae. *Feddes Repertorium* 95: 425–521.

Mahelka V, Fehrer J, Krahulec F, Jarolímová V. 2007. Recent natural hybridization between two allopolyploid wheatgrasses (*Elytrigia*, Poaceae): ecological and evolutionary implications. *Ann Bot.* 100: 249–260.

Mahelka V, Kopecký D. 2010. Gene capture from across the grass family in the allohexaploid *Elymus repens* (L.) Gould (Poaceae, Triticeae) as evidenced by ITS, GBSSI, and molecular cytogenetics. *Mol Biol Evol.* 27:1370–1390.

Mahelka V, Kopecký D, Paštová L. 2011. On the genome constitution and evolution of intermediate wheatgrass (*Thinopyrum intermedium*: Poaceae, Triticeae). *BMC Evol Biol.* 11:127.

Mahelka V, Suda J, Jarolímová V, Trávníček P, Krahulec F. 2005. Genome size discriminates between closely related taxa *Elytrigia repens* and *E. intermedia* (Poaceae: Triticeae) and their hybrid. *Folia Geobot.* 40: 367–384.

Malinská H, Tate JA, Matyášek R, Leitch AR, Soltis DE, Soltis PS, Kovařík A. 2010. Similar patterns of rDNA evolution in synthetic and recently formed natural populations of *Tragopogon* (Asteraceae) allotetraploids. *BMC Evol Biol.* 10:291.

Mayol M, Rosselló JA. 2001. Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. *Mol Phylogenet Evol.* 19: 167–176.

Muir G, Fleming CC, Schlötterer C. 2001. Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur L*. *Mol Biol Evol.* 18:112–119.

Negi MS, Rajagopal J, Chauhan N, Cronn R, Lakshmikumaran M. 2002. Length and sequence heterogeneity in 5S rDNA of *Populus deltoides*. *Genome* 45:1181–1188.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet.* 39:121–152.

Nylander JAA. 2004. MrModeltest v2. Program distributed by the author. Uppsala (Sweden): Evolutionary Biology Centre, Uppsala University.

Orioli A, Pascali C, Pagano A, Teichmann M, Dieci G. 2012. RNA polymerase III transcription control elements: themes and variations. *Gene* 493:185–194.

Pieler T, Hamm J, Roeder RG. 1987. The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell* 48:91–100.

Pinhal D, Yoshimura TS, Araki CS, Martins C. 2011. The 5S rDNA family evolves through concerted and birth-and-death evolution in fish genomes: an example from freshwater stingrays. *BMC Evol Biol.* 11: 151.

Renny-Byfield S, Chester M, Kovařík A, et al. (11 co-authors). 2011. Next generation sequencing reveals evidence of genome downsizing and elimination of paternally derived repetitive DNA sequences in tobacco. *Mol Biol Evol.* 28:2843–2854.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Rooney AP. 2004. Mechanisms underlying the evolution and maintenance of functionally heterogeneous 18S rRNA genes in Apicomplexans. *Mol Biol Evol.* 21:1704–1711.

Rooney AP, Ward TJ. 2005. Evolution of a large ribosomal RNA multigene family in filamentous fungi: birth and death of a concerted evolution paradigm. *Proc Natl Acad Sci U S A.* 102:5084–5089.

Rozas J, Sánchez-Delbarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497.

Sajdak LS, Reed KM, Phillips RB. 1998. Intraindividual and interspecies variation in the 5S rDNA of coregonid fish. *J Mol Evol.* 46:680–688.

Sastri DC, Hilu K, Appels R, Lagudah ES, Playford J, Baum BR. 1992. An overview of evolution in plant 5S DNA. *Plant Syst Evol.* 183: 169–181.

Scoles GJ, Gill BX, Xin Z-Y, Clarke BC, McIntyre CL, Chapman C, Appels R. 1988. Frequent duplication and deletion events in the 5S RNA genes and the associated spacer regions of the Triticeae. *Plant Syst Evol.* 160:105–122.

Selig C, Wolf M, Müller T, Dandekar T, Schultz J. 2008. The ITS2 database II: homology modelling RNA structure for molecular systematics. *Nucleic Acids Res.* 36:D377–D380.

Skalická K, Lim KY, Matyášek R, Matzke M, Leitch AR, Kovařík A. 2005. Preferential elimination of repeated DNA sequences from the paternal, *Nicotiana tomentosiformis* genome donor of a synthetic, allotetraploid tobacco. *New Phytol.* 166:291–303.

Smirnov AV, Entelis NS, Krasheninnikov IA, Martin R, Tarassov IA. 2008. Specific features of 5S rRNA structure—its interactions with macromolecules and possible functions. *Biochemistry* 73:1418–1437.

Smit S, Widmann J, Knight R. 2007. Evolutionary rates vary among rRNA structural elements. *Nucleic Acids Res.* 35:3339–3354.

Soltis PS, Soltis DE. 1991. Multiple origins of the allotetraploid *Tragopogon mirus* (Compositae): rDNA evidence. *Syst Bot.* 16: 407–413.

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland (MA): Sinauer Associates.

Tajima F. 1993. Simple methods for testing molecular clock hypothesis. *Genetics* 135:599–607.

Takezaki N, Rzhetsky A, Nei M. 1995. Phylogenetic test of the molecular clock and linearized tree. *Mol Biol Evol.* 12:823–833.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28:2731–2739.

Tang ZX, Yang ZJ, Fu SL, Yang MY, Li GR, Zhang HQ, Tan FQ, Ren Z. 2011. A new long terminal repeat (LTR) sequence allows to identify J genome from JS and St genomes of *Thinopyrum intermedium*. *J Appl Genet.* 52:31–33.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876–4882.

Uslu E, Reader SM, Miller TE. 1999. Characterization of *Dasypyrum villosum* (L.) Candargy chromosomes by fluorescent in situ hybridization. *Hereditas* 131:129–134.

Veldman MG, Klootwijk J, van Heerikhuizen H, Planta RJ. 1981. The nucleotide sequence of the intergenic region between the 5.8S and 26S rRNA genes of the yeast ribosomal RNA operon. Possible implications for the interaction between 5.8S and 26S rRNA and the processing of the primary transcript. *Nucleic Acids Res.* 9:4847–4862.

Vizoso M, Vierna J, González-Tizón AM, Martínez-Lage A. 2011. The 5S rDNA gene family in mollusks: characterization of transcriptional regulatory regions, prediction of secondary structures, and long-term evolution, with special attention to *Mytilidae mussels*. *J Hered.* 102:433–447.

Volkov RA, Borisjuk NV, Panchuk II, Schweizer D, Hemleben V. 1999. Elimination and rearrangement of parental rDNA in the allotetraploid *Nicotiana tabacum*. *Mol Biol Evol.* 16:311–320.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Mol Biol.* 42: 225–249.

Wendel JF, Schnabel A, Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc Natl Acad Sci U S A.* 92:280–284.

Wicke S, Costa A, Muñoz J, Quandt D. 2011. Restless 5S: the re-arrangement(s) and evolution of the nuclear ribosomal DNA in land plants. *Mol Phylogenet Evol.* 61:321–332.

Yu H, Zhang C, Ding C, Wang X, Zhang H, Zhou Y. 2010. Genome constitutions of *Pseudoroegneria geniculata*, *P. geniculata* ssp. *scythica* and *P. geniculata* ssp. *pruinifera* (Poaceae: Triticeae) revealed by genomic in situ hybridization. *Acta Physiol Plant.* 32: 645–650.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415.