### **Deep-Level Diagnostic Value of the rDNA-ITS Region**

### Mark A. Hershkovitz\*† and Louise A. Lewis‡

\*NCBI (Genbank), National Library of Medicine, National Institute of Health; †Laboratory of Molecular Systematics, Smithsonian Institution; and ‡Department of Biology, University of New Mexico

The similarity of certain reported angiosperm rDNA internal transcribed spacer (ITS) region sequences to those of green algae prompted our analysis of the deep-level phylogenetic signal in the highly conserved but short 5.8S and hypervariable ITS2 sequences. We found that 5.8S sequences yield phylogenetic trees similar to but less well supported than those generated by a ca. 10-fold longer alignment from rDNA-18S sequences, as well as independent evidence. We attribute this result to our finding that, compared to 18S, the 5.8S has a higher proportion of sites subject to vary and greater among-site substitution rate homogeneity. We also determined that our phylogenetic results are not likely affected by intramolecular compensatory mutation to maintain RNA secondary structure nor by evident systematic biases in base composition. Despite historical homology, there appears to be no ITS2 primary sequence similarity shared between fungi, green algae, and angiosperms. ITS2 sequences within each of these groups, however, share sufficient similarity to cluster correctly on the basis of alignability. Our results indicate that ITS region sequences can diagnose organismal origins and phylogenetic relationships at many phylogenetic levels and provide a useful paradigm for molecular evolutionary study.

#### Introduction

Diagnosing the organismal identity or affinities of a DNA sequence can be especially challenging in the absence of adequate evolutionary characterization of phylogenetic signal in the molecule. The internal transcribed spacer (ITS) region separating 18S and 26S nuclear ribosomal DNA (rDNA), which includes two spacers (ITS1 and ITS2) and the intervening 5.8S coding sequence, has become well characterized across interspecific-intergeneric-level divergences (reviewed in Baldwin et al. 1995) but is rarely compared at deeper levels. The spacers are reputedly poorly conserved at deeper levels, e.g., above the family level in angiosperms (Baldwin et al. 1995). The 5.8S coding region, by contrast, is little used in phylogenetics because of its substantial conservation at deeper levels (Troitsky and Bobrova 1986; Troitsky et al. 1991; Suh, Thien, and Zimmer 1992), which, considering its small size (164– 165 bp), translates to little informative variation relative to more exploited sequences, e.g., 18S rDNA. For example, a 5.8S analysis that included sequences from three fungi, one green alga, a moss, and three angiosperms (Troitsky and Bobrova 1986) produced the "correct" topology among the land plants but an unresolved trichotomy among land plants, the green alga, and fungi. Casting further suspicion on 5.8S phylogenetic signal are the potential for character nonindependence because of secondary structural constraints (Wheeler and Honeycutt 1988) and mutational saturation such as characteristic of 5S rDNA (Steele et al. 1991).

In the course of analyzing conserved sequence and structural motifs in angiosperm ITS sequences (Hershkovitz and Zimmer 1996), we discovered (via FASTA searches) similarities between 5.8S/ITS sequences re-

© 1996 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

ported for eight species of monkeyflower (Mimulus state tatus complex; Scrophulariaceae; Ritland and Straus 1993; Ritland, Ritland, and Straus 1993) and those of green algae. Subsequently, a typical angiospermous **F**S region was recovered from Mimulus guttatus (A. Liston personal communication), confirming our suspicion that the reported sequences were contaminants, but also prompting this reconsideration of deep-level phylogenetic signal in the ITS region. In this study, we compare 5.8S phylogenetic trees with those derived using rDNA-18S sequences and independent evidence and compare the amount of sequence variation, relative base substitution patterns, and among-site rate heterogeneity in 538S and 18S. We also compare conservation patterns in ITS2 sequences from fungi, green algae, and angiosperms and determine the degree to which ITS2 sequences from these taxa cluster on the basis of pairwise alignability. 69

#### Methods

#### **Taxon Sampling**

The 5.8S analysis included representative anglosperms, green algae (including two of the putative *Mimulus* sequences), available representatives of other chaorophytes, and fungal and protistan outgroups (table 1). Phylogenetic signal of the 5.8S data was compared with that of an 18S data set that was similarly representative taxonomically but differing in composition because of sequence availability.

We limited the spacer analysis to ITS2 because of sequence availability and because we found that it contained greater sequence conservation than ITS1. The sample comprised 2 of the putative *Mimulus* sequences. 5 fungi, 4 chlorophycean (here, for simplicity, including trebouxiophycean—see Friedl 1995) green algae, 2 conifers, and 14 diverse angiosperms, including representatives of 6 families of the monophyletic (Cronquisi 1981; Chase et al. 1993) angiosperm order Caryophyllales. The Caryophyllales were included to demonstrate interfamilial phylogenetic signal. ITS2 conservation in angiosperms is considered in detail elsewhere (Hershkovitz and Zimmer 1996).

ģ

guest

Key words: ITS, 5.8S, rDNA evolution, eukaryotes, maximumlikelihood, among-site rate variation.

Address for correspondence and reprints: Mark A. Hershkovitz, NCBI (Genbank), National Library of Medicine, National Institute of Health, Bldg. 38A, 8600 Rockville Pike, Bethesda, Maryland 20984. E-mail: hershkov@ncbi.nlm.nih.gov.

 Table 1

 Taxa, Acronyms, and Genbank Accessions

Acronym	Taxon	Family	Broader Affiliation	Genbank	Source
1. 5.8S sequences	<u> </u>				
Angiosperms <sup>a</sup>					
ARABI CANEL CUCUR LACTU LYCOP POTAM	Arabidopsis thaliana Canella winterana Cucurbita melo Lactuca sativa Lycopersicon esculentum Potamogeton patans	Brassicaceae Winteraceae Cucurbitaceae Asteraceae Solanaceae Potamogetonaceae	Magnoliopsida; eudicot Magnoliopsida; magnoliid dicot Magnoliopsida; eudicot Magnoliopsida; eudicot Magnoliopsida; eudicot Liliopsida	X52320 L03844 M36377 L13957 X52265	DNA DNA DNA DNA DNA RNA
SILEN TRITI VICIA	Silene dioica Triticum vulgare Vicia faba	Caryophyllaceae Poaceae Fabaceae	Magnoliopsida; eudicot Liliopsida Magnoliopsida; eudicot	X86830 M10469 M10471	DNA RNA RNA
Gnetophytes					
EPHED GNETU	Ephedra kokonika Gnetum gnemon	Ephedraceae Gnetaceae	Gnetopsida, Ephedrales Gentopsida, Gnetales	X15676 <sup>b</sup>	RNA RNA <sub>O</sub>
Conifers <sup>c</sup>					Inw
	Picea mexicana Pinus contorta	Pinaceae	Coniferopsida, Coniferales	U24251 U23956	
TAXUS	Taxus baccata	Taxaceae	Coniferopsida, Taxales	X93991	DNA
Ferns			,,		fro
MARSI OSMUN	Marsilea quadrifolia Osmunda regalis	Marsileaceae Osmundaceae	Filicopsida; Marsiliales Filicopsida; Filicales	X15939 X63199	
Moss					s://a
MNIUM	Mnium rugicum	Mniaceae	Bryopsida	X13432	RNA
Chlorophycean, tr	ebouxiophycean, micromonado	phycean, green algae <sup>d</sup>			den
CHLOR	Chlorella ellipsoidea	?e	?	D13340	
CHLAM	Chlamydomonas reinhardtii	Chlamydomonaceae	Chlorophyceae; Volvocales;	X65621	DNA
GONIU	Gonium pectorale	Volvocaceae	Chlorophyceae; Volvocales	U23534	DNA <sup>O</sup>
	Mimulus nasutus"	?' ?f	? ?	L02/99 L02801	
PANDO	Pandorina morum	Volvocaceae	Chlorophyceae: Volyocales	L02001	DNA
SPERM	Spermatozopsis similis	?g	Chlorophyceae; Volvocales?	X69488	
TETRA	Tetraselmis striata	Chlorodendraceae	Micromonadophyceae	X65967	DNA
Ulvophycean gree	en alga				cle/
CLADO	Cladophora albida	Cladophoraceae	Ulvophyceae	h	DNA
Fungi <sup>i</sup>					9/1:
GLOMU	Glomus mossae	Endogonaceae	Zygomycotina; Endogonales	U16756	
BIPOL	Bipolaris oryzae	Pleosporaceae	Ascomycotina; Loculoascomycetes	X78122	DNA
SACCH	Saccharomyces cervesiae	Saccharomycetaceae	Ascomycotina; Hemiscomycetes	K01048	$RNA_{+}$
SCLER	Sclerotinia sclerotiorum	Sclerotiniaceae	Ascomycotina; Discomycetes	M96382	DNA
	Volvariella volvacea	Pluteaceae	Basidiomycotina: Agaricales	L115973	
Danum alaa	Vorranena vorracea	Theeeede	Busicioni ycotina, riganeaics	013975	Dungu
SCYTO	Scytosiphon lomentaria	Scytosiphonaceae	Phaeophyceae	D16558	DNA
Red algae					21
CYANI SARCO	Cyanidium caldarium Sarcodiotheca furcata	Porphyridiaceae Solieriaceae	Porphyridiales Gigartinales	X65077 U21346	
Dinoflagellates	· · · ·				5t N
CRYPT PRORO	Crypthecodinium cohnii Prorocentrum micans	Crypthecodiniaceae Prorocentraceae	Dinophyceae; Peridiniales Dinophyceae; Prorocentrales	M25116 M14649	
Oomycete					
РНҮТО	Phytophthora megasperma	Pythiaceae	Peronosporales	X75632	DNA
2. 18S sequences					
Angiosperms <sup>a</sup>					
DRIMY18	Drimys winteri	Winteraceae	Magnoliopsida; magnoliid dicot	L24089	DNA
GLYCI18	Glycine max	Fabaceac	Magnoliopsida; cudicot	X02623	RNA
	Gossypium hirsutum Hydrocotyle sibthornicidez	Malvaceae	Magnoliopsida; eudicot	L24145	
LYCOP18	Lycopersicon esculentum	Solanaceae	Magnoliopsida: eudicot	X51576	DNA
ORYZA18	Oryza sativa	Poaceae	Liliopsida	X00755	RNA
PRUNU18	Prunus persica	Rosaceae	Magnoliopsida; eudicot	L28749	DNA
RIBES18	Ribes aureum	Grossulariaceae	Magnoliopsida; eudicot	L28143	DNA
SINAP18	Sinapis alba	Brassicaceae	Magnoliopsida; eudicot	X66325	DNA
SPINAIS	spinacia oleracea	Chenopodiaceae	magnonopsida; eudicot	L24420	DNA

#### Table 1 Continued

Acronym	Taxon	Family	Broader Affiliation	Genbank	Source
Gnetophyte					
EPHED18	Ephedra sinica	Ephedraceae	Gnetopsida, Ephedrales	D38242	RNA
Conifer					
PINUS18	Pinus griffithii	Pinaceae	Coniferopsida	X75080	DNA
Cycadophyte					
ZAMIA18	Zamia pumila	Zamiaceae	Cycadopsida	M20017	DNA
Ferns					
OSMUN18	Osmunda cinnamomea	Osmundaceae	Filicopsida; Filicales	U18516	DNA
SALVI18	Salvinia natans	Salviniaceae	Filicopsida; Salviniales	X90413	DNA
Moss					Do
MNIUM18	Mnium hornum	Mniaceae	Bryopsida	X80985	DNA
Chlorophycean, tr	rebouxiophycean, micromonado	phycean green algae <sup>d</sup>			oad
CHLAMI8	Chlamydomonas reinhardtii Chlaralla allingoidea	Chlamydomonaceae	Chlorophyceae; Volvocales	M32703 X63520	
SPERM18	Spermatozopsis similis	?s	Chlorophyceae: Volyocales?	X65557	DNA
TETRA18	Tetraselmis striata	Chlorodendraceae	Micromonadophyceae	X70802	DNA
<b>VOLVO18</b>	Volvox carteri	Volvocaceae	Chlorophyceae; Volvocales	X53904	DNA
Ulvophycean gree	en alga				s://a
CLADO18	Cladophora albida	Cladophoraceae	Ulvophyceae	Z35317	DNA
Fungi <sup>i</sup>					den
GIGAS18	Gigaspora margarita	Gigasporaceae	Zygomycotina; Gigasporales	X58726	DNĂ
GLOMU18	Glomus intraradices	Endogonaceae	Zygomycotina; Endogonales	X58725	DNA
	Omphalina umbellifera Pleurotus ostreatus	Polyporaceae	Basidiomycotina, Agaricales Basidiomycotina: Aphyllophorales	U23543 U23544	DNA
SCLER18	Sclerotinia sclerotiorum	Sclerotiniaceae	Ascomycotina; Discomycetes	L37541	DNA
Brown alga					nbe
SCYTO18	Scytosiphon lomentaria	Scytosiphonaceae	Phaeophyceae	L43066	DNĂ
Red algae <sup>k</sup>					ticle
PORPH18	Porphyridium aerugineum	Porphyridiaceae	Porphyridiales	L27635	RNA
RHODE18	Rhodella maculata	Porphyridiaceae	Porphyridiales	U21217	RNA
SARCO18	Sarcodiotheca furcata	Solieriaceae	Gigartinales	U43553	DNA
Dinoflagellate					6/9
<b>CRYPT</b> 18	Crypthecodinium cohnii	Crypthecodiniaceae	Dinophyceae; Peridiniales	M64245	DNA
Oomycete					691
<b>PHYTO18</b>	Phytophthora megasperma	Pythiaceae	Peronosporales	X54265	DNA
3. ITS2 sequences					gue
Angiosperms <sup>a</sup>					t st
AMARA	Amaranthus retroflexus	Amaranthaceae	Magnoliopsida; eudicot; Caryophyllales	L48798	
CANEL	Arabiaopsis inaliana Canella winterana	Winteraceae	Magnoliopsida; eudicol Magnoliopsida: magnoliid dicot	A32320 L03844	
CUCUR	Cucurbita melo	Cucurbitaceae	Magnoliopsida; eudicot	M36377	DNA
LACTU	Lactuca sativa	Asteraceae	Magnoliopsida; eudicot	L13957	DNA
LYCOP	Lycopersicon esculentum	Solanaceae	Magnoliopsida; eudicot	X52265	DNA
	Malanthemum racemosum Mollugo verticillata	Liliaceae	Liliopsida Magnoliopsida: eudicot: Carvonhyllales	U24041 1 48700	DNA
PHYTL	Phytolacca americana	Phytolaccaceae	Magnoliopsida; eudicot; Caryophyllales	L48800	DNA
SILEN	Silene dioica	Caryophyllaceae	Magnoliopsida; eudicot; Caryophyllales	X86830	DNA
TALIN	Talinum paniculatum	Portulacaceae	Magnoliopsida; eudicot; Caryophyllales	L48801	DNA
TETRG	Tetragonia tetragonioides	Aizoaceae	Magnoliopsida; eudicot; Caryophyllales	L48802	DNA
VICIF	Vicia faba	Fabaceae	Magnoliopsida: eudicot	X17535	DNA
Conifers			<b>6 1 7 1</b>		
PICEA.	Picea mexicana	Pinaceae	Coniferopsida, Coniferales	U24251	DNA
PINUS	Pinus contorta	Pinaceae	Coniferopsida, Coniferales	U23956	DNA
Chlorophycean, th	rebouxiophycean, micromonado	phycean green algae <sup>d</sup>			
CHLOR	Chlorella ellipsoidea	?e	?	D13340	DNA
	Gonium pectorale	Volvocaceae	Chiorophyceae; Volvocales	U23534	DNA DNA
MIURF	"Mimulus nusulus"	2f	?	L02801	DNA
PANDO	Pandorina morum	Volvocaceae	Chlorophyceae; Volvocales	U23530	DNA
SPERM	Spermatozopsis similis	?s	Chlorophyceae; Volvocales?	X69488	DNA
SPONG	Spongiochloris spongosa	Chlorococcaceae	Chlorococcales; Chlorophyceae	U34776	DNA

Downloaded from https://acac

Table 1 Continued

Acronym	Taxon	Family	Broader Affiliation	Genbank	Source
Fungi <sup>i</sup>					
BIPOL	Bipolaris oryzae	Pleosporaceae	Ascomycotina; Loculoascomycetes	X78122	DNA <sup>j</sup>
GIGAS	Gigaspora margarita	Gigasporaceae	Zygomycotina; Gigasporales	X84233	DNA
HETER	Heterobasidion annosum	Schizophyllaceae	Basidiomycotina; Aphyllophorales	X70024	DNA
SACCH	Saccharomyces cerevesiae	Saccharomycetaceae	Ascomycotina; Hemiscomycetes	K01048	RNA
VOLVA	Volvariella volvacea	Pluteaceae	Basidiomycotina; Agaricales	U15973	DNA

NOTE.—Within sequence data subsets, taxa are grouped by broader affiliation. The acronyms are those used in tables 2 and 4 and figures 1–8. The classification is intended only as a guide to the diversity—see footnoted references for additional information on presumed interrelationships within groups represented by three or more samples.

<sup>a</sup> Cronquist (1981); Chase et al. (1993).

<sup>b</sup> Not currently accessioned in Genbank; transcribed from Troitsky et al. (1991).

<sup>c</sup> Page (1990).

<sup>d</sup> Mattox and Stewart (1984); Friedl (1995).

e Chlorella is polyphyletic; species identification and affinities remain problematic; Huss and Sogin (1990); Wilcox et al. (1992); Friedl (1995).

f Organismal origin and precise phylogenetic relationships not inferrable from present data.

<sup>g</sup> Position among green algae remains unresolved; Steinkötter et al. (1994).

<sup>h</sup> Not currently accessioned in Genbank; from Bakker, Olson, and Stam (1995; fig. 1, "C. albida Roskoff").

<sup>i</sup> Gargas et al. (1995).

<sup>j</sup> The Genbank documentation erroneously indicates RNA.

<sup>k</sup> Garbary and Gabrielsen (1990).

#### Multiple Sequence Alignments

Multiple alignments were constructed with the aid of the PILEUP program of GCG version 8.1 (Genetics Computer Group 1994), CLUSTALW (Thompson, Higgens, and Gibson 1994), and CLUSTALV (Higgins, Bleasby, and Fuchs 1992): for (1) total 5.8S data; (2) green algal ITS2 sequences; and (3) green algal, fungal, and seed plant ITS2 sequences. Alignments and boundary approximations were adjusted manually using the GDE version 2.2 (Smith 1994) multiple sequence editor. For the 18S sequences, we constructed a manual alignment for the region spanning positions 24-1,783 of the Oryza sativa sequence (table 1). We truncated the alignment because the 5' and 3' regions were unavailable for many of the sequences. The broader ITS2 alignments were derived using CLUSTALV, which applies the "fast, dirty" algorithm used in CLUSTALW. We used a 10-base window to detect conserved motifs and high (8-15) gap-opening combined with low (2-4) gap-extension penalties to extend unconserved sequence as necessary to accommodate conserved motif alignment. To avoid input-order bias, we interspersed sequences from the different major groups. We excluded CLADO from the alignment because we found no primary sequence similarity to green algal or any other samples. A FASTA search inputting the CLADO ITS2 retrieved no other ITS2 sequences in the database.

# Phylogenetic/Comparative Sequence Analysis 5.85 Sequences

Phylogenetic analyses of a 5.8S data set modified from the alignment were carried out using maximum likelihood (ML) and maximum parsimony (MP) procedures in prerelease test versions d27–d45 of PAUP\* 4.0 (Swofford 1996). Delimitation of the 5.8S sequence was approximate because of apparent inconsistencies in the reported boundaries among the taxa (see Results). We excluded the highly diverged 5.8S sequence of the dinoflagellate Crypthecodinium (table 1) from the final analysis but determined in preliminary analyses that<sub>o</sub>it always paired with the Prorocentrum (PRORO) sequence. Using ML under the six-parameter general timereversible model with empirically determined base frequencies, we obtained preliminary estimates of the relative substitution rates and parameters describing among-site rate heterogeneity ( $\alpha$ , the shape parameter  $\overline{\mathfrak{D}}\mathfrak{f}$ the gamma distribution, Yang 1994; and  $\rho$ , the propertion of invariant sites; Swofford et al. 1996) over MP trees. An initial ML tree was generated using the initial rate parameter estimates and the heuristic search procedure in PAUP\* with the tree-bisection/reconnection (TBR) swapping algorithm. Substitution rate and heterogeneity parameters were reestimated over the initial ML tree, and these were used to perform TBR swapping on the initial tree to generate a second and final MaL tree. This recursive procedure was repeated, but the resulting ML topology was the same. Parameters were reestimated over the final tree and used to estimate pairwise distances. We also compared base compositional bias in angiosperms and green algae. Finally, the data set was analyzed and bootstrapped (Felsenstein 1985) 1,000 times using MP, and we derived MP and ML trees constrained for a topology supported by independent evidence.

#### 18S Sequences

The 18S sequences were examined primarily to compare their phylogenetic signal and evolutionary pattern with those of 5.8S. Phylogenetic analysis was limited to the MP procedure above, including bootstrapping and determining the parsimony score for a tree constrained for a plausible independently supported topology. We estimated, as above, substitution rates and siteto-site rate heterogeneity using ML over the parsimony tree and calculated pairwise distances using the ML parameters.

#### **ITS2** Sequences

Similarities among groups of ITS2 sequences were estimated using the guide tree feature in CLUSTALW, which yields a neighbor-joining tree based on similarity scores for pairs of sequences aligned optimally for given gap-opening and gap-extension penalties (see Hershkovitz and Zimmer 1996). Nominally, the guide tree topology functions as a phylogenetic template for successive sequence alignment, but we found that the guide tree provides a phenogram for comparing, without imposition of a fixed multiple alignment, a set of divergent and length-variable sequences. The most similar sequences form terminal bifurcations in the tree, and the internal branches cluster mutually more similar sequences. The branch lengths, however, are distance distorted in that the pairwise optimal alignments are noncommutative. For example, an optimal multiple alignment of the sequences AGAGAA, AGGAA, and AAGAA would not maintain the optimal pairwise alignments, but the guide tree branch lengths are based on the optimal pairwise scores. Nonetheless, the procedure will cluster sequences sharing alignable motifs relative to those that do not. For the ITS2 sequences, we generated trees using the "slow, accurate" algorithm, and, as in the multiple alignment described above, high (15–50) gap-opening combined with low (0-1) gap-extension penalties.

#### Results

5.8S and 18S Sequences Sequence Alignment

The 5.8S sequences are substantially alignable (fig. 1). In the case of apparent alignment ambiguities, we optimized the alignment for groups of sequences that were otherwise most similar. Aside from the 5' and 3' ends, the 5.8S sequences show few indels, particularly within chlorophytes and fungi. A length-variable region between positions 121 and 150 generally aligns better within versus between major taxonomic groups. In the empirically deduced yeast 5.8S secondary structure model (Yeh and Lee 1991), the length-variable region corresponds to a small stem structure that begins with the G residue at position 118 in the alignment and extends to C at 149. This stem is reproduced in other taxa folded according to the yeast model (CHLAM, Thompson and Herrin 1994, Fig. 7c; CLADO, Bakker, Olsen, and Stam 1995, Fig. 2) or other criteria (Suh, Thien, and Zimmer 1992; Ritland and Straus 1993; Troitsky and Bobrova 1986).

The 5' and 3' 5.8S boundaries reported for the various sequences are inconsistent. This might be attributable to a paucity of RNA-based sequences, misalignment with existing RNA-based sequences (e.g., CLA-DO; Bakker, Olsen, and Stam 1995; cf. Thompson and Herrin 1994), or clerical error in Genbank documentation (e.g., GONIU, cf. Coleman, Suarez, and Goff 1994; CANEL, cf. Suh, Thien, and Zimmer 1992). Another factor might be developmentally regulated 5.8S rRNA length variability. In yeast, a long form, possibly derived from an alternative pre-rRNA processing pathway (Henry et al. 1994; Lygerou et al. 1994; cf. Lindahl, Archer, and Zengel 1994; van Nues et al. 1994), extends 6–7 bases 5' from the short form. The *Saccharomyces* (SACCH) sequence in figure 1 represents the short (cf. van Nues et al. 1994; fig. 2A) and apparently more abundant (Henry et al. 1994) form; the long form extends to the 5' end of the figure 1 alignment or 1 base upstream. Long 5.8S forms reportedly exist in other eukaryotes as well (Henry et al. 1994).

Another potential alignment problem is sequencing error, particularly in the RNA-based sequences. Although these are useful for delineating 5.8S boundaries, they lack complementary-strand confirmation. Thompson and Herrin (1994), for example, found 12 errors in the RNA-based *Chlamydomonas* sequence in Genbank.

#### Properties of 5.8S Sequences

Table 2 shows that the observed proportion of variable sites in 5.8S is roughly twice that for 18S, although the absolute number of variable 5.8S sites is much less. For both data sets, the observed proportion of variable sites increases as more distantly related taxa are included. When estimated simultaneously,  $\alpha$  and  $\rho$  both generally increase as more distantly related taxa are included (table 3). When each parameter is estimated without allowance for the other,  $\alpha$  decreases and  $\rho$  increases with the inclusion of more distantly related taxa. For 5.8S, the  $\alpha$  and  $\rho$  estimates for the fungal plus chlorophyte subset differs markedly from those of the next larger and smaller taxon subsets. Obviously, the proportion of invariant sites cannot decrease as more distantly related taxa are excluded. In order to examine the effect of scoring 18S regions unalignable with angiosperms as "masing," we estimated parameters over an 18S parsimeny tree with the embryophytes pruned. This removes ca. 80 of the most variable positions from the alignment (table 2), although the fungi and protist sequences still have many sites scored as missing. In any case, the amongsite rate heterogeneity estimates for this taxon complement are  $\alpha/\rho = 0.58/0.26$ ,  $\alpha = 0.33$ , and  $\rho = 0.53$  (cf. table 3).

The 5.8S divergences appear to be roughly twice those of 18S (table 2). The 18S divergences are increasingly underestimated with increasing distance from angiosperms, however, because nonangiosperm 18S regions not alignable with angiosperms were scored as missing. We presume that poor alignability reflects extreme sequence divergence.

In both data sets, C–T transition rates are ca. twice that of A–G transitions and five times that of each transversion (table 3). The G–T rate among the 5.8S sequences is ca. half that of other transversions.

We examined base compositional patterns because we noticed that many A residues shared between the reported *Mimulus* sequences and green algae aligned with G residues in angiosperms. This suggested the possibility that similarities between the *Mimulus* and green algal sequences represented a convergent increase in adenosine. Table 4 shows base composition for subsets of positions varying between the nine angiosperm and six chlorophycean algae. These comparisons are not statistically rigorous because they ignore phylogenetic

		1		_						90	
				I_				V			
2	LYCOP	ACCA AACGACTCT	C GGCAACGGAT	ATCTCGGCTC	ICGCAICGAI	GAAGAACGIA	GCGAAAIGCG	AIACIIGGIG	IGAAIIGCAG	AAICCCGIGA	
3	CUCUR	taaC A								GC	
4	triti	CACc			• • • • • • • • • • •	• • • • • • • • • • •		c		c	
5	LACTU	CaCA A	• • • • • • • • • • • • • • • • • • • •	T	A	• • • • • • • • • • •	A	• • • • • • • • • • •	• • • • • • • • • • •	• • • • • • • • • • •	
7	vicia	atla A	• • • • • • • • • • • • •	<b>.</b>							
8	CANEL	TCAA G.T								c	
9	potam	AT C.T		T					G		
10	ephed	СТ Т									
11	gnetu	CC C	A		TT	••••	<u></u>		•••••		
12	PINUS	*GA A	T	•••••	.T.TTA	• • • • • • • • • • •	• • • • • • • • • • • •	••••	• • • • • • • • • • • •	• • • • • • • • • • •	
14	TAYUS	CT TGGC	• • • • • • • • • • • • •		CA						
15	marsi	.G	. A		.T			GA		TC	_
16	osmun		. A	T	.TA	c.		G	c	TC	
17	mnium	.TAC	. A	T	.TA	C.		G		TC	
18	CHLAM	ccaa gacA	. AA	T	G		• • • • • • • • • •	G	• • • • • • • • • • •	ATA	
19	GONIU	ccaa GA	. AA	<b>T</b>	AG	c.	•••••	G	•••••	ATA	
20	MILIPH	CCAA GA	. AA	т	AG	·····.	•••••				D
22	MIURF	CaAA G. AG.	G TA	T				G		<b>T</b>	Š
23	SPERM	caaa gacA	. AA	T	T.A	c.		G		<b>T</b>	nl
24	TETRA	*A GA	. AA	<b>T</b>	.TAA	c.		<u>G</u>	• • • • • • • • • • •	<b>T</b>	oa
25	CHLOR	aaaT GA	. AA	T	C	c.	A 60	G	• • • • • • • • • • •	T	de
20	UETED	_tyat aytaaca	32 La	<u> </u>	C.A	<u> </u>		AC A		T.A	<u> </u>
28	VOLVA	taaa ta.AT.	. AA	Ст				AGA		T.A	fro
29	SCLER	tagt taAAT.	. AA	СТТ	.G	C.		AGA		T.A	Ä
30	sacch	tatt aaAAT.	. AA	Стт.	· · · · · · · · · · · ·	c.	• • • • • • • • • • •	GA	• • • • • • • • • •	<u>T</u> . <u>.</u>	ht
31	BIPOL	ttat ta.AT.	. AA	Стт	.G	c.	• • • • • • • • • • •	G	••••••••••	T.A	tp
32	CYANT	CGAN AGANC	AA	C	TTCA	<u> </u>	- A A	A. TG	A	GT.TT	- 0
34	SARCO	aaat ta.AA	T .AG	GT	CAA	cc.		.AGA	cc		a)
35	PHYTO	AT AG.AT.	. AG	GA	G.A	CT	c	GA	c	GT.A	â
36	SCYTO	GTTG T.AAT.	. A	GT	C.AA	c.	• • • • • • • • • <u>•</u> •	GT	cc	A	de
37	crypt	AT.	AG.T	тстт	AGA.C	GG.ACT	····T.	TGG.A	••••	GCATG	⊒.
30	FRORU	ycat t	. A	G	GAA						<u>.</u>
											P
									170		oup.c
		91   TT/	τ/	V			. V'		170		oup.cor
1	ARABI	91   ACCATCGAGT CTT	I'	<u></u>	CAAGCC TTCT	-GGCCG	. <u>V'</u>	TGCCTG GGTG	170   TCACAA		pup.com/r
1 2	ARABI LYCOP	91   Accatcgagt CTT	<u>1'</u> TGAACGC AAGT	 IGCGCC C	CAAGCC TTCT GAAGCC ATTT	-GGCCG	. <u>V'</u> Agggca cgtc A	TGCCTG GGTG 	170   TCACAA ACG-		oup.com/mb
1 2 3	ARABI LYCOP CUCUR	91   ACCATCGAGT CTT C	L' TGAACGC AAGT	 IGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC TTCT	-GGCCG	. <u>v'</u> Agggca cgtc A	ТGCCTG GGTG С. 	170   TCACAA ACG- ACGC		pup.com/mbe/;
1 2 3 4	ARABI LYCOP CUCUR triti	91   ACCATCGAGT CTT C	L' TTGAACGC AAGT	<u> </u>	CAAGCC TTCT GAAGCC ATTT GGAGCC ATCT GAGGCC ATCC	-GGCCG -GGCCG -GGCCG -GCCCG	. <u>V'</u> AGGGCA CGTC A	TGCCTG GGTG C. C.	170   TCACAA ACG- ACGC ACGC		oup.com/mbe/art
1 2 3 4 5 6	ARABI LYCOP CUCUR triti LACTU SILEN	91   <u></u> ACCATCGAGT CTT C	I' TGAACGC AAGT		CAAGCC TTCT GAAGCC ATTT GAGGCC TTCT GAAGCC ACTC GAAGCC ATCC GAAGCC - TTC	-GGCCG -GGCCG -GGCCG -GGCCG -GGCTG	. <u>V'</u> AGGGCA CGTC A AC. AC. AC.	TGCCTG GGTG C. C. C. C.	170   TCACAA .ACG- .ACGC .ACGC .ACGC .ACGC		pup.com/mbe/articl
1 2 3 4 5 6 7	ARABI LYCOP CUCUR triti LACTU SILEN vicia	91   ACCATCGAGT CTT C	I' TTGAACGC AAGT	 FGCGCC C C C C C	CAAGCC TTCT GAAGCC ATTT GAGGCC TTCT GAGGCC ATCC GAAGCC ATCC GAAGCC ATTA	-GGCCG -GGCCG -GGCCG -GGCCG -GGCTG -GGTTG	. <u>V'</u> AGGGCA CGTC AAC. AC. AC.	TGCCTG GGTG C. C. C. C.	170   TCACAA ACG- ACGC ACGC ACGC ACGC ACGT		pup.com/mbe/article/
1 2 3 4 5 6 7 8	ARABI LYCOP CUCUR triti LACTU SILEN VICIA CANEL	91   ACCATCGAGT CTT C	L' TTGAACGC AAGT	 IGCGCC C C C C C C	GAAGCC ATTT GGAGCC ATTT GAGGCC ATCT GAAGCC ATCC GAAGCC ATCC GATGCC ATTA GAGGCC ACTA	-GGCCG -GGCCG -GGCCG -GGCCG -GGCTG -GGTTG -GGTTG	. <u>V'</u> AGGGCA CGTC A A A A A A A	TGCCTG GGTG C. C. C. C. C.	170   TCACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC		pup.com/mbe/article/13,
1 2 3 4 5 6 7 8 9	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam	91   ACCATCGAGT CTT C	L' TTGAACGC AAGT	 TGCGCC C C C C C C C C C	GAAGCC TTCT GAAGCC ATTT GAGGCC ATCT GAAGCC ATCC GAAGCC ATCC GATGCC ATTA GAGGCC ACTA TAAGCT TCCG	-GGCCG -GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG	. <u>V'</u> AGGGCA CGTC AAC. AC. AC. AA.C. AA.C.	TGCCTG GGTG 	170    ACACAA .ACGC- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC		oup.com/mbe/article/13/9/
1 2 3 4 5 6 7 8 9 10	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed	91   ACCATCGAGT CTT C	1/ TIGAACGC AAGT		GAAGCC TTCT GGAGCC ATTT GAGGCC ATTT GAGGCC ATCT GAAGCC ATCC GAAGCC ATTA GAGGCC ACTA TAAGCT TCCG GAAGCCTC	-GGCCG -GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCCG -GGCCG	. <u>V'</u> AGGGCA CGTC AC. AC. AC. AC. AT.C. AT.C. AA.C.	TGCCTG GGTG 	170   CACAA .ACGC- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .CCAA .GCAA		pup.com/mbe/article/13/9/12
1 2 3 4 5 6 7 8 9 10 11	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PTNUS	91   ACCATCGAGT CTT C T T T TA	I' TIGAACGC AAGT		CAAGCC TTCT GAAGCC ATTT GGAGCC ATTT GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ACTA GAAGCCTC G-AGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG	. <u>V'</u> AGGGCA CGTC AC. AC. AC. AA. AA. AA. AA. AA. AA. AA. AA. AA.	TGCCTG GGTG C. C. C. 	170   TCACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA		pup.com/mbe/article/13/9/1276
1 2 3 4 5 6 7 8 9 10 11 12 13	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA	91   ACCATCCAGT CTT C	L' TIGAACGC AAGT 		CAAGCC TTCT GAAGCC ATTT GGAGCC TTCT GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ACTA GAAGCCTC GAAGCCTC GAAGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG	 AGGGCA CGTC A A A A	TGCCTG GGTG C. C. C. C. 	170   TCACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACAT .ACgc .ACAT .GCAA .GCAA .GCAA .GCAA .GCAT CC		pup.com/mbe/article/13/9/1276/9
1 2 3 4 5 6 7 8 9 10 11 12 13 14	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS	91   ACCATCCAGT CTT C	IL TGAACGC AAGT 		CAAGCC TTCT GGAGCC ATTT GGAGCC ACTC GAAGCC ACTC GAAGCC ACTA GAGGCC ACTA GAGGCC ACTA GAGGCCTC GAGCCCTC GAGCCCTC GAGCCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG	. <u>V'</u> AGGGCA CGTC AC. AC. AC. AA.C. AA.C. AA.C. AA.C. AA.C. AA.C. AA.C. AA.C.	TGCCTG GGTG C. C. C. C. C. C. C. C. C. C. C. C. C.	170     CACAA .ACGC .ACGC .ACGC .ACGC .ACAT .ACgc .ACAT .GCAA .GCAA .GCAA .GCAA .GCAT CC .GCAC		pup.com/mbe/article/13/9/1276/991
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS marsi	91   ACCATCGAGT CTT CT. T. T. T. T. T. T. T. T. T. T. T.	L' TTGAACGC AAGT	 TGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTT GAGGCC ATCC GAAGCC ATCC GATGCC ATTA GAGGCC ATTA GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG	. <u>V'</u> AGGGCA CGTC A	TGCCTG GGTG 	170   CACAA .ACGC- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAT .GCAA .GCAT .CCAC		oup.com/mbe/article/13/9/1276/99146
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS marsi osmun	91   ACCATCGAGT CTT C	1/ TIGAACGC AAGT 		CAAGCC TTCT GAAGCC ATTT GGAGCC ATTCT GAAGCC ATCC GAAGCC ATCC GATGCC ATTA GAGGCC ACTA GAGGCCTC GAAGCCTC GAGGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGTCG -GGCCG -GTCCG -GTCCG	V'	TGCCTG GGTG C. 	170     .ACG- .ACGC- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAT CC .GCAC .CAC .CAC		oup.com/mbe/article/13/9/1276/991469
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17	ARABI LYCOP CUCUR LACTU SILEN VICIA CANEL DOLAM ephed gnetu PINUS PICEA TAXUS marsi osmun milum curiu	91   ACCATCGAGT CTT C T	1/ TGAACGC AAGT 	 TGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC C ATTT GAGGCC ATCC GAAGCC ATCC GAGCC ATTA GAGGCC ACTA GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGTCG -GGCCG -GTCCG -GTCCG -GTCCG	V'	TGCCTG GGTG C. 	170     .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAA .GCAC .CCC .CC		oup.com/mbe/article/13/9/1276/991469 by
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18	ARABI LYCOP CUCUR triti LACTU SILEN VICIA CANEL potam ephed gnetu PINUS PICEA TAXUS marsi osmun mnium CHLAM GONTU	91   ACCATCCAGT CTT C T	1/ TGAACGC AAGT 		CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ATCC GAAGCC ATTA GAGGCC ACTA GAGCCTC GAGCCTC GAGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG	 AGGGCA CGTC AC. AC. AC. AC. AA.C.	TGCCTG GGTG C. C. C. C. C. C. C. C. C. C. C. A.C. C.TTA. A.C. C A.C.	170     CACAA .ACGC .ACGC .ACGC .ACGC .ACAT .ACgc .ACAT .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAC .CCC .CC		oup.com/mbe/article/13/9/1276/991469 by g
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	ARABI LYCOP CUCUP triti LACTU SILEN SILEN SILEN POTAM ephad gnetu PINUS PICEA TAXUS marsi osmun CHLAM GONIU PANDO	91   ACCATCCAGT CTT C	<u>I</u> TGAACGC AAGT 	 TGCGCC C	CAAGCC TTCT GGAGCC ATTT GGAGCC ACTC GAAGCC ACTC GAAGCC ACTC GAAGCC ACTA GAGGCC ACTA GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTCC GAGGCCTCC GAGGCCTCC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GGCCA	V'	TGCCTG GGTG C. 	170     ACGCAA .ACGC .ACGC .ACGC .ACGC .ACAT .ACGC .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAC .CAC	+26 bases)	oup.com/mbe/article/13/9/1276/991469 by gue
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21	ARABI LYCOP CUCUP triti LACTU SILEN Vicia CANEL POCAM ephed gnetu PINUS PICEA TAXUS MATSI OSMUN MATSI CHLAM GONIU PANDO MIURH	91 	1/ TTGAACGC AAGT 	 FGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTCT GAAGCC ATCC GAAGCC ATCC GATGCC ATTA GAGGCC ATTA GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCTTC GAGGCTTC GAGGCTTC GAGGCTTC GAGGCTTC GAGGCTTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GCCA -GGCCA -GGCCA	.         V'           AGGGCA         CGTC           A.         .           A.         T.           A.         T.           A.         T.           A.         T.	TGCCTG GGTG C. 	170     .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACAT .ACgc .ACAT .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCAT .GCGT .ACC .GCGT .ACC .GCGT .ACC .GCGT .ACC .GCGT .ACC .GCAT .ACC .GCAT .ACC .GCAT .ACC .GCAT .ACC .ACC .ACC .ACC .ACC .ACC .ACC	+26 bases)	pup.com/mbe/article/13/9/1276/991469 by guest
1 2 3 4 5 6 7 8 9 10 11 13 14 15 16 17 18 19 20 21 22	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PINUS PINUS Marsi marsi osmun mnium CHLAM GONIU PANDO MIURF	91 	1/ TIGAACGC AAGT 	 TGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTT GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGGCCTC GAAGCCTC GAGCCTC GAGCCTC GAGGCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GCCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA	V'_ AGGGCA CGTC A	TGCCTG GGTG C. 	170   CACAA . ACG- . ACGC . GCAA . GCAA . GCAA . GCAC . CCC . ACC . GGGT TA . GGGT TA . GGGT TA . GGGT TA . GGGT TA	+26 bases)	oup.com/mbe/article/13/9/1276/991469 by guest or
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24	ARABI LYCOP CUCUR LACTU SILEN VICIA CANEL POLAM GPHEU PILES TAXUS MAININ MIUS SPERM TETES	91   ACCATCGAGT CTT C T T T T T T T T T T T T T T T A T A	1/ TGAACGC AAGT 	 TGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ATCC GAAGCC ATTA GAGGCC ACTA GAGGCCTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GCCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA	.         V'	TGCCTG GGTG C. 	170     CACAA .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACAT .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAT .GCAC .CCC .CCC .GCGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA	+26 bases)	oup.com/mbe/article/13/9/1276/991469 by guest on 2
1 2 3 4 5 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25	ARABI LYCOP CUCUR LACTU SILEN VICIA CANEL POTAM ephed gnetu PICEA TAXUS marsi osmun mnium CHLAM GONIU PANDO MIURH MIURF SPERM TETRA CHLOR	91   ACCATCCAGT CTT C T T T T T T T T T T T T T T T T A	1/ TGAACGC AAGT 	 TGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ATTA GAGGCC ATTA GAGGCC ATTA GAGGCC -TCC GAGCC -TC GAGGCC -CTC GAGGCC -CTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GCCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA	 AGGGCA CGTC AC. AC. AC. AC. AA.C. AA.C. AA.C. AA.C. AA.C. AA.C. AT.A. AT.A.	TGCCTG GGTG C. 	170     CACAA .ACGC .ACGC .ACGC .ACGC .ACGC .ACAT .ACGC .ACAT .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAC .CTC .CTC .CTC .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGT TT .GGT CA	+26 bases)	oup.com/mbe/article/13/9/1276/991469 by guest on 21 .
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 9 20 21 22 23 24 25 26	ARABI LYCOP CUCUP triti LACTU SILEN Vicia CANEL POTAM ephed gnetu PINUS PICEA TAXUS Marsi Osmun MIUR MIURF SPEEM MIURF SPEEM MIURF SPEEM	91   ACCATCGAGT CTT CT. T. T. T. T. T. T. T. T. T. T. T. T. T. T. T. 	<u>I</u> TIGAACGC AAGT TA TA TA TA TA TA TA TA TA TA TA TA TA	V FGCGCC C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ACTC GAAGCC ACTA GAGGCC ACTA GAGGCC ACTA GAGGCCTCC GAGGCCTCC GAGCCTCC GAGCCTCC GAGCCTCC GAGCCTCC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GTCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA	V'	TGCCTG GGTG C. C. C. C. C. C. C. C. C. C. C. C. A.C. A.C. C. A.C. C. A.C.	170     ACGC- .ACGC- .ACGC .ACGC .ACGC .ACGC .ACGC .GCAT .GCAA .GCAA .GCAA .GCAA .GCAC .CTC .CAC .CAC .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGTT TAAA	+26 bases) TGGC	pup.com/mbe/article/13/9/1276/991469 by guest on 21 Au
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 19 20 21 22 23 24 25 26 27	ARABI LYCOP CUCUR triti LACTU SILEN Vicia CANEL potam ephed gnetu PINUS FICEA TAXUS MATSI OSMUN MIURA MIURA MIURA MIURA MIURA SPERA CHLOR CLADO HETER	91 	1/ TGAACGC AAGT 	V           IGCGCC           C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGGCCTC GAAGCCTC GAGCCTC GAGCCTC GAGCCTC GAGGCTC GAGGCTC GAGGCTTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA	.         V'	TGCCTG GGTG C. 	170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGA .GCAA .GCAA .GCAA .GCAT .GCAT .GCAT .GCAT .GCAT .GCGT .GCGT .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGT TA .GGT .GCAA	+26 bases) TGGC	pup.com/mbe/article/13/9/1276/991469 by guest on 21 Augu
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 9 20 21 22 23 26 27 28	ARABI LYCOP CUCUR Triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PINUS MILEA TAXUS marsi osmun milum CHIAM GONIU PANDO MIURF SPERM TETRA CHLOR CLADO HETER VOLVA	91  ACCATCGAGT CTT C	L/ TGAACGC AAGT 	V           IGCGCC           T           C           T           C           T           C           T           C           T           C           T           C <tr< th=""><th>CAAGCC TTCT GAAGCC ATTT GAAGCC ATTT GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGGCC -TTC GAAGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TTC GAGGCC -TTC GAGGC - TTC GAGGC - TTC GAGGC - CTC GAGGC - CTC GAGGC - CTC GAGGC - CTC GAGGC - CTC GAGCC - CTC GACC - CTC</th><th>-GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA </th><th>.         V'</th><th>TGCCTG GGTG C. </th><th>170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAA .GCAT .GCAC .CTC .CTC .GGGT TA .GGGT TA .GGTT TAAA .GGTT TAAA .GT- GAA .GT- GAA</th><th>+26 bases) TGGC</th><th>oup.com/mbe/article/13/9/1276/991469 by guest on 21 Augus</th></tr<>	CAAGCC TTCT GAAGCC ATTT GAAGCC ATTT GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGGCC -TTC GAAGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TTC GAGGCC -TTC GAGGC - TTC GAGGC - TTC GAGGC - CTC GAGGC - CTC GAGGC - CTC GAGGC - CTC GAGGC - CTC GAGCC - CTC GACC - CTC	-GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA 	.         V'	TGCCTG GGTG C. 	170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAA .GCAT .GCAC .CTC .CTC .GGGT TA .GGGT TA .GGTT TAAA .GGTT TAAA .GT- GAA .GT- GAA	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991469 by guest on 21 Augus
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20 21 22 23 24 25 26 27 28 29 20	ARABI LYCOP CUCUR triti LACTU SILEN SILEN SILEN SILEN POTAM ephad gnetu PINUS PICEA TAXUS MAIN CHIAM GONIU PANDO MIURH MIURH MIURH MIURH SPERM CHLOR CLADO HETER VOLVA SCLER	91   ACCATCCAGT CTT C T A T A T A T A T A T A T A T A A A A A A A A A A A A A A A A A T A A A A T A A A A T T A T	L' TGAACGC AAGT 		CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ACTC GAAGCC ATTA GAGGCC ACTA GAGCC -TTC GAGCC -TTC GAGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGC - TTC GAGGC - TTC GAGCC - TTC GACC - TTC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GCCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA	 AGGGCA CGTC AC. AC. AC. AC. AA.C. AA.C. AA.C. AA.C. AA.C. AA.C. AT.C. AT.C. A.A.T.A.A.T.A. A.A.T.A.A.T.A. A.A.T.A.A.T.A. A.A.T.A.A.T.A. A.A.T.A.A.T.A. A.A.T.A.A.T.A. A.A.T.C. A.A.T.C. A.A.T.C. A.A.T.C. A.A.T.C.	TGCCTG GGTG C. 	170   TCACAA .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAC .CTC .CTC .CTC .GGGT TA .GGGT TA .GGTT TAAA .GGTT TAAA .GTT- GAA .ATTC CA .ATTC CA .ATTC CA .ATTC CA .CTC	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	ARABI LYCOP CUCUP TILI LACTU SILEN Vicia CANEL POTAM ephed gnetu PINUS PICEA TAXUS MATSI OSMUN MIURH MIURF SPEEM MIURH MIURF SPEEM CHLOR CLADO HETER SCLER SACCH SCLER SACCH	91 	1/ TTGAACGC AAGT 	V           IGCGCC           T           C           T           C           T           C           T           C           T           C           T           C           T           C           T	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ACTC GAAGCC ACTA GAGGCC ATTA GAGGCC CATTA GAGGCCTC GAGCCTC GAGGCTC GAGCCTC C C AC 	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA 	.         V'	TGCCTG       GGTG	170     ACGC- .ACGC- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAA .GCAT .GCAT .GCAT .GCGT .CTC .CTC .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGTT TAAA .GGTT TAAA .GGTT CA .GGTT CA .GCTT CA .GGTT CA .GGTT CC .GGTT CA .GGTT CA .GGTA	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 202
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 32 24 25 26 27 28 29 30 31 32	ARABI LYCOP CUCUR triti LACTU SILEN Vicia CANEL POCAM ephed gnetu PINUS FICEA TAXUS MATSI OSMUN MATSI OSMUN PANDO MIURF SPERM MIURF SPERM MIURF SPERM MIURF SPERM CHLOR CLADO HETER VOLVA SCLER SACCH BIPOL GLOMU	91 	1/ TGAACGC AAGT TGAACGC AAGT T T T T T T T T T T T T T T T T T T	V           IGCGCC           C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ATCC GAAGCC ATCC GAAGCC ATCA GAGGCC ATTA GAGGCCTC GAAGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCTC GGT -ATT GGT -ATT	-GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA 	.         V'           AGGGCA         CGTC           A         C.           A         A.           A         A.           A         T.C.           A         T.C.           A.         T.C.           A.         T           A.         T           A.         T           A.         T           A.         T           A.A         T.C.           A.A         T.C.	TGCCTG       GGTG	170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAT .GCAT .GCAT .GCAT .GCGT .GCGT .GCGT .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGGT TA .GGT TA .GGT TA .GGT TA .GTT .GT	+26 bases) TGGC	pup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 9 20 21 22 23 24 25 26 27 28 29 30 31 32 33	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS	91  ACCATCGAGT CTT CT. T. T. T. T. T. T. T. T. T. T. T. T. T. T. T. T. 	L/ TGAACGC AAGT 	V           IGCGCC           T           C           T           C           T           C           T           C           T           C	CAAGCC TTCT GAAGCC ATTT GAAGCC ATTT GAAGCC ATCC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGGCC ATTA GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCCTC GAGGCTC GAGCCTC GAGGCTC G GAGCTC G GAGCTC	-GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GTCCG -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA 	V'	TGCCTG       GGTG	170   ACG- ACGC ACGC ACGC ACGC ACGC ACGC ACGC ACGC ACGC ACGC CCC CC	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20 21 22 23 24 25 26 27 28 29 30 31 32 33 33 34	ARABI LYCOP CUCUR CUCUR SILEN VICIA CANEL POLAM GPHOL PINUS MILUR PINUS MIUR SPERM TAXUS MIUR SPERM TETRA CHLOR CLADO HETER CLADO HETER SACO CYANI	91	1/ TGAACGC AAGT 	V           IGCGCC           C	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGGCC -TC GAGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGCC -TC GAGGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAC GAC -TC GAGGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGGC -TC GAGGC -TC GAGGC -TC GAGGC -TC GAGGC -TC GAGGC -TC GAGGC -TC GAGGC -TC GAGC -TC GAGGC -TC GAGGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -TC GAGC -T	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GCCCA -GGCCA 	.         V'	TGCCTG       GGTG	170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAT CC .GCAT CC .GCAT CC .GCAT .GCGT TA .GGGT CA .GGTT TAAA .GTT CAA .ATTT CAA .GGTA GAA .ATTT CAA .ATTT CAA	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991468 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 35 35 35 35 35 35 35 35 35	ARABI LYCOP CUCUR triti LACTU SILEN Vicia CANEL Potam ephed gnetu gnetu FINUS FICEA TAXUS Marsi osmun milum MIURH MIURF SPEEM MIURH MIURF SPEEM CHLAM CHLAM CHLAM CHLAM SCLER sacch BIPOL GLADO CYNNI SARCO PHTO SARCO	91	1/ TTGAACGC AAGT 	V 	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ACTC GAAGCC ATCC GAAGCC ATTA GAGGCC ATTA GAGCC -TTC GAGCC -TTC GAGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGCC -TTC GAGGC ATT- -GGT ATT- -GGT ATT- -GGT ATT- -GGT ATT- -GGT ATT-	-GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GTCCG -GCCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -GGCCA -CGCC A -CCC A -CCCA A -CCCA G -CCCA G -CCCA G	.         V'	TGCCTG       GGTG	170     CACAA .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAA .GCAC .CTC .CTC .GGGT TA .GGGT CA .GGTT TAAA .GTT TAAA .GTT GAA .ATTC ca .ATTC ca .ATTC ca .CTC .CTC .CTC .GGT CA .GGT TA .GGT CA .GGT CA .GGT CA .GGT CA .GTT TAAA .GTT CA .ATTC ca .ATTC ca .CTC .CTC .CTC .CTC .CCC .CCC .GGT CA .GGT CA .GGT CA .GCT CA .GC	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 9 10 11 12 13 14 15 16 17 18 19 20 21 22 24 22 24 22 24 22 24 22 24 30 31 33 33 33 34 35 36 37	ARABI LYCOP CUCUR triti LACTU SILEN Vicia CANEL POTAM ephed gnetu PINUS PICEA TAXUS TAXUS CALON CHLAM GONTU PANDO CHLAM GONTU PANDO HETER CHLOR CLADO HETER SCLER SACCO PHYTO SCIVO SCIVO	91	1/ TTGAACGC AAGT TTGAACGC AAGT TT TT TT TT TT TT TT TT TT TT TT TT T	V           IGCGCC           T           C           T           C           T           C           T           C           T           C           T           C           T           C           T           C           T           C           T           C <th>CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ATCC GAAGCC ATCC GAAGCC ATCC GAAGCC -TCC GAGCC -TC GAGCC -TC GAGCC -TC GAGCC -TC GAGCC -TC GAGGC -TC -TC GAGT ATT- -GGT ATT- -GGT ATT- -GGT ATC -GGT ATC -GGT ATC -GGT ATC -GGT ATC -GGT ATC -GGT ACC</th> <th>-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GGCCA</th> <th>.         V'           AGGGCA         CGTC           A         C.           A         C.           A         C.           A         C.           A         C.           A         C.           A         A.           A         A.           A         A.           A         T.C.           A         T.           A         T.           A         T.           A         T.           A         T.           A         T.           A.A         T.           G.A</th> <th>TGCCTG       GGTG        </th> <th>170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAC .CCC .CAC .CAC .GGGT TA .GGGT TA .GGTA TA .GTT TAAA .GT- GAA .ATTT caa .ATTT caa .ATTT caa .ATTT caa .GTTA gaa .CTCt tga .CGCA cat .CGCA CAA .GTA CAA .GTA GAA .ATTT CAA .GTA GAA .GTA GAA .ATTT CAA .GTA GAA .CTC CAA .CTC CAA .CTC CAA .GTTA GAA .ATTT CAA .GTTA GAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA</th> <th>+26 bases) TGGC</th> <th>oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2022</th>	CAAGCC TTCT GAAGCC ATTT GGAGCC ATTC GAAGCC ATCC GAAGCC ATCC GAAGCC ATCC GAAGCC -TCC GAGCC -TC GAGCC -TC GAGCC -TC GAGCC -TC GAGCC -TC GAGGC -TC -TC GAGT ATT- -GGT ATT- -GGT ATT- -GGT ATC -GGT ATC -GGT ATC -GGT ATC -GGT ATC -GGT ATC -GGT ACC	-GGCCG -GGCCG -GGCCG -GGCTG -GGCTG -GGCTG -GGCCG -GGCCG -GGCCG -GGCCG -GGCCG -GTCCG -GTCCG -GGCCA	.         V'           AGGGCA         CGTC           A         C.           A         C.           A         C.           A         C.           A         C.           A         C.           A         A.           A         A.           A         A.           A         T.C.           A         T.           A         T.           A         T.           A         T.           A         T.           A         T.           A.A         T.           G.A	TGCCTG       GGTG	170   CACAA .ACG- .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .ACGC .GCAA .GCAA .GCAA .GCAA .GCAC .CCC .CAC .CAC .GGGT TA .GGGT TA .GGTA TA .GTT TAAA .GT- GAA .ATTT caa .ATTT caa .ATTT caa .ATTT caa .GTTA gaa .CTCt tga .CGCA cat .CGCA CAA .GTA CAA .GTA GAA .ATTT CAA .GTA GAA .GTA GAA .ATTT CAA .GTA GAA .CTC CAA .CTC CAA .CTC CAA .GTTA GAA .ATTT CAA .GTTA GAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA .CTCT CAA	+26 bases) TGGC	oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2022

FIG. 1.—Provisional alignment of 5.8S rDNA sequences. Taxon acronyms, definitions, and sequence documentation are given in table 1. Acronyms in lowercase letters denote RNA-based sequences. Upper- and lowercase bases denote, respectively, the 5.8S and flanking ITS regions according to Genbank documentation, except for the SCYTO sequence, in which the boundaries were not indicated. Where sequence data are available the 5'-end of the alignment extends to include the 5'-most bases reported for the CYANI sequence, and the 3' ends are, except for the green algae, extended to include the 3'-most bases reported within each major taxonomic group. The anomalous 3' extensions for PANDO (not shown) and CLADO are presumably erroneously included in the ITS2 sequence. The first 3 bases of the CANEL sequence are actually ITS1 (Suh, Thien, and Zimmer 1992). The SACCH 5.8S is RNA based, but the flanking sequence is DNA based (table 1). The 5' flanking sequence, however, is part of a "long" 5.8S form (see text). Position 1 in the alignment corresponds to character 1 in the phylogenetic analysis data set (appendix). The length-variable region (ca. 120–150) is aligned optimally within major groups but arbitrarily between them. In the yeast 5.8S secondary structure model, the superposed regions I–V pair with I'–V'. Pairing relationships in the length-variable region (V–V') probably varies among the taxa.

Table 2Divergence Matrices for 5.8S and 18S rDNA Sequences

	Sites	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
5.8S sequences																<u> </u>
1 ARABI	162		•	•	•	•										
2 LYCOP	161	0.03		•	•	•	•	•	•	•	•	·	·	•	•	•
4 TRITI	162	0.03	0.04	0.07	· 			•		•	:	•		•	÷	:
5 LACTU	162	0.08	0.05	0.08	0.07											
6 SILEN	159	0.05	0.02	0.05	0.05	0.03			•	•						
7 VICIA	162	0.08	0.06	0.11	0.10	0.11	0.07			•	•	•	•	•	·	•
9 POTAM	161	0.12	0.07	0.10	0.04	0.08	0.07	0.09	0.14	•	÷	•	•	•	÷	•
10 EPHED	158	0.08	0.06	0.11	0.09	0.12	0.07	0.12	0.13	0.17						
11 GNETU	157	0.08	0.06	0.11	0.09	0.13	0.09	0.10	0.11	0.17	0.06				•	•
12 PINUS	158	0.14	0.11	0.17	0.13	0.16	0.10	0.13	0.18	0.24	0.12	0.10		•		÷
14 TAXUS	156	0.16	0.14	0.18	0.13	0.19	0.13	0.13	0.18	0.23	0.12	0.12	0.04	0.15		Do
15 MARSI	155	0.20	0.17	0.20	0.15	0.23	0.17	0.20	0.20	0.25	0.17	0.17	0.17	0.19	0.18	h
16 OSMUN	156	0.23	0.19	0.21	0.15	0.24	0.19	0.23	0.21	0.27	0.17	0.21	0.21	0.25	0.22	0,87
17 MNIUM	157	0.23	0.20	0.23	0.18	0.26	0.20	0.22	0.22	0.28	0.22	0.25	0.21	0.22	0.22	0,90
19 GONIU	160	0.28	0.22	0.28	0.23	0.29	0.23	0.28	0.27	0.33	0.21	0.21	0.25	0.30	0.28	0.96
20 PANDO	160	0.28	0.22	0.28	0.23	0.30	0.23	0.30	0.27	0.34	0.23	0.23	0.27	0.30	0.34	0.38
21 MIURH	160	0.24	0.19	0.25	0.19	0.26	0.20	0.26	0.24	0.30	0.21	0.21	0.22	0.26	0.27	0.14
22 MIURF	160	0.29	0.24	0.30	0.24	0.32	0.25	0.32	0.29	0.35	0.29	0.23	0.25	0.30	0.33	0.59
23 SPERM 24 TETRA	159	0.24	0.20	0.26	0.20	0.27	0.21	0.26	0.25	0.32	0.23	0.19	0.20	0.23	0.27	0.13
25 CHLOR	160	0.30	0.25	0.31	0.21	0.30	0.25	0.27	0.20	0.34	0.24	0.23	0.23	0.28	0.31	0,003
26 CLADO	159	0.49	0.23	0.53	0.48	0.47	0.44	0.50	0.50	0.58	0.44	0.42	0.46	0.48	0.53	0.39
27 HETER	157	0.22	0.46	0.27	0.22	0.27	0.24	0.26	0.28	0.32	0.18	0.19	0.25	0.28	0.29	0.37
28 VOLVA 29 SLCER	160	0.26	0.24	0.31	0.25	0.31	0.27	0.28	0.32	0.31	0.26	0.27	0.31	0.36	0.36	0.98
30 SACCH	158	0.25	0.30	0.33	0.29	0.33	0.24	0.25	0.33	0.33	0.30	0.31	0.34	0.38	0.38	0.29
31 BIPOL	160	0.26	0.24	0.29	0.24	0.29	0.25	0.27	0.30	0.30	0.24	0.26	0.28	0.32	0.34	0.68
32 GLOMU	160	0.33	0.34	0.39	0.34	0.40	0.35	0.36	0.36	0.39	0.31	0.31	0.38	0.41	0.40	0.26
33 CYANI	146	0.47	0.47	0.49	0.51	0.53	0.48	0.48	0.54	0.68	0.47	0.43	0.45	0.50	0.50	0.38
35 PHYTO	141	0.42	0.44	0.31	0.47	0.30	0.43	0.46	0.54	0.51	0.37	0.38	0.45	0.49	0.53	0.003
36 SCYTO	145	0.46	0.46	0.52	0.45	0.52	0.47	0.46	0.53	0.52	0.46	0.50	0.51	0.52	0.55	0.32
37 PRORO	147	0.54	0.54	0.59	0.53	0.56	0.54	0.55	0.60	0.61	0.55	0.67	0.61	0.65	0.66	0.00
18S sequences																13/9
1 PRUNU18	1733	0.02	•	•	•	•	•	•	•	•	·		•	·	•	9/.1.2
3 DRIMY18	1728	0.03	0.04		•	•	:	•	÷	÷		•	•			276
4 GOSSY18	1717	0.03	0.03	0.04	_											9/9
5 RIBES18	1725	0.03	0.03	0.03	0.03	•	•		•	•						01.2
6 SINAPI8	1736	0.03	0.04	0.04	0.03	0.04		·	·		•	•				69-
8 HYDRO18	1609	0.03	0.03	0.03	0.03	0.03	0.04	0.04		•		•	•		•	by
9 ORYZA18	1744	0.04	0.04	0.05	0.04	0.05	0.06	0.05	0.06	_						ĴŪ/
10 LYCOP18	1735	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.04	0.04	_	•				Jes
11 EPHED18	1712	0.09	0.08	0.09	0.09	0.08	0.09	0.09	0.10	0.09	0.08			·	·	t o
13 PINUS18	1737	0.09	0.07	0.08	0.07	0.07	0.08	0.08	0.09	0.08	0.07	0.08	0.07	•	•	D N
14 SALVI18	1739	0.10	0.09	0.10	0.10	0.10	0.11	0.10	0.11	0.11	0.10	0.11	0.09	0.10		
15 OSMUN18	1675	0.10	0.09	0.10	0.09	0.09	0.10	0.10	0.11	0.10	0.10	0.10	0.08	0.09	0.05	Αų
16 MNIUM18	1706	0.09	0.08	0.10	0.09	0.09	0.10	0.09	0.11	0.09	0.09	0.09	0.08	0.09	0.04	0.04
17 CHLAM18	1642	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.14	0.13	0.13	0.13	0.13	0.13	0.11	0.11 0.12
19 SPERM18	1640	0.13	0.12	0.13	0.12	0.12	0.14	0.13	0.14	0.13	0.13	0.13	0.13	0.13	0.10	0.0
20 TETRA18	1634	0.13	0.12	0.13	0.13	0.13	0.13	0.13	0.14	0.13	0.13	0.12	0.12	0.13	0.10	0.11
21 CHLOR18	1631	0.12	0.11	0.12	0.11	0.12	0.12	0.11	0.12	0.11	0.11	0.11	0.11	0.12	0.09	0.09
22 CLADOI8 23 SCLER18	1028	0.27	0.26	0.27	0.27	0.26	0.27	0.27	0.30	0.26	0.26	0.26	0.26	0.27	0.25	0.27
24 OMPHA18	1473	0.18	0.18	0.19	0.18	0.19	0.19	0.15	0.21	0.15	0.18	0.10	0.10	0.17	0.15	0.15
25 PLEUR18	1485	0.18	0.18	0.19	0.17	0.18	0.18	0.18	0.20	0.18	0.18	0.19	0.18	0.20	0.16	0.17
26 GIGAS18	1554	0.16	0.15	0.16	0.16	0.16	0.15	0.16	0.18	0.17	0.16	0.15	0.16	0.17	0.14	0.15
27 GLOMUI8 28 SARCO18	1527	0.18	0.17	0.18	0.18	0.18	0.17	0.17	0.20	0.18	0.18	0.17	0.18	0.18	0.16	0.17
29 RHODE18	1484	0.34	0.55	0.35	0.19	0.35	0.35	0.35	0.43	0.35	0.35	0.36	0.33	0.36	0.33	0.36
30 PORPH18	1480	0.17	0.17	0.17	0.17	0.17	0.18	0.17	0.20	0.17	0.17	0.17	0.17	0.17	0.15	0.17
31 PHYTO18	1497	0.18	0.18	0.19	0.18	0.18	0.18	0.19	0.21	0.19	0.18	0.19	0.19	0.20	0.17	0.18
32 SCYTO18	1493	0.15	0.16	0.16	0.15	0.16	0.15	0.16	0.18	0.16	0.15	0.15	0.15	0.16	0.14	0.15
55 CR1110	1473	0.25	0.25	0.23	0.22	0.23	0.24	0.23	0.20	0.23	0.23	0.25	0.22	0.24	0.22	0.23

Note.—Taxon acronyms are listed in table 1. Divergences were calculated using substitution rates and among-site rate heterogeneity ( $\alpha$  and  $\rho$ ) corrections (se table 3) estimated using maximum likelihood (ML) over (for 5.8S) the ML tree (fig. 2) or (for 18S) the parsimony tree (fig. 5). The number of sites refers to positions scored as nucleotides. Distances from positions scored as missing or gaps were extrapolated from nucleotide positions. This extrapolation probably underestimates true distance because poorly aligning, hence more extremely divergent, regions were scored as missing.

Table 2	
Extended	

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
·	•		•	•		•	·	•	•	•	•	•	•	•		•	•	•	•	•
							•					•			•	•	•		•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	·	÷	:	•	:	:	•
					•				•				•			•	•	•	•	•
•	•		•	•	·	·	÷	·	•	:	:	·	•	•	:	•	•	:	:	•
			•					•		•		•			•		•	•	•	•
•	•	·	•	·	•	•	•	·	•	·	·	•	:	•	•	•	•	•	•	•
		•	•						•						•			•		÷
•			·	·	•	•	•	•	÷	•	•	•	:	•	•	•	•	:	:	Qow
•		•								•		· ·				•		•		nło
0.11	•	•	÷	•	•	•	·	•	÷	·	•	•	•	•	:	•		:	•	ade
0.11	0.22			•											•		•	•		¢d-fr
0.21	0.24	0.01		•	•	•	•	•	·	·	•	•	•	•	•	•	•	•	•	.om
0.20	0.22	0.01	0.01	0.05			•							•		•	•			htt
0.24	0.24	0.08	0.09	0.10	0.05	0.05	•	•		•	•	•	•	•	•	·	•	÷	÷	ps:/
0.18	0.18	0.06	0.06	0.07	0.05	0.05	0.07	· 		•							•			/aca
0.19	0.21	0.06	0.07	0.08	0.05	0.11	0.08	0.08		•	•	•	•	•	•	•	•	•	•	ade
0.42	0.42	0.23	0.24	0.25	0.24	0.24	0.28	0.27	0.22	0.27		•							•	mic
0.19	0.20	0.17	0.19	0.19	0.16	0.22	0.16	0.18	0.14	0.25	0.05		•	·	•	•	•	•	•	.ou
0.22	0.21	0.21	0.25	0.26	0.22	0.28	0.24	0.23	0.21	0.28	0.09	0.09	0.03				•			p.c
0.17	0.17	0.17	0.20	0.20	0.17	0.22	0.19	0.18	0.16	0.27	0.09	0.08	0.05	0.05		•	•	•	•	om/
0.25	0.26	0.21	0.23	0.21	0.21	0.27	0.23	0.21	0.19	0.29	0.17	0.15	0.15	0.14	0.19	0.42		•		dm,
0.32	0.35	0.36	0.39	0.37	0.33	0.40	0.34	0.37	0.31	0.46	0.36	0.31	0.40	0.38	0.38	0.39	0.42	0.41	•	e/ai
0.41	0.45 0.36	0.30	0.30	0.32	0.29	0.37	0.33	0.29	0.27	0.32	0.28	0.25	0.32	0.26	0.31	0.31	0.48	0.41	0.28	rtidl
0.42	0.48	0.45	0.45	0.47	0.37	0.42	0.39	0.34	0.41	0.57	0.38	0.37	0.40	0.40	0.40	0.57	0.68	0.57	0.42	0 <u>§</u> 6
																				/9/1
•	•							•	•	•	•	•		•	•	•				276
•	•	•	•	•	•	:	•	÷	•	•	÷	•	•	:	:	•				96/
	•						•						•							146
·	•	•	•	•	•	·	•	•	•	•	•	÷	÷	•	•	•				d 60
			•		•	•				•	•			•						b Ai
•	•	•	·	•	•	•	•	•	•	·	•	•	•	•	:	•				ues
•	:	•					•	•												st or
•	•	•	•	•		•	•	•	•	•	·	•	·	•	•	•				ר 21
	÷						•							•	•	•				A
•	•		•	•	•	•	•		•	•	•	•	•	•	•	·				nBr
0.10	· 	•	:				:													st 2
0.10	0.01		•	•	•	•	•	•	•	•	·	•	•		•	•				022
0.09	0.05	0.05	0.07	· 	•	•	• .		•	•		•	:	:	•					10
0.08	0.05	0.06	0.05	0.04		•	·			•	•	•	•	٠	•	•				
0.24	0.23	0.23	0.22	0.23	0.20	0.29	· 	•		•	•	•		:	:					
0.17	0.19	0.19	0.20	0.19	0.18	0.33	0.10		•	·	•	•		•	•	•				
0.16	0.18	0.18	0.19	0.19 0.15	0.17	0.31	0.09	0.03	0.11	· 	•	•	:	•	•	•				
0.14	0.20	0.19	0.19	0.19	0.17	0.32	0.11	0.14	0.14	0.05			•	•	•	•				
0.33	0.38	0.38	0.40	0.39	0.38	0.52	0.37	0.35	0.35	0.33	0.35	0.32	·	:	:					
0.18	0.22	0.18	0.23	0.18	0.15	0.23	0.20	0.22	0.22	0.20	0.20	0.33	0.17		•	•				
0.16	0.20	0.20	0.19	0.17	0.17	0.32	0.20	0.21	0.21	0.18	0.21	0.38 0.38	0.25 0.22	0.21 0.19	0.12	· 				
0.21	0.13	0.25	0.23	0.23	0.22	0.34	0.24	0.25	0.26	0.23	0.26	0.46	0.27	0.27	0.23	0.24				

Table 3						
Properties	of	5.8S	and	<b>18S</b>	rDNA	Sequences

			Taxa S	Subset	S				Rei	ATIVE S	Substit	UTION F	RATES		·	
	1	2	3	4	5	6	VAR	%Var	AC	AG	AT	CG	СТ	- α/ρª	$\alpha^{b}$	ρ
5.8S Sequences	x	х	х	х	х	x	113	0.66	2.5	6.2	2.6	1.9	11.4	0.62/0.00	0.62	0.30
	х	х	х	x	х		90	0.53	2.4	5.0	2.3	1.7	9.7	1.05/0.26	0.49	0.43
	х	x	х	х			81	0.48	1.8	4.3	3.0	1.8	9.7	0.52/0.00	0.52	0.43
	х	х	х				76	0.45	2.1	4.4	2.7	1.9	10.6	0.50/0.00	0.50	0.44
	х	х					64	0.38	2.1	3.1	1.8	1.2	9.9	0.82/0.23	0.45	0.49
	х						53	0.31	2.5	4.3	3.0	1.9	13.3	0.96/0.33	0.40	0.54
18S Sequences	x	x	x	x	x	х	864	0.48	0.8	2.6	1.0	1.1	4.1	0.48/0.22	0.32	0.48
	х	х	х	х	х		617	0.34	0.8	2.6	1.0	1.0	5.7	0.55/0.33	0.29	0.57
	х	х	х	x			570	0.31	0.7	2.5	1.0	1.0	5.6	0.61/0.40	0.27	0.61
	х	х	х				470	0.26	0.7	2.5	1.0	0.9	5.7	0.68/0.49	0.22	0.67
	х	х					370	0.20	0.6	2.3	1.2	1.0	6.3	0.66/0.56	0.17	0.73
	х						304	0.17	0.5	2.4	1.6	0.9	7.3	0.68/0.62	0.11	0.77

NOTE.—Based on sequences listed in table 1. Taxon subsets are: 1—seed plants; 2—seedless embryophytes; 3—Chlorophyceae (including Trebouxiophyceae and Micromonadophyceae); 4—*Cladophora*; 5—fungi; 6—protists. Reversible substitution rates (relative to a G-T rate = 1) and among-site rate heterogeneity parameters (a, the gamma distribution shape parameter, and r, the proportion of constant sites) were estimated using maximum likelihood (ML) over the figure 2 ML tree or portions thereof, with the excluded taxa pruned. The observed number of variable sites (Var) includes only sites with nucleotide polymorphisms. The observed proportions of variable sites (%Var) are rough estimates because they were calculated over the total alignment length (170 sites for 5.8S and 1,813 figure 18S) and ignore variability scored as missing data in poorly alignable regions.

<sup>a</sup> Estimated simultaneously with substitution rates.

<sup>b</sup> Estimated with substitution rates shown and  $\rho$  set to 0.

<sup>c</sup> Estimated with substitution rates shown, without  $\gamma$  rates.

structure within the sample (for which reason we do not include variance statistics) and are based on low numbers of sites. Nevertheless, these comparisons show that a high A bias in Chlorophyceae occurs at 22 sites (lines 5 and 6) that are invariant among the angiosperm sample and, therefore, not likely subject to nonstationarity effects within angiosperms. Sites observed to vary among the angiosperm sample may indeed have a higher A content (line 4), but these sites are conservedly low A in chlorophycean algae. Base compositions in the *Mimulus* sequences fit the algae pattern. These comparisons illustrate that base compositional biases in specific angiosperms or chlorophycean algae are not likely to obscure phylogenetic affinities.

We explored the justification for differential weighting of the 5.8S position for apparent compensatory mutations (Wheeler and Honeycutt 1988) but concluded that the impact of reweighting would be negligible. The experimentally determined pre-rRNA 5.8S secondary structure for yeast (Yeh and Lee 1991; #f. Thompson and Herrin 1994) includes 54 positions paired in cis (fig. 1). These positions are highly conserved among eukaryotes, and the instances of variation appear mostly uncorrelated. Position 114 (C/T) in chaorophytes and fungi, however, usually complements position 139 (G/A). Positions 116 and 137 appear corelated in fungi, but no compensatory correlation is egident for chlorophytes. Otherwise, pairing in the proposed yeast model is in *trans*: the 5' and 3' ends pair with, respectively, the ITS1-3' and 26S-5' ends, and the terminal two bases of the 5.8S are paired with ITS2: 28–29. This pattern is evident in models proposed for diverse eukaryotes (Coleman, Suarez, and Goff 1994; Fritz et al. 1994; Holmberg, Melander, and Nygard 1994; Suh, Thien, and Zimmer 1992). 9

https://a

#### Phylogenetic Trees

The consensus of 5.8S topologies derived heughtically using ML (fig. 2;  $-\ln$  likelihood = 1,803.43%7)

Table	4
-------	---

Dase Composition of Anglosperms, Chorophyceae, and Reported Munulus 5.65 rDNA seq
---

_										-				0
			ANGIOSPERMS				CHLOROPHYCEAE				"Mimulus"			20
	CLASS OF SITES	n	А	С	G	Т	Α	С	G	Т	A	С	G	$\eta^{N}$
1.	Variable in either angio <sup>a</sup> or chloro <sup>b</sup>	56	0.14	0.36	0.27	0.23	0.26	0.27	0.18	0.29	0.23	0.29	0.19	0.30
2.	Variable in angio	32	0.18	0.40	0.10	0.31	0.11	0.32	0.21	0.36	0.12	0.30	0.22	0.37
3.	Variable in chloro	20	0.08	0.46	0.24	0.23	0.20	0.33	0.13	0.34	0.13	0.38	0.18	0.33
4.	Variable in angio, constant in chloro	24	0.23	0.36	0.10	0.31	0.14	0.27	0.27	0.32	0.16	0.27	0.25	0.32
5.	Constant in angio, variable in chloro	12	0.08	0.42	0.33	0.17	0.32	0.25	0.18	0.25	0.20	0.38	0.21	0.21
6.	Constant in angio and chloro, but vary-													
	ing between them	10	0.00	0.10	0.80	0.10	0.70	0.20	0.10	0.00	0.65	0.20	0.10	0.05
7.	Variable in both angio and chloro	8	0.06	0.53	0.10	0.32	0.02	0.45	0.06	0.46	0.00	0.50	0.13	0.38

NOTE.—The data represent only the sequences included in table 1. The sites examined are derived from the figure 1 alignment, as classified here. Thus, the values for Chlorophyceae for Class 2 (variable in angiosperms) represent the base frequencies of sites in Chlorophyceae that align with variable sites in angiosperms. n = Number of sites in each class.

<sup>a</sup> Angiosperms.

<sup>b</sup> Chlorophyceae, including Trebouxiophyceae and Micromonadophyceae.



FIG. 2.—Maximum-likelihood (ML) tree for the 5.8S data set (appendix 1) modified from the 5.8S alignment (fig. 1). Taxon acronyms are listed in table 1. The tree ( $-\ln$  likelihood = 1,803.4408) was derived using the heuristic ML procedure (stepwise taxon addition, tree bisection-reconnection branch swapping) in PAUP\* with estimated substitution and rate-heterogeneity parameters on line 1 of table 3. Missing data (including gap scores) are ignored in calculating branch lengths. Branch lengths are equal to the expected number of substitutions per site (scale bar).

conflict with morphological and/or other molecular evidence in the arrangement of angiosperms (Chase et al. 1993), monophyly of seedless embryophytes (MARSI, OSMUN, MNIUM; Garbary, Renzaglia, and Duckett 1992; Kranz et al. 1995), inclusion of red algae (CY-ANI, SARCO) among chlorophytes (Delwiche, Kuhsel, and Palmer 1995), paraphyly of fungi (Bruns et al. 1992), and relationships among fungi (Gargas et al. 1995). The inclusion of red algae among chlorophytes is perhaps the most striking anomaly, but character state reconstruction using parsimony under ACCTRAN option in PAUP\* indicates that of the 11 apomorphies potentially supporting such a relationship, 10 subsequently reverse. Figure 3 shows one of 1,170 377-step trees derived using MP. Like the ML tree, the MP trees conflict with broader evidence in the arrangement of angiosperms, monophyly of seedless embryophytes, and relationships among fungi. In addition, the chlorophycean alga *Chlorella* (CHLOR) erroneously grouped with the ulvophycean *Cladophora* (CLADO; Mishler et al. 1994; Friedl 1995, 1994). Chlorophytes are monophyletic in the MP trees, but the bootstrap support is low. The best ML score among MP trees is 1,806.4889, while the parsimony score for the ML tree (fig. 2) is 383.

Using the rate parameters from the ML tree above, the ML score for a 5.8S tree constrained for independently supported groupings shown in fig. 4 is 1,820.0009. With parameters reestimated over this constrained topology, the score improves only slightly to 1,819.7321, while the unconstrained topology becomes 1,803.7472. The best parsimony tree consistent with the figure 4 constraints has length 387.

The 18S parsimony tree (fig. 5) also contains several groupings contrary to broader evidence: the angiosperm arrangement, monophyly of gymnosperms (EPHED, PINUS, ZAMIA; Doyle, Donoghuc, and Zimmer 1994), monophyly of seedless embryophytes, and polyphyly of the stramenopiles (CRYPT, PHYTO, SCY-TO; Leipe et al. 1994). The 18S tree has more bootstrap recoveries  $\geq$  70% (at 16/32 nodes) than the 5.8S parsimony tree (6/36 nodes). One of these high recoveries (PINUS plus EPHED; 77%) supports a likely incorrect node. We did not explore secondary structural constraints on 18S sequence evolution across this data set. although such have been proposed in other 18S analyses (Wheeler and Honeycutt 1988). The parsimony score for a tree constrained for independently supported relationships (analogous to fig. 4) was 2,380.

#### ITS2 Sequences Alignment

Figure 6 shows the CLUSTALV alignment of green algal, fungal, and angiosperm ITS2 sequences. The alignment was not adjusted a posteriori; it demonstrates mechanical detection of conserved motifs in highly divergent sequences rather than optimal multiple or pairwise alignment. As with 5.8S, the ITS2/26S boundary is not always clear, but even fewer RNA-based sequences exist for the latter. We presume that the 3' end of the alignment approximates the boundary. There are no unambiguous conserved motifs shared between the algae, fungi, and plants, although a G-rich region is evident at positions 190-210, usually containing one or two YGG trinucleotides. Its conservation appears bona fide among angiosperms (Liu and Schardl 1994; Hershkovitz and Zimmer 1996), as does the alignment to conifers, but the homology is less clear in the other groups, especially the fungi. Within major groups, motif conservation is amply evident, especially among Caryophyllales, for which many of the gaps are nearly perfectly aligned, e.g., between positions 70 and 250. The most alignable regions among Caryophyllales superimpose over motifs shared with other angiosperms, e.g., position 75–97 and positions 131-141. Both regions are shared by conifers, as well, and the latter corresponds to the angiospermconserved motif previously identified by Liu and Schardl (1994). The conifers share broad regions of alignability, but the taxa are too closely related to infer broadly conserved conifer-specific motifs.

Because one issue of the present paper is the identity of the reported *Mimulus* sequences, the chlorophycean sequences are aligned more optimally in figure 7. Most striking is the similarity between the reported *Mimulus* sequences (MIURF, MIURH, and MIURM) and the chlorococcalean (SPONG) and volvocalean (GONIU and PANDO) taxa. Worth reiterating, however, is the apparent lack of primary sequence conservation between these green algae and the ulvophycean CLA- DO, as evidenced by comparison (not shown) and failure of this sequence to retrieve algal or any other ITS sequences from the databases.

The alignment of the fungi appears less cohesive than for the other major groups, although the 5' end appears conserved. Greater similarity is evident within the fungal classes, i.e., the basidiomycete (VOLVA and HETER) and ascomycete (BIPOL and SCLER) pairs. The zygomycete (GLOMU) most resembles the ascomycetes. Conservation within the fungal groups would likely be better characterized by incorporating more of the available sequences, but consideration of fungi here is primarily for perspective in relation to green algae and plants.

#### Guide Tree Analysis

Figure 8 shows the CLUSTALW guide tree generated using gap-opening/-extension penalties of 2520.3. Caryophyllales and green algae, including the putative Mimulus sequences, each appear distinct. The Mimulus sequences (MIURF and MIURH) cluster with the calorococcalean Spongiochloris (SPONG) and, in turn, with two volvocaleans, Pandorina (PANDO) and Goratum (GONIU). The close relationship between the MIURF and Spermatozopsis (SPERM) in the 5.8S sequences is not evident in ITS2, although sampling among chlorophycean algae is sparse for both sequences, and no Spongiochloris 5.8S sequence is available. Separation of other taxonomic groups is less pronounced, although the tree groups the basidiomycete, ascomycete, and confifer pairs. Mild perturbation of the gap penalties causes all of the "resolution" to disappear. Still, we found no guide tree with comparable internal branch resolution that showed intermixing of algal, fungal, and seed plant sequences. /9/

#### Discussion

We undertook the present analysis because phylogenetic signal in the ITS region had not been characterized sufficiently to diagnose the organismal origin  $\Im$  a suspected contaminant sequence. Properties of the JTS region sequences posed several analytical challenges. For the 5.8S sequences, these included (aside from the more broadly applicable problems of site- and lineagespecific evolutionary rate variation, nonstationary substitution patterns, and sampling limitations) the potentially inadequate number of variable sites and the possibility that the variable sites were either too functionally constrained and/or mutationally saturated. In the case of ITS2 sequences, diagnostic signal across the phylogenetic space considered here had never been demonstrated (cf. Baldwin et al. 1995). Besides resolving our original query, the exercise yielded a broader assessment of the evolutionary properties of 5.8S and ITS2 sequences and useful observations on sequence analysis in general.

#### 5.8S Sequences

Ribosomal DNA 5.8S sequences yield phylogenetic trees qualitatively similar to those from 18S but less well supported. The 5.8S and 18S trees conflict similarly with



independent evidence in the spurious grouping of ferns plus moss. The similarities in the trees probably reflect, in part, the genetic and functional linkage of these sequences. This linkage is perhaps best demonstrated by the high divergence of *Cladophora* in both the 5.8S and 18S data sets, evident in the figure 2, 3, and 5 trees. But linkage per se has no bearing on amount and/or quality of phylogenetic signal, as evidenced by the likewise linked internal spacers. The lower bootstrap recoveries for groups in the 5.8S tree is probably attributable mainly to the relatively low number of variable sites at a given phylogenetic level.

Sequence divergences are much greater in the 5.8S than 18S sequences, but, at any level of phylogenetic

comparison, the proportion of variable sites is also much higher, and likelihood estimates of among-site rate heterogeneity are less extreme (table 3). This difference in the variation pattern offsets, to some degree, the >10fold advantage of 18S in the number of analyzable sites. Divergences in 18S measured from angiosperms increase gradually with the inclusion of more distantly related chlorophytes (excluding the highly divergent *Cladophora*) and more with the inclusion of fungi and protists (table 2). Some angiosperm–fungi and angiosperm– protist divergences appear similar, but this discounts divergence in poorly alignable regions scored as missing. The 5.8S divergences from angiosperms to chlorophycean algae and to fungi are approximately the



FIG. 4.—Consensus tree of relationships supported by diverse lines of evidence. Support for the numbered nodes is discussed and/or referenced in (1) eudicots, (2) monocots, and (3) angiosperms (Chase et al. 1993; Doyle, Donoghue, and Zimmer 1994) 4—8; (4) gnetophytes and (5) anthophytes: Doyle, Donoghue, and Zimmer 1994); (6) Pinaceae and (7) conifers: Page 1990; (8) seed plants: Doyle, Donoghue, and Zimmer 1994); (9) ferns, (10) vascular plants, and (11) embryophytes: Mishler et al. 1994; (12) Volvocales, (13) Chlorophyceae/ Trebouxiophyceae/Micromonadophyceae and 14) Ulvophyceae plus node 13 (Mishler et al. 1994); (15) chlorophytes (Bhattacharya and Medlin 1995); (16) basidiomycetes, (17) filamentous ascomycetes, (18) ascomycetes, (19) septate fungi; and (20) fungi (Gargas et al. 1995); (21) heterokont stramenopiles and (22) alveolates plus node 21 (Leipe et al. 1994; Bhattacharya and Medlin 1995) and (23) rhodophytes (Bhattacharya and Medlin 1995).

same, suggesting mutational saturation at this divergence level, but the calculation ignores the length-variable region (fig. 1, positions 121–150, in which angiosperms are clearly more similar to green algae than to fungi. Moreover, the angiosperm–protist divergences are much higher than all other divergences from angiosperms.

Our estimates of  $\alpha$  and  $\rho$  over taxon subsets of both the 18S and 5.8S data appear to reflect either variation in among-site rate heterogeneity throughout the phylogenv or the sensitivity of the estimate to divergence patterns among the sequences (i.e., sampling; cf. Yang 1994), or both. For the 18S data, the estimated variation possibly reflects, at least in part, our scoring as missing data the poorly alignable, hence hypervariable, regions in nonangiosperms. The  $\rho$  estimate, however, should be lowest rather than highest for the seed plant data subset, which includes a higher proportion of potentially variable sites. The  $\alpha$  and  $\rho$  estimates for the tree with embryophytes pruned (see Results) fall in between those estimated over the entire taxon set and the set excluding protists. That the  $\rho$  estimate in this case (0.53) is still much less than those for the embryophyte and seed plant sets is counterintuitive, because the nonembryophyte sequences include all sites that align with embryophytes (i.e., the most conserved regions) minus those that a not. The 5.8S rate heterogeneity estimates vary more erratically over the taxon subsets than the 18S data, which probably reflects the small number of total sites, hence high variance. Considering variance, there is presumably a range of combinations of substitution and among-site heterogeneity parameters that could explain the data more or less equally well, but recovery and evaluation of all combinations is not readily feasible  $\vec{s}$ ing current software. In any case, variation in estimates of rate heterogeneity are problematic, because the ML algorithm in PAUP\* (and essentially all commonly applied phylogenetic algorithms) assumes that all rate marameters are constant.

Our results provide a practical illustration of proplems that obfuscate phylogenetic reconstruction from DNA sequences. While the phylogenetic utility of 5. rDNA is obviously constrained by its short length, our comparison with 18S rDNA indicates that sequence length is not the sole determinant of tree accuracy, as demonstrated by relatively strong bootstrap support for the presumably inaccurate grouping of *Ephedra* with *E*nus in our 18S tree. Large divergences, character nonindependence, and lineage-specific evolutionary rates, coupled with analytical methods that fail to correct for any or all of these conditions, can create an inverse correlation between data quantity and support for the true tree (Hillis, Huelsenbeck, and Swofford 1994; Lockhart et al. 1994). <u>s</u>

#### **ITS2** Sequences

The major factors limiting broader elucidation of ITS sequence conservation patterns within and among eukaryote groups are scattered and sparse sampling, limitations on secondary structural comparison imposed by sequence hypervariability, and an increased probability that hypervariable molecules in distantly related taxa will randomly include short stretches of sequence similarity. The present work demonstrates how a more taxonomically intensive sampling will bear upon understanding of eukaryotic ITS evolution. Previously, recognition of short conserved motifs in otherwise unalignable ITS sequences have been limited to examples

20



Downloaded from https://academic.oup.com/mbe/article/13/9/1276/991469 by guest on 21 August 2022

FIG. 5.—Maximum-parsimony (MP) tree for the 18S data set (available upon request). Taxon acronyms are listed in table 1. Numbers above branches are percent bootstrap recovery. Asterisks denote branches collapsing in the strict consensus. Branch lengths are proportional to the number of changes assigned to that branch (see scale bar). The tree (length = 2,315, CI excluding uninformative characters = 0.45; RI = 0.61; RC = 0.33, 535 parsimony-informative sites) was derived using the heuristic MP procedure (100 trees held at each step of stepwise taxon addition, tree bisection-reconnection branch swapping) in PAUP\*.

in vertebrates (Crouch and Bachellerie 1986; cf. Schmickel et al. 1990) and angiosperms (Liu and Schardl 1994). Conserved motifs have been reported for fungi but these have been limited to ascomycetes, especially yeasts (Crouch and Bachellerie 1986; van Nues et al. 1995). Wesson, Porter, and Collins (1992) reported among insects, frog, and sea urchin, "a region [of ITS2] showing good homology" with sequence flanking ribosomal processing sites identified in rat and yeast. The specified region, however, is unconserved even among yeasts and, besides, is substantially dispensable for yeast rRNA processing (cf. van der Sande et al. 1992; van Nues et al. 1995). Conserved motifs in ITS2 (fig. 6) are responsible for the partitions between seed plants, green algae, and fungi in the guide tree (fig. 8). The motifs shared between conifers and angiosperms evidence conservation over at least 350 million years since these shared a common ancestor (Stewart and Rothwell 1993). Likewise, the distinctiveness of Caryophyllales ITS2 sequences among angiosperms evident in the alignment and figure 8 exemplifies suprafamilial phylogenetic signal in the spacer (see also Hershkovitz and Zimmer 1996). The degree of ITS2 sequence conservation evident among green algae (Fig. 7) appears greater than or equal to that among seed plants (cf. fig. 6). The divergence time

4.	tetrg	atogogtctt c-c-c-acccacatcgc-gtggg-ag	g-ggaaggag	gatggootcoogtg	-cctcac	c-gggcgcg-gctggcctaa	atagggagccc	a oggogaogagt	-agccgcggcgatt
3.	mollu	atagegtctc c-c-c-gcacctaccgc-gcgatg	g-ggaaggag	gatggoctcoogtg	-cctcac	c-gggcgcg-gttcgcctaa	agaoggagccc	a cogcatogage	-ogccgoggogati
4.	phytl	atogogtctn c-cc-toccataoctagt-g-gag-gg	gag - ggg aaggag	gatgcoctcoogtg	-ccttac	c-gggtgcg-gctggcctaa	aaatggagtoo	cogg-aatgagt	-tgccgcggcgatt
5.	amara	actgcgtctc c-c-c-caa-cocacctagctgtgg	gaggggcgaggag	gatggtctcccatg	-cctcac	c-gggcgtg-gatggcctaa	aacaggagccc	a cggtttcgagc	-tgctgcggcgati
6.	silen	atogogt-ctcc-c-c-aaatccaccacgag-tg	gagggg-aaggad	gatggtttcccgcg	-actcac	c-ggtogog-gotggootaa	aattggagooo	t oggcatoggga	-tgeogogogatz
7.	tritu	caaaaca-ogctc ccaacccct catcogg	gaatcaggatgcggca	tctgqtccctcqt	-ctoqca	aqqqqcq-qtqqaccqaa	gategggetgeegeetee	g coccopicaopicg-ctgtopc	gtccgcgcat
8.	maian	toacytoyctc c-gcyc-acoc-tyccct cccyaaccg-gag-ag	-g cygcyggcgcygatgcyga	attgyoccccgtg	-ct	geggaegeg-gegggtegaa	gtgctggcc-	- googgecggg	acygacycygcyagt
9.	canel	cctcacattoctc c-c-c-aocaccotogg-tg	atoggggatocoga	attquecqtecate	-ctat	ttgtgcgtg-gtcggcctaa	acttgggood	c ttatagctagcg-cggc	atgacacqqcgaqt
10.	arabi	a tog togte - c ctcaccat cct ttgc	tgatgc-gggacggaa	getggtetecogtg	tgttac	cgcacgogttggcctaa	atoc-gagoca	aggaogootggage	gtacogacator
11.	lycop	atogogtogccc c-togc-acqc-ogca-aq	g ctttag-ogoggggggggaa	getggeeteeegtg	-cqcc	gagegegeg-geeggeetaa	atgogagtoc	a ogtogacgg	acgtcgcggcaag
12.	vicif	atogaageete ettgecaattte cotgattattgtg	caqq-gtgqat	gttggoctcoogtgage	tctttogtctc	atggttggttgaa	aattgagace	t tggtagggt	gtgccatgatagat
13.	cucur	atcoctoccc c-caccac-acaactct ccccatocopg-gt	xy ttytgaaggcagggacacad	actggootcoogta	-ccac	cgtcgtgcg-gatggcttaa	attt-gagtoc	t ogatgetegtegtegeg	acactacggtggtt
14.	lactu	atoge-togetc cocaccat-acet cccctacggttog	a tggtgttggggggggat	aatggcctcccgtg	-cttgtgttt	cqqttqqcctaa	ata	agagtteettegge	ggacgcacgactagt
15.	pinus	atoc-cattctocaacococt ccctocaa	ta toctagogagcagcoga	actautoatocata	occaacoc	ggtgcg-gttggctgaa	atacctc	a agcgatgtttcgtggc	gtgcgt-cggcgag
16.	picea	catectae-atacacgognt cectgnaagg	a tqqqqqaqqaqcqqa	gatggtogtoogtt	occaaogt	gqntcg-gtcggntgaa	atac	a tgccatgttccgtgtc	ttgtgt-cggcgag
17	obren	aaaac-tcacactaaaaaac ctoctttititg-ototogaa	ctract	atctcatat-acacac	acaaacatoogca	ta	gottogttgaagttt	a canctoctattacata-gac	conctrat-ougectetac
18.	goniu	anto contractor contractor contractor anto contractor c	a a macquatctooct	atctoa	acaatcatataa	tta00g	aatcagctgaagt	a cagagetteatecate	conctaat-opportctact
19.	miurf	aatatoactoraatoaaccttotottoa	at acaatagatetaget	ottecaatac-otcatoaoc	ttactoatag	tttaacdtgatg	gatcagctgaagt	g cagaggetaatgeaag-gae	cc-tt-at-gggccgcaact
20.	miurh	aatatoaccaatoattatactata-totataa	c act g-gtggatctggct	attecaetctecec-ae	ctagetgeetagtg-	-tocatctg	gatetgctgaagt	a cagaggetaac-caag-gac	corte-aa-gggccgcaact
21.	SDODG	aatatcacactactactac	t -ccatatagacctago	attocanttoantac	togaccacctttactoutga	togaatcagtcacoctg	gatcogctgaaga	a cagaggetaac-teta-gac	couct-aa-gogccocaatt
22.	sperm	caccectcoctccaccac cecttcaccouncutatteg coccectoctcocgaga	c accagaacagatataac	ttoccag	gaccoct	tcaaaccgggggctg	ggttggttgaagac-	- cagaggetagagea-g-gae	cogtcaagggcttcaact
23.	chlor	cacce-ceteoceaee eteteeteteetete	a gaag-googatetogoo	atetetatae	tcctogaga	gaa Coq	gotcogctgaaga	a cagaggettgagcatg-gtc	ccoctgatagggcgccagct
24	volva	antartantana ana ana ana ana ana ana ana ana	and tagattag			torm-atometetortea	aag	. catcagcag-ggc	coadcocadt-coocct
25.	heter	gaa tete cange of ugeogge teres ogget teteg	r acatamettamettam	agett	tactagtee	togogatogoctectetea	aat	g cattagcga-gac	ccttatactaccacca
26	hipol	galactor can the second and the second	taxe	atttttttato	tootettte	toppagactcocctta	aaacoattg	CancC	cta
27.	scler	-caacoctcaac toac	taad		ccatotcaotaa	toocagoctcta	aaat	- cagtogcggc	ac-cqct
28.	aiaas	actacoctcaaac ocaac	taga	tcatexte		-000000ac000000	aaag	a canconcouc	acatato
						····	J	5	5
		201							385
1.	talin	-ggtggtggacgagget tgaggecc ctgtttgcatc	ogcoqc-qca	-cqcacqc-cqtcqqq	aqqqqctcqtcqqaccct	t-aqqtqqcqaaa		gc-cac-aaacogtt	
2.	tetra	-gotogattaaaacctca cggatogogc t	qcq-	caaqcacqt-cqcc-aq	qqqaqctcq~-aaqqacccc	q-aq		gctcac-aa-cogtt	
3.	mollu	-ggtggattgaacggcoccgagccacgt caatttgta	cq tcqc-qtq-	-cqaqcqaqq-tqcccqa	qqqaqctcqctqqacccc	g-agcgcga		<ul> <li>gctcat-accctt</li> </ul>	
4.	phytl	33-33			topagetontaggacoct	tgatgage	ctog	tortcaaaacottt	
5.	F-4	- ootooatoaoaoooctc	1catcat-ata-	· -tocacat-toccaot					
6.	amara	-ggtggatgagaaggeete	gogtogt-gtg- gogtogc-ac	<ul> <li>-tgcacat-tgccagt</li> <li>aqtqcqt-tqacctt</li> </ul>	gtggccttga-ggacoct	a-gagegttge	ccga	ggggggaccaa-cogtt	
	amara silen	-ggtggatgagaaggoctcggtcgtggttcmnnt -ggtggtgtgcaaggoctcagoctagaatgcaatc -ggtgutgaacaaggoctc	gogtogt-gtg- gogtogc-ac oqtogc-gogt	agtgogt-tgacett coccagtgccaac	gtggccttga-ggaccct	a-gagegttge	ccga	gggcgaccaa-ccgtt	
7.	amara silen tritu	-ggtgggtggagaggoctcggcogtggttcmntc -ggtggtgtgcaaggoct	gogtogt-gtg- gogtogc-ac ogtogc-goqt toactoroaconcol	agtgegt-tgacett cocagtgecaac	gtggccttga-ggacoct gtgagctogaaggacoct	a-gagegttge t-tgatgttge	ccga cttq	gggggaccaa-cogtt ttgcaaccaaaccgtt c otogtcttogacc	
7.	amara silen tritu majan	gytgytgyacgagactc ggcggg ggtgytgyacgagactc accord ggcggtg ggtgytgyacgagactc agatga at	g - og		gtggcettga-ggacoet gtggetogaaggacoet aatgacoet	a-gagcgttgc t-tgatgttgc gcaaacg-cccaac	cttq agcaca		
7.	amara silen tritu maian canel	gytgytgyacgaggoctc gycygr gytgytgyacgaggoctc agoct agoct agaatgoarat	g -ogtogt-gtg g -ogtogc-ac togc-gogt togc-gogt 		gtggccttga-ggacoct gtgactogaaggacoct oggcgcatgcaaggaacoc- gtggtgaggaacoc-	a-gagcgttgc t-tgatgttgc gcaaacgccgagc g-aaacgccgagc	ccga cttgagcaca		
7. 8. 9.	amara silen tritu maian canel arabi	ggtggtgtgacgaggoctc ggcogtg ggtggtggcaggoctc ageagac ac	g - cgtcgt-gtg g - cgtcgc-gcq cgtcgc-gcq tgactccgaccgcgq a - cggcgctgaaccgctgcct a - cgcgt-gcg g - tagat		gtggccttga-ggacct gtgactcgaaggacct aatgacca gtggtggggggctt 	a-gagcgttgc t-tgatgttgc gcaaacg- acgccgagc g-aaac	ctgagcaca	- gyggaccaa-cogtt - ttgcaaccaaacogtt c gtogtcttogacc acccctoggaaca	
7. 8. 9. 10. 11.	amara silen tritu maian canel arabi lycop	ggtggtgtgacaggoctc	g - cgtcgt-gtg g - cgtcgc-ac- cgactcogacogcg a - cggcgctgaacogctgcct a - cggcgctgaacogctgcct g - tagat	ggacat-tgccagt agtggt-tgactt cccagtgccaac cggcatgg-tggcccaa cogcacc-oggccccag gaatgtgc-tgcgg 	gtggccttga-ggacoct gtggcctgaaggacoct oggcgcatgcaaggaacce- gtggtggggggcctt aaag-toc	a-gagcgttgc t-tgatgttgc gcaaacg	ccga 	-gyogaccaa-cogtt c gtogtcttogacc c gtogtcttogacc accoctoggacca ccaaggatgc	  
7. 8. 9. 10. 11. 12.	amara silen tritu maian canel arabi lycop vicif	ggtggttgacaggoctc		gaacar-ugcoagt agtogt-tgcoagt coccagtgccaac coccagtgcco-ogg gaatgtgc-tgcc-gg gaatgtgc-tgcc-gg 	gtggcttga-ggacct gtggctcgaaggacct cggcgcatgcaaggaccc gtggtggggggctt aagtcc ccgcgct	a-gagcgttgc t-tgatgttgc gcaaac			   t-ga-
7. 8. 9. 10. 11. 12. 13.	amara silen tritu maian canel arabi lycop vicif cucur	ggtggtggacgagagggctc		gacat-ugcoagt 	giggottig - a-ggacot giggottig - a-ggacot 		cttqagcaca agcaca ata cttaaaaggacg		  t-ga-
7. 8. 9. 10. 11. 12. 13. 14.	amara silen tritu maian canel arabi lycop vicif cucur lactu	ggtggtggatggagggctc			giggottig-a-ggaoot giggottig-a-ggaoot giggottig-aaggaoot giggottig-aaggaoot giggottig-aaggaoot cogoottigaattig-or-gooto tigaattig-or-goot totacaagaoot		cttq		  t-ga-
7. 8. 9. 10. 11. 12. 13. 14.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus	ggtggttggatgagaggottc	9 - og tog - tog - - og		giggottig-a-ggacot giggottig-a-ggacot cggogotagcaaggaacoo- giggtgggygoct cogogot	a-gagegttgc t-tygatgttgc gcaaac		gycgacca cogt- - tgyaaccaacogt- - cgogtotoga-co- - acoctoggaca- - getoc- - getoc- - gottt-coga- - acoctog- - acoctoggaca- - getoc- - get	  t-ga- t-tt-
7. 8. 9. 10. 11. 12. 13. 14. 15.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea	ggtggtggatggaggottc			giggoctig-a-ggacoct digactor-aaggacoct aadgacoca giggo-tig-oggggoctic 			gycgaccaa cogtt trocaccaacogtt gycgaccaacogtt gytogatcyga-cc- 	  t-ga-  t-tt- t-tc-
7. 8. 9. 10. 11. 12. 13. 14. 15. 16.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea pando	ggtggttgaacaggoctc - ggcogtg - gttrannt	- ogtogt-gt- - og	t - gacat 'g - coagt agtogt 'g - act' c - ggcatgg- tg-act c - ggcatgg- tg-octcaa coggac - c - oggcoccaa - gactgg- tg - c - gg coggactoc - ggactca agtogg- tg - c tct c agtogg- tg - c tct ggcttgfgsgactoc ggttgfgfgsgattc	giggottig-a-ggacot giggottig-a-ggacot adgactor-aagacot ogogotgagaaggaacoc- gigg-tg-aggagoct cogogot-aagacot tigaattig-gcacot tigaattig-tigottigottit ggaactigottigottit ggaactigottigottit			ggcgaccaa.cogtt- ggcgaccaa.cogtt- cgtogttbgacc- 	 t-ga-  t-tt- t-tc- tctc-
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. <u>16.</u> 17.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea pando coniu	ggtggtgagaggactagaaggoctc ggcggtg ggtggacggaaggoctc ggcggtg aggact aggacgaaggoct ggcggacggacgaggoct ggcggacggacggacgacgacgacgacgacgacgacgac			giggoctig a ggaocot diggoctig a ggaocot asigacos cguygotgoaaggaacoc diggo-tig-ogggygott 			gycgaccaa cogtt -ttpcaaccaa cogtt -ttpcaaccaacogtt -gtocttoggacca -gctoc	  t-ga-  t-tt- t-tc- tetc- tetc-
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. 16. 17. 18.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea pando goniu miurf	ggtggtggatggagggctc     ggtggtggatggaggctc       ggtggtggatggaaggctc     agaatgca-atc       ggtggtggatggaagggctc     agaatgca-atc       ggtggtgaacgggt     ggtggtggatgga       ggtggtggatggatgga     gatggatggatgga       ggtggtgaacgga     gatggatggatgga       ggtggtgaacgga     gatggatggatggatggatggatggatggatggatggat	j - og togt - gt toga togt - gg toga	c - goacat 'g - coag - agtog: fg - acctt c - goacat - g - goacat c - goacat - c - ggoacag - gaatgug - tg - c - gg c or - gaacat - c - agtog - tg - c - ta - c - agtog - tg - c - ta - c - agtog - tg - c - ta - goattag - gaacat - ggattag - gaacat - ggattag - tg - tg - tg - tg - tg - tg	giggottig-a-ggaoot giggottig-a-ggaoot adgaota-aaggaoot giggig-ggyggott tigaattiggotoc ttcacgaoot 			ggycgaccaa cogtt- -ttgcaaccaa.cogtt- -gtogtttogacc- -acocctoggaca -gctoc -gac -gottt-coaacggtgg -acgtt-togacc- -atgct-togacc- 	t-ga-  t-ga-  t-tt- t-to- tetc- tetc- tetc- 
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. 17. 18. 19. 20	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus pinus pando goniu miurf	ggtggtggatggagggctc - ggcggtg - gttrmntaggtggtggatggaaggctc - aggatga - agatga - agatga - at	- ogtog-ac		digactica a-gracoct digactor - aspacoct adgractor - aspacoct ogyogratycaagaacco- digar-ti- oggogoct - aag-toc ogogoct - ticaaca-gacoct togactigatig goctoc - ticac gacoct - tocaca acaggaa	-a cgage		gygogaccaa cogit -tigaaccaa cogit -gigogaccaa cogit -gotoctogaacca -gotoctogaacca -gotoctogaacca -gotoctogaacca -gotoctogaacca -agotoctogaacca -gotottagaacc	t-ga- t-tac- tetc- tetc- tetc- tetc- -etc-
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. <u>16.</u> 17. 18. 19. 20. 21.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea pando goniu miurf miurf	ggtggtggatggaggottc     ggtggtggatggaggottc       ggtggtggatggaaggottc     aggatgatata       ggtggtggatggaaggottc     aggatgatata       ggtggtggatggatggatggatggatgga     ggtggtggatggatggatggatggatggatggatggatg	- og togt-gt- - og togt-ga 		giggottig-a-ggaoot dtgactor-aaggaoot agtgot-ig-aaggaoot gigg-tig-aaggaoot cogogot-aaag-toc cogogot-togagot- togactig-tig-agoot 			gycgaccaa cogtt -ttcaaccaacogtt gycgaccaacogtt gycgacca- -accccoggaca- -gotoct-oggaca- -gotoc-gaco- tgogtt-cogacc adget-togacc 	t-ga- t-ga- t-tt- t-tt- tc-cc- tctc- tctc- tctc- tctc- tctc-
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. <u>16.</u> 17. 18. 19. 20. 21. 22.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea pando goniu miurf miurh spong	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	- ogtog-ac	tgacat-g-ccag agtogt-g-act- c-ggcatgg-tg-cca- cogga- cogga- gactgg-tg-cc- gacc ggcatgg-tg- ggcatgg-tg- ggcatgg-tg- ggcatgg- ggcatgg- ggtatgg- gacca- ggcatgg- ggcatgg- ggcatgg- ggcatgg- ggcatgg- ggcatgg- ggcatgg- ggcatgg- 	digactig a gaacoct digactor aagacoct ogogoctacaagaacoc digactor gogogoct cogocot togactig			gycgaccaa.cogtt- gyggaccaa.cogtt- cgtogttbgacc- 	  t-ga-  t-tt- t-tt- t-tt- tetc- tetc- -etc- -etc- ettc- tetc-
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. <u>16.</u> 17. 18. 19. 20. 21. 22. 23.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus picea pando goniu miurf spong sperm chlor	ggtggtggatggatggaggotte     ggtggtggatggaaggotte       ggtggtggatggaaggotte     aggatggatggaaggotte       ggtggtggatggaaggotte     aggatggatate       ggtggtggatggaaggotte     ggtggtggaaggotte       ggtggtggatggaaggotte     ggtggtggatggaaggatate       ggtggtggatggatggatggatggatggatggatggatg	- ogtayt-stay-ac		giggoctig-a-ggaocct dtgactog-aaggaocct oggogoctgoaaggaacco- giggo-tg-oggggoctt togaottg-oggggoct togaottg-oggggoct togaottg-ttog- ggaactucttog- accaggaa- 			gycgaccaa cogtt -ttccaaccaacogtt cytogttoga-cc- -accctoggaca -gctoc gaoc tgogttt-ccaacogtg acgtt-toga -gctottgg-ttccaac -gtocttgg-ttccaac -gtocttgg-ttccaac -gtocttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtotttgg-ttccaac -gtottgg-ttccaac -gtotttgg-ttccaac -gtottgg-ttccaac -gtottgg-ttccaac -gtottgg-ttccaac -gtottgg-ttccaac -gtottgg-ttccaac -gtottgg-ttc-ac -gtott	  t-ga-  t-tt- t-tc- tctc- tctc- -ctc- -ctc- -ctc- -ttc- 
7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23.	amara silen tritu maian canel arabi lycop vicif cucur lactu picea pando goniu miurh sporm sporm chlor	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	- og tog t-gt- - og tog - ac- - tog - tog - cog - - tog - tog - cog - - og cog - - cog cog - - cog cog - - cog gta - - cog gta - - cog	t - gazat - g - coagt - agtogt - g - actt cocgutor - coggoodag - gazdgg - gazotag - gazdgg - gazotag - gazdgg - gazotag - coggood - gazotag - gazdtag - gazotag - gazttag - gazotag - ggttgg - gazotag - ggttgg - gazotag - ggttgg - gazotag - ggt - t coggooda - tg t coggooda - gg	giggottig-a-ggaoot giggottig-a-ggaoot giggottig-aaggaoot giggig-ggygott tigaattiggotto 			ggcgaccaa cogtt 	t-ga- t-ga- t-ga- t-tt- t-to- teto-
7. 8. 9. 10. 11. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinus pinus pinus pinus goniu miurf miurf spong sperm chlor volva	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	- og		giggoctig-a-ggaocot digactig-a-ggaocot adigactor-aaggaocot adigactor-aaggaocot giggo-tig-oggagoct togaactig-oggagoct togaactig-tig-oggaocot togaactig-tig-oggaa accaggaa 			gycgaccaa cogtt -trocaccaacogtt gycgaccaa cogtt -togaccaacogtagac- -gotoctogaaca -gotoctogaaca -gotoctogaaca -gotoctoga -gotoctoga 	t-tt- t-tt- t-tt- t-tc- tctc- tctc- tctc- tctc- tctc- tctc- tctc- tctc- tct tct
7. 8. 9. 10. 11. 12. 13. 14. 15. 15. 17. 18. 19. 20. 21. 23. 24. 25.	amara silen tritu maian canel arabi lycop vicif cucur lactu pinas picea pando goniu miurh sporm sporm sporm chior volva heter	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	- og togt-gt - og togt-ggt 		giggottig-a-ggaoot dtgactor-aaggaoot aatgacoo gtgg-tig-aggaoot gtgg-tig-aggaoot cogogot-aaag-toc cogogot-aaag-toc cogogot-tig-aggoot tigaattiggoot togaactigttigettig- tigaattig			ggrogaccaa cogtt -tracaccaaacogtt ggrogaccaa cogtt -acocctoggaca -acocctoggaca 	  t-ga-  t-tt- t-nc- t-tc- tcc- tcc-  ctc-  ctc-  ctc-  ctc-  ctc-  ctc-      
7. 8. 9. 10. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 23. 24. 25. 26. 27.	amara silen tritu maian canel arabi lycop vicif cucur lactu picea pando gondu miurh miurh spong spera spor chor volva heter bipol	$\begin{array}{cccccccccccccccccccccccccccccccccccc$		t - goacat 'g - coagt - agiogy 'tg - acctt c - goacat 'g - coagt c - goacag 'goacoa cospactors' - cospococag atgoge 'g - ctcta c - atgoge 'g - ctcta c - atgoge 'g - ctcta c - atgoge 'g - ctcta c - gottgtgtggacattor aggottgtgtggacattor aggottgtgtggacattor tg - t tg - octa tg - tt tg - octa tg - tt tg - octa tg - tt cg - octa tg - tt cg - octa tg - tt cg - octa tg - t cgacc tt ccaacc tt t - caacc	gigsoctig a-ggaocot digactig a-ggaocot digactog -aagacot oggogotgoaaggaacoc diggotoggogot togaactigggoot 			gycgaccaa cogit -tigcaaccaa cogit -tigcaaccaa cogit -gitogttoga - cc- -accctoggaaca -ggitogttoga - cc- acgitogtoga - cc- acgitogtoga - cc- -acgitoga - cc- -acgitoga - cc- titag - ciccaac titag - ciccaac titag - ciccaac titag - ciccaac titag - ciccaac titag - ciccaac 	t-ga- t-ta- t-ga- t-ta- tete tete tete tete tete tete tete tete tete tete tete tete tete- tete tete- tete tete- tete- tete- tete- tete- tete- tete- tete- tete- tete- tete- te
7. 8. 9. 10. 11. 12. 13. <u>14.</u> 15. <u>16.</u> 17. 18. 19. 21. 22. 23. 24. 25. 26. 27. 28.	amara silen tritu maian canel arabi lycop vicif cucur jactu picea pando goniu miurf miurf miurf miurf spong sperm chlor volva heter bipol scler gicas	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	- og togt-gt- - og		giggottig-a-ggaoot dtgactor-aaggaoot astgacot gtgg-tig-aaggaoot gtgg-tig-aaggaoot cogogot-aaag-toc cogogot-togagot togaactig-tig- togactig-tig- cogogot- togactig- togactig- togactig- togactig- togactig- togactig- togac- gaaggaa- accaggaa- accaggaa- accaggaa- accaggaa- accaggaa- accaggaa- accaggaa- accaggaa- gaaggaacc- gcaggaatc- gcaggaatc- gcaggaatc- gtcogagagac- actaggagaac- accaggaa- accagaa- accagaa- accagaa- accagaa- ac			gycgaccaa cogtt- -ttgcaaccaacogtt- gytggaccaa cogtt- -coccoggaca- 	t-ga- t-ga- t-tt- t-tt- tctctc- tct- tctc

1

talin

1.

FIG. 6.—CLUSTALV alignment of ITS2 sequences of angiosperms, chlorophycean algae, and fungi. Window size = 10, gap-opening penalty = 10, gap-extension penalty = 3. Taxon acronyms (are listed in table 1. Taxa are organized by major group: caryophyllalean angiosperms: (1-6); other angiosperm (7-14); conifers (15-16); green algae (17-23); and fungi (24-28).

1290

200

100

										100
SPERM	tt <u>c</u> acccctc	GCTCCACCAC	CCCTTCACCG	GGCGTGTTCG	CCG-CGCCTG	CTCGCGAGAC	CGGCCAGGAG	CGGATGTGAC	CTTCCCAGGA	CCCCCTTCAA
CHLOR	ttca <u>c</u> ccc-c	TCGCCAC	CCTCTCCCCT	GTGGAGAGAG	AAGGCGGAT-	CTGGCCGT	CT	CTGTG		-CTCCT-CGA
PANDO	-taaaa-ctc	АСАСТААААА	ACCTGCTTTT	TTGGTGTG-G	AACTGA	CTGTCTCAT-	GTG	CGCGCAC-A-		AACA
GONIU	-t <u>a</u> atc-ctc	GCTCTCT	CCCTCCCTGT	GTGGGATGAG	AACGGAT-	CTGGCTG-TC	тс	GGCAATC-A-		TATG
SPONG	-taatatcac	-ACCCACATA	CCCTGTGTAT	GTGAT-TCCG	TGTGGAC-	CTGGCCGTTC	CAGTTCAGTG	CTCGACC-AC	CTTTACTGGT	GATC-GAATC
MIURM	-taatatcac	TCTCCCTATA	CCCTGTGTAT	AAC-TGAG	TGGAC-	CTGGTTGTTC	C		GGT	GATCCTT-TC
MIURH	- <u>L</u> AATATCAC	CAATCATTAT	ACTTATGTAT	AAT-CACT	GGTGGAT-	CTGGCTGTTC	CAGCTCTG	C-CGC-AG	CCTAGCTGC-	C-TAGT-
MIURF	-tAATATCAC	T-CCAATCAA	CCTTGTGTTG	AATGCA	-GTGGAT-	CTGGCTGTTC	CAATAC-GT-	CATGAGC	TTGCTCGT	GGTT-TAAC-
										••••
	101									200
SPERM	ACCGGGGGCTG	GGTTGGTTGA	AGACCAGA	GGCTAGAGCA	-GGACCCG-T	CAA-GGGCTT	CAACTGGGTA	GGCGTCCAAG	CTCTTCGGGA	GCTTTCCGCA
CHLOR	GAGAACGG	GGTCGGCTGA	AGAACAGA	GGCTTGAGCA	TGGTCCCGCT	GATAGGGCGC	CAGCTTGGTA	GGTAG	CTCAT	GCTGCATGCC
PANDO	TGCGCATG	GGTTGGTTGA	AGTTTGCAGC	TGCTATTACA	TGGACCCGCT	CAT-GGGCCT	CTACTGGGTA	GGCAG	TTTTTTAATT-	GCTAAATGCT
GONIU	ATTGCC-G	GGTCAGCTGA	AGTGCAGA	GGTTGATGCA	TGGACCCGCT	AAT-GGGCCT	CTACTGGGTA	GGCAA	TTCGTT-	GCTAA-TGCT
SPONG	AGTCACCCTG	GATCGGCTGA	AGAGCAGA	GGCTAAC-TC	TGGACCCGCT	-AA-GGGCCG	CAATTGAGTA	GGCAC	TTGTT-	GCTA-ATACT
MIURM	GGTGAACC-G	GATCAGCTGA	AGAACAGA	GGTTAGCACA	AGGACCCGTT	-AA-GGGCCG	-CAATTGGGTA	GGCAG	TTTTCT	GCTC
MIURH	-GTGCATCTG	GATCTGCTGA	AGTACAGA	GGCTAAC-CA	AGGACCCGTC	-AA-GGGCCG	CAACTGGGTG	GGCAG	AAAT	GCTATATTCT
MIURF	-GTGATG	GATCAGCTGA	AGTGCAGA	GGCTAATGCA	AGGACCC-TT	-AT-GGGCCG	CAACTGGGTA	GGCAG	CTCAC	GCTTGCTATT
	201								289	
SPERM	CAAAGTTG-T	CGCCTGGGAC	CTGC-TGG	-CGGCCCG	CAGGAATC	GTGCCCCGTG	CAC-G	ATCTC-T-AT	-CTtcgacc	
CHLOR	GCTGACAC	CCCTAGGGAC	CTGTGC-TGT	AAAGCCGG	CAGGAATC	GGGGCGCAAG	CTCCGA	ACTCT-TAAT	-T-tcgacc	
PANDO	T-TAGTAGAT	GGCTTAGGAT	CTGTGTGTAT	-TGGCACA-A	ACCAGGAAA-	-TT	A	ACCTTCTTTT	-TCtcgacc	
GONIU	T-TAGTAGAT	GGCTTGGGA-	CTGTGCTTGT	-CGACCCA-A	ACCAGGAAC-	TCAGCCTCGT	GCTGAGCAAA	ACCCT-TTTT	-TCtcgacc	
SPONG	T-TAATTGTT	GGTCTGGGA-	CTATTGN	TCGNCCCT-A	ACCAGGAAA-	ATT	A	ACC-TTATTT	TTCtcgacc	
MIURM	C-TAATTGTT	GGCTTGGGAC	CATGAGC	TCGACCCTCA	ATCAGGAA	-TT	A	ACCCTTATTT	-TCtcgacc	
MIURH	TAGTTGTT	GGCTTGGGA-	CTTTGCTTGT	TCGGCCCTCA	AACAGGAACC	ATT	A	ACC-T-ATTT	-TCtcgacc	
MIURF	TCTAGTTGTT	GGCTTGG-A-	CCATGTGT	TTG-CCCTCA	AACAGGAA	-TT	A	ACC-TTATTT	-TCtcgacc	

FIG. 7.—Manually optimized alignment of chlorophycean ITS2 sequences. Taxon acronyms are listed in table 1. Lowercase letters signify putative 5' and 3' adjacent coding regions. Underlined bases indicate the distalmost bases of ITS2 according to Genbank documentation.

among these groups is unclear, in part because their relationships remain poorly resolved, but also because fossilized coccoid green algae cannot be confidently identified and might belong to many different lineages. Vol-

1

vocales are known from the upper Jurassic (ca. 200 million years ago; Bold and Wynne 1985). The three fungal classes are distinct before the Permian, with zygomycetes and "probable" ascomycetes known as early as



FIG. 8.—Guide tree for ITS2 sequences derived using CLUSTALW. Taxon acronyms are listed in table 1. The tree is derived using neighborjoining on scores for independent pairwise "optimal" alignments using a gap-opening penalty of 25 and gap-extension penalty of 0.3. The tree is a phenogram, rather than phylogram or cladogram: the apparent divergences do not reflect identically aligned base positions across all taxa.

the Silurian (Stewart and Rothwell 1993). Collectively, these results reveal that ITS2 has retained detectably conserved sequence/structure through 350 or more million years of evolution.

ITS hypervariability and poor alignability across major taxonomic groups had previously yielded the opinion that "most of ITS [is] devoid of sequence-specific functions" (Crouch and Bachellerie 1986), i.e., is free to vary randomly. In vivo analyses in yeasts, however, have demonstrated that heterologous ITS2 sequences from only "closely" and not more distantly related yeast species were functional and that mutations/ deletions of conserved ITS2 motifs reduced or eliminated 26S accumulation (van der Sande et al. 1992; van Nues et al. 1994, 1995). Deleting certain combinations of unconserved regions reduced function, and even mutations in unconserved regions reduce cell growth and viability. Evidence for functional constraints on ITS sequence elements in other taxa is lacking, but the yeast experiments suggest that patterns of ITS sequence variation are not random but rather reflect rapid coevolution between sequence and ribosomal processing factors (ribonucleolar proteins and RNAs; van Nues et al. 1995). If this is the case, ITS variation might be useful for guiding taxonomic sampling of the processing factors themselves, which might provide useful phylogenetic evidence, even at levels where ITS itself does not.

#### Conclusions

The purpose of this study was to provide an updated comparison of 5.8S and ITS2 sequences in diverse eukaryotes in order to determine whether the organismic origin of such sequences could be confidently estimated. Our analyses yield several conclusions:

1) Although our results do not challenge the notion that these sequences are not ideal for dccp-level phylogenetic analysis, they do suggest that the deep-level phylogenetic signal contained in 5.8S and ITS2, however small, is generally consistent with that from other molecular and nonmolecular data. Because the ITS region contains at least some diagnostic signal spanning interkingdom- to interspecific-level divergences, the organismal origins, as well as phylogenetic relationships, of any ITS region sequence can be accurately approximated. This result is significant given the increasing popularity of ITS region sequences for lower-level phylogenetic analyses and concomitant increase in the number of taxa for which only ITS sequences are now or ever likely to become available.

2) Our results suggest that the diagnostic limitations of 5.8S sequences do not strictly reflect short sequence length but are also influenced by factors affecting sequences of any length.

3) Because the ITS region comprises, across a relatively short sequence, three structurally and functionally distinct molecules, each of which is the focus of ongoing experimental analysis in yeast and other model organisms, ITS region comparisons among both closely and distantly related organisms should provide a readily exploitable tool for deriving unique insights into DNA sequence evolution in general.

4) Our results demonstrate that both 5.8S and spacer sequence variation at relatively deep phylogenetic levels is nonrandom and thus probably reflects the evolution of underlying elements of the translational apparatus. Thus, ITS region variation should prove useful for experimental analyses of translational apparatus evolution, which in turn might yield phylogenetic data not directly derivable from the ITS region itself.

5) Finally, the importance of exercising caution at all stages of DNA-based research is underscored by the extent of apparent errors and ambiguities in DNA sequence database documentation that we uncovered in the process of analyzing suspected PCR contaminant sequences.

#### Acknowledgments

We are indebted to Dave Swofford for accession and permission to publish results using prerelease versions of PAUP\* 4.0,. We thank Dave Swofford, Paul Lewis, and Jack Sullivan for discussions and insights into inference methods; Liz Zimmer, Matt Kane, Tod Reeder, Robb Brumfield, Bruce Baldwin, and an annymous reviewer for critically reading the manuscript; and participants of the Laboratory of Molecular Systematics journal club for critical comments. M.A.H. was supported by a Smithsonian Postdoctoral Fellowship in Molecular Evolution. L.A.L. was supported by a Postdoctoral Fellowship from a Mellon Foundation grant awarded to the Botany Department, Duke University

#### LITERATURE CITED

- BAKKER, F. T., J. L. OLSEN, and W. T. STAM. 1995. Evolution of nuclear rDNA ITS sequences in the *Cladophora albidal* sericea clade (Chlorophyta). J. Mol. Evol. **40**:640–651
- BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, and M. J. DONOGHUE. 1995.
   The ITS regions of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. 82:247–277.
- BHATTACHARYA, D., and L. MEDLIN. 1995. The phylogen of plastids: a review based on comparisons of small-submit ribosomal RNA coding regions. J. Phycol. **31**:489–4989
- BOLD, H. C., and M. J. WYNNE. 1985. Introduction to the Algae. Prentice-Hall, Englewood Cliffs, N.J.
- BRUNS, T. D., R. VILGALYS, S. M. BARNS, D. GONZALEZED. S. HIBBETT, D. J. LANE, L. SIMON, S. STICKEL, T. M. SZA-RO, W. G. WEISBURG, and M. L. SOGIN. 1992. Evolutionary relationships within the Fungi: analysis of nuclear small subunit rRNA sequences. Mol. Phyl. Evol. 1:231–241.
- CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD et al. (39 co-authors). 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. Ann. Mo. Bot. Gard. 80:528–580.
- COLEMAN, A. W., A. SUAREZ, and L. J. GOFF. 1994. Molecular delineation of species and syngens in volvocacean green algae (Chlorophyta). J. Mol. Evol. 30:808–90.
- CRONQUIST, A. 1981. An integrated system of classification of flowering plants. Columbia Univ. Press, New York, N.Y.
- CROUCH, R. J., and J.-P. BACHELLERIE. 1986. Ribosomal RNA processing sites. Pp. 47–80 in S. K. DUTTA, ed. DNA systematics, Vol. 1. Evolution. CRC Press, Boca Raton, Fla.

- DELWICHE, C. F., M. KUHSEL, and J. D. PALMER. 1995. Phylogenetic analysis of tufA sequences indicates a cyanobacterial origin of all plastids. Mol. Phyl. Evol. 4:110–128.
- DOYLE, J. A., M. J. DONOGHUE, and E. A. ZIMMER. 1994. Integration of morphological and ribosomal RNA data on the origin of angiosperms. Ann. Missouri Bot. Gard. 81: 419-450.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.
- FRIEDL, T. 1995. Inferring taxonomic positions and testing genus level assignments in coccoid green lichen algae: a phylogenetic analysis of 18S ribosomal RNA sequences from *Dictyochloropsis reticulata* and from members of the genus *Myrmecia* (Chlorophyta, Trebouxiophyceae cl. nov.). J. Phycol. **31**:632-639.
- FRITZ, G. N., J. CONN, A. COCKBURN, and J. SEAWRIGHT. 1994. Sequence analysis of the ribosomal internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culidae). Mol. Biol. Evol. **11**:406–416.
- GARGAS, A., P. T. DEPRIEST, M. GRUBE, and A. TEHLER. 1995. Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. Science **268**:1492–1495.
- GARBARY, D. J., and P. W. GABRIELSEN. 1990. Taxonomy and evolution. Pp. 477–498 in K. M. COLE and R. G. SHEATH, eds. Biology of the red algae. Cambridge University Press, Cambridge.
- GARBARY, D. J., K. S. RENZAGLIA, and J. G. DUCKETT. 1992. The phylogeny of land plants: a cladistic analysis based on male gametogenesis. Plant Syst. Evol. **188**:237–269.
- GENETICS COMPUTER GROUP. 1994. Program Manual for the Wisconsin Package, version 8. Genetics Computer Group, Madison, Wis.
- HENRY, Y., H. WOOD, J. MORRISEY, E. PETFALSKI, S. KEAR-SEY, and D. TOLLERVEY. 1994. The 5' end of yeast 5.8S rRNA is generated by exonuclease from an upstream cleavage site. EMBO J. 13:2452-2463.
- HERSHKOVITZ, M. A., and E. A. ZIMMER. 1996. Conservation patterns in angiosperm rDNA-ITS2 sequences. Nucl. Acids. Res. 24:2857–2867.
- HIGGINS, D. G. A., J. BLEASBY, and R. FUCHS. 1992. Clustal V: improved software for multiple sequence alignment. CA-BIOS 8:189–191.
- HILLIS, D. M., J. P. HUELSENBECK, and D. L. SWOFFORD. 1994. Hobgoblin of phylogenetics? Nature **369**:363–364.
- HOLMBERG, L., Y. MELANDER, and O. NYGÅRD. 1994. Probing the conformational changes in 5.8S, 18S, and 28S rRNA upon association of derived subunits into complete 80S ribosomes. Nucleic Acids Res. 22:2776–2783.
- HUSS, V. A. R., and M. L. SOGIN. 1990. Phylogenetic position of some *Chlorella* species within the Chlorococcales based upon complete small-subunit ribosomal RNA sequences. J. Mol. Evol. **31**:432–442.
- KRANZ, H. D., D. MIKS, M.-L. SIEGLER, I. CAPESIUS, and C. W. SENSEN. 1995. The origin of land plants: phylogenetic relationships among chlorophytes, bryophytes, and vascular plants inferred from complete small-subunit ribosomal RNA gene sequences. J. Mol. Evol. 41:74–84.
- LEIPE, D. D., P. O. WAINWRIGHT, J. H. GUNDERSON, D. POR-TER, D. J. PATTERSON, F. VALOIS, S. HIMMERICH, and M. L. SOGIN. 1994. The stramenopiles from a molecular perspective: 16S-like sequences from *Labyrinthuloides* and *Cafeteria roenbergensis*. Phycologia 33: 369–377.
- LINDAHL, L., R. H. ARCHER, and J. M. ZENGEL. 1994. Alternate pathways for processing the internal transcribed spacer 1 in pre-rRNA of *Saccharomyces cerevisiae*. Nucleic Acids Res. 22:5399–5407.

- LOCKHART, P. J., PENNY, D., M. A. STEEL, and M. D. HENDY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605–612.
- LIU, J.-S., and C. L. SCHARDL. 1994. A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. Plant Mol. Biol. 26:775–778.
- LYGEROU, Z., P. MITCHELL, E. PETFALSKI, B. SÉRAPHIN, and D. TOLLERVEY. 1994. The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins. Genes & Dev. 8:1423-1433.
- MATTOX, K. R., and K. D. STEWART. 1984. Classification of the green algae: a concept based on comparative cytology. Pp. 29–72 *in* D. E. G. IRVINE and D. M. JOHN, eds. Systematics of the green algae. Academic Press, London.
- MISHLER, B. M., L. A. LEWIS, M. A. BUCHHEIM, K. S. REN-ZAGLIA, D. J. GARBARY, C. F. DELWICHE, F. W. ZECHMAN, T. S. KANTZ, and R. L. CHAPMAN. 1994. Phylogenetic relationships of the "green algae" and "bryophytes". Ann. Mo. Bot. Gard. **81**:451–483.
- PAGE, C. N. 1990. Interrelationships between families of conifers. Pp. 292–293 in K. U. KRAMER and P. S. GREEN, eds. The families and genera of vascular plants. Springer-Verlag, Berlin.
- RITLAND, C., K. RITLAND, and N. A. STRAUS. 1993. Variation in the internal transcribed spacers (ITS1 and ITS2) among eight taxa of the *Mimulus guttatus* species complex. Mol. Biol. Evol. 10:1273-1278.
- RITLAND, C., and N. A. STRAUS. 1993. High evolutionary divergence of the 5.8S ribosomal DNA in *Mimulus glauces*cens (Scrophulariaceae). Plant Mol. Biol. 22:691–696.
- SCHMICKEL, R. D., J. SYLVESTER, D. STAMBOLIAN, and I. L. GONZALES. 1990. The ribosomal DNA sequence for the study of the evolution of human, chimpanzee, and gorilla. Pp. 11–32 *in* S. K. DUTTA and W. P. WINTER, eds. DNA systematics, Vol. 3. Human and higher primates. CRC Press, Boca Raton, Fla.
- SMITH, S. 1994. GDE version 2.2. University of Illinois, Urbana, Ill.
- STEELE, K. P., K. E. HOLSINGER, R. K. JANSEN, and D. W. TAYLOR. 1991. Assessing the reliability of 5S rRNA sequence data for phylogenetic analysis. Mol. Biol. Evol. 8: 240–248.
- STEINKÖTTER, J., D. BHATTACHARYA, I. SEMMELROTH, C. BI-BEAU, and M. MELKONIAN. 1994. Prasinophytes form independent lineages within the Chlorophyta: evidence from ribosomal RNA sequence comparisons. J. Phycol. **30**:340– 345.
- STEWART, W. N., and G. W. ROTHWELL. 1993. Paleobotany and the evolution of plants, 2nd edition. Cambridge University Press, Cambridge, England.
- SUH, Y., L. B. THIEN, and E. A. ZIMMER. 1992. Nucleotide sequences of the internal transcribed spacers and 5.8S rRNA gene in *Canella winterana* (Magnoliales; Canellaceae). Nucleic Acids Res. 20:6101-6102.
- SWOFFORD, D. L. 1996. PAUP\* version 4.0. Sinauer Assoc., Inc., Sunderland, Mass. (in press).
- SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic Inference. Pp. 407–514 in D. M. HILLIS, C. MORITZ, and B. K. MABLE, eds. Molecular systematics, 2nd edition. Sinauer Associates, Inc. Sunderland, Mass.
- THOMPSON, A. J., and D. L. HERRIN. 1994. A chloroplast group I intron undergoes the first step of reverse splicing into host cytoplasmic 5.8S rRNA. J. Mol. Biol. **236**:455–468.
- THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, po-

sition specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-4680

- TROITSKY, A. V., and V. K. BOBROVA, 1986, 23s rRNA-derived small ribosomal RNAs: their structure and evolution with references to plant phylogeny. Pp. 137-170 in S. K. DUTTA, ed. DNA systematics, Vol. 2. Plants. CRC Press, Boca Raton, Fla.
- TROITSKY, A. V., Y. F. MELEKHOVETS, G. M RAKHIMOVA, V. K. BOBROVA, K. M. VALJEJO-ROMAN, and A. S. ANTONOV. 1991. Angiosperm origins and early stages of seed plant evolution deduced from rRNA sequences. J. Mol. Evol. 32: 253-261.
- VAN DER SANDE, C. A. F. M., M. KWA, R. W. VAN NUES, H. VAN HEERIKHUIZEN, H. A. RAUÉ, and R. J. PLANTA. 1992. Functional analysis of internal transcribed spacer 2 of Saccharomyces cerevisiae ribosomal DNA. J. Mol. Biol. 223: 899-910.
- VAN NUES, R. W., J. M. J. RIENTES, S. A. MORRÉ, E. MOLLEE, R. J. PLANTA, J. VENEMA, and H. A. RAUÉ. 1995. Evolutionarily conserved structural elements are critical for processing of internal transcribed spacer 2 from Saccharomyces cerevisiae precursor ribosomal RNA. J. Mol. Biol. 250: 24 - 36.
- VAN NUES, R. W., J. M. J. RIENTES, C. A. F. M. VAN DER SANDE, S. F. ZERP, C. SLUITER, J. VENEMA, R. J. PLANTA, and H. A. RAUÉ. 1994. Separate structural elements within

internal transcribed spacer 1 of Saccharomyces cerevisiae precursor ribosomal RNA direct the formation of 17S and 26S rRNA. Nucleic Acids Res. 22:912-919.

- WESSON, D. M., C. H. PORTER, and F. H. COLLINS. 1992. Sequence and secondary structure comparisons of ITS rDNA in mosquitoes (Diptera: Culicidae). Mol. Phylogenet. Evol. 1:253-269.
- WHEELER, W. C., and R. L. HONEYCUTT. 1988. Paired sequence differences in ribosomal RNAs: evolution and phylogenetic implications. Mol. Biol. Evol. 5:90-96.
- WILCOX, L. W., L. A. LEWIS, P. A. FUERST, and G. L. FLOYD. 1992. Assessing the relationships of autosporic and zoosporic chlorococcalean green algae with 18S rDNA sequence data. J. Phycol. 28:381-386.
- YANG, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. 39:306-314.
- YEH, L.-C. C., and J. C. LEE. 1991. Higher-order structure of the 5.8S rRNA sequence within the yeast precursor ri氧osomal RNA synthesized in vitro. J. Mol. Biol. 217:649-659. ded from https://aca

BARBARA A. SCHAAL, reviewing editor

Accepted July 24, 1996

APPENDIX

# Alignment of 5.8S Sequences Used for Phylogenetic Analysis in This Study, Including Gap and Missing Data Designations<sup>a</sup>

		1									100	
1	ARABI	AACGACTCTC	GGCAACGGAT	ATCTCGGCTC	TCGCATCGAT	GAAGAACGTA	GCGAAATGCG	ATACTTGGTG	TGAATTGCAG	AATCCCGTGA	ACCATCGAGT	
2	LYCOP	AACGACTCTC	GGCAACGGAT	ATCTCGGCTC	TCGCATCGAT	GAAGAACGTA	GCGAAATGCG	ATACTTGGTG	TGAATTGCAG	AATCCCGTGA	ACCATCGAGT	
3	CUCUR	AACGACTCTC	GGCAACGGAT	ATCTCGGCTC	TCGCATCGAT	GAAGAACGTA	GCGAAATGCG	ATACTTGGTG	TGAATTGCAG	GATCCCGCGA	ACCACCGAGT	
5	LACTU	AACGACTCTC	GGCAACGGAT	ATTTCGGCTC	ACGCATCGAT	GAAGAACGTA	GCAAAATGCG	ATACTTGGTG	TGAATTGCAG	AATCCCGTGA	ACCATCGAGT	
6	SILEN	AACGACTCTC	GGCAACGGAT	ATCTCGGCTC	TCGCATCGAT	GAAGAACGTA	GCGAAATGCG	ATACTT-GTG	TGAATTGCAG	AATCCCGTGA	ACCATCGAGT	
7	vicia	AATGACTCTC	GGCAACGGAT	ATCTAGGCTC	TTGCATCGAT	GAAGAACGTA	GCGAAATGCG	ATACTTGGTG	TGAATTGCAG	AATCCCGTGA	ACCATCGAGT	
8	CANEL	GATGACTCTC	GGCAACGGAT	ATCTCGGCTC	TCGCATCGAT	GAAGAACGTA	GCAAAATGCG	ATACTTGGTG	TGAATTGCAG	AATCCCGCGA	ACCATCGAGT	
10	ephed	TACGACTCTC	GGCAATGGAT	ATCTCGGCTC	TCGCATCGAT	GAAGAACGTA	GCGARATGCG	ATACTTAGTG	TGAATTGCAG	AATCCCGTGA	ATCATCGAGT	_
11	gnetu	CACGACTCT-	GACAATGGAT	ATCTCGGCTC	TCGT-TCGAT	GAAGAACGTA	GCGAAATGCG	ATACTTGGTG	TGAATTGCAG	AATCCCGTGA	ATCATCGAAT	
12	PINUS	AACGACTCTC	GTCAACGGAT	ATCTCGGCTC	TTGTTACGAT	GAAGAACGTA	-CGAAATGCG	ATACTTAGTG	TGAATTGCAG	AATCCCGTGA	ATCATCGAGT	_
13	PICEA	?ATGACTCTC	GGC-A-GGA-	ATCTCGGCTC	TTGTCAACA-	GAAGAACGTA	GCGA-ATGCG	ATACTTAGTG	TGAATTGCAG	AATCCCGTGA	ATCATCGAGT	
15	marei	2AGGACTCTC	ACCAACGGAT	ATCTCGGCTC	TUGC-AUGAT	GAAGAACGTA	GCGAAATGCG	ATACTTAGTG	TGAATTGCAG	AATCCCGTGA	ATGATCGAGT	-0
16	osmun	?ACGACTCTC	AGCAACGGAT	ATCTTGGCTC	TTGCAACGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAACTGCAG	AATTCCGCGA	ATAATCGAGT	8
17	mnium	?ATAACCCTC	AGCAACGGAT	ATCTTGGCTC	TTGCAACGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AATTCCGCGA	ATCATCGAGT	<u>n</u>
18	CHLAM	???AACTCTC	AACAACGGAT	ATCTTGGCTC	TCGGATCGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AAATACGTGA	ATCATCGAAT	- Q
19	GONIU	GACAACTCTC	AACAACGGAT	ATCTTGGCTC	TCAGATCGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AAATACGTGA	ATCATCGAAT	de
20	MTURH	GACAACTCTC	AACAACGGAT	ATCTTGGCTC	TCAGATCGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AAATACGTGA	ATCATCGAAT	d
22	MIURF	GACAACTGTG	TACAACGGAT	ATCTTGGCTC	TCGTAACGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AATTCCGTGA	ACCATCAAAT	5
23	SPERM	GACAACTCTC	AACAACGGAT	ATCTTGGCTC	TCGTAACGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AATTCCGTGA	ACCATCGAAT	В
24	TETRA	GACAACTCTC	AACAACGGAT	ATCTTGGCTC	TTACAACGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AATTCCGTGA	ACCATCGAAT	htt
25	CHLOR	GACAACTCTC	TACAACGGAT	TTCTTGGCTC	CCGCATCGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAGTG	TGAATTGCAG	AATTCCGTGA	ATCATCGAAT	sd
27	HETER	TACAACTTTC	AACAATGGAT	CTCTCGGCTC	TCGCATCGAT	GAAGAACGCA	GCGAAATGCG	ATAAGTAATG	TGAATTGCAG	AATTCAGTGA	ATCATCGAAT	-
28	VOLVA	TACAACTTTC	AACAACGGAT	CTCTTGGCTC	TCGCATCGAT	GAAGAACGCA	GCGAAATGCG	ATAAGTAATG	TGAATTGCAG	AATTCAGTGA	ATCATCGAAT	ac
29	SCLER	TAAAACTTTC	AACAACGGAT	CTCTTGGTTC	TGGCATCGAT	GAAGAACGCA	GCGAAATGCG	ATAAGTAATG	TGAATTGCAG	AATTCAGTGA	ATCATCGAAT	ad
30	BIROL	??AAACTTTC	AACAACGGAT	CTCTTGGTTC	TCGCATCGAT	GAAGAACGCA	GCGAAATGCG	ATACGTAATG	TGAATTGCAG	AATTCCGTGA	ATCATCGAAT	en
32	GLOMU	GATCACTITC	AACAACGGAT	CTCTTGGCTC	TCGCATCGAT	GAAGAACGCA	GCGAAATGCG	ATAAGTAGTG	TGAATTGCAT	AATTTTGTGA	ATCATCGAAT	lic
33	CYANI	AGAAGCTCTC	AACGGTGGAT	ATCTCGGCTC	TCTTGACGAT	GAAGAACGCA	-CAAAAAGCG	AAATGTAGTG	TGAACTGCAA	GTTTTCGTGA	ATCATTGAAT	- <u>`</u> 2
34	SARCO	TACAACTCAT	GACGGTGGAT	GTCTTGGCTC	CAGCAACGAT	GACGAACGCA	GCGAATTGCG	AAA-GTAGTA	CGAACTGCAG	AA-CCCGTGA	ATCATCGAAT	٦p.
35	PHYTO	AGCAACTTTC	AGCAGTGGAT	GTCTAGGCTC	GCACATCGAT	GAAGAACGCT	GCGAACTGCG	ATACGTAATG	CGAATTGCAG	GATTCAGTGA	GTCATCGAAA	8
37	crvpt	?ACAACTTTC	AGCAGTTGAT	TCCTTGGTTC	AGACCTCGAT	GAAGGGCACT	GCGAAA-GTG	A-ATGGCATG	TGAA-TGCAG	GCATCCGGGA	ATTGAGAGCT	m)
38	PRORO	TACAACTTTC	AGCGACGGAT	GTCTCGGCTC	GAACAACGAT	GAAGGGCGCA	GCGAAATGTG	ATAAGCAATG	TGAATTGCAG	AATTCCGTGA	ACCAATAGGG	m
												õ
												Ð
												e/art
		101						170	[ 12	2-145	j	e/articl
1	ADART	101   СТТТСААССС		C222CAAGCC	TTCT2CCCCC	2222866668	CGTCTCCCTC	170   	[ 12	2-145	]	e/article/1
1 2	ARABI LYCOP	101   CTTTGAACGC CTTTGAACGC	AAGTTGCGCC	C???CAAGCC C???GAAGCC	TTCT?GGCCG ATTT?GGCCG	????AGGGCA ????AGGGCA	CGTCTGCCTG	170   GGTGTCACAA GGCGTCACG?	[ 12   ???????????????????????????????????	2-145	]   ?? ??	e/article/13/
1 2 3	ARABI LYCOP CUCUR	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC	TTCT?GGCCG ATTT?GGCCG TTCT?GGCCG	????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG	170   GGTGTCACAA GGCGTCACG? GGCGTCACGC	[ 12 ] ?????????? ????????? ??????????????	<b>2-145</b> ????????????????????????????????????	]   ?? ?? ??	e/article/13/9/1
1 2 3 4 5	ARABI LYCOP CUCUR triti	101 j CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC	C???CAAGCC C???GAAGCC C???GAGCC C???GAGGCC	TTCT?GGCCG ATTT?GGCCG TTCT?GGCCG ACTC?GGCCG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGCCTGCCT	170   GGTGTCACAA GGCGTCACG? GGCGTCACGC GGCGTCACGC	[ 12 ] ?????????? ????????? ??????????????	2-145 ????????????????????????????????????	] 77 77 77 77	e/article/13/9/127
1 2 3 4 5 6	ARABI LYCOP CUCUR triti LACTU SILEN	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC	C???CAAGCC C???GAAGCC C???GGAGCC C???GAAGCC C???GAAGCC C???GAAGC2	TTCT?GGCCG ATTT?GGCCG TTCT?GGCCG ATCC?GGCCG ATCC?GGCTG -TTC?GGCTG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGCCTGCCT	170 J GGTGTCACAA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC	[ 12 ????????? ????????? ????????? ????????	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ??	e/article/13/9/1276/
1 2 3 4 5 6 7	ARABI LYCOP CUCUR triti LACTU SILEN vicia	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC	C???CAAGCC C???GAAGCC C???GAGGCC C???GAAGCC C???GAAGCC C???GAAGC? C???GATGCC	TTCT?GGCCG ATTT?GGCCG ACTC?GGCCG ACTC?GGCCG -TCC?GGCTG -TTC?GGCTG ATTA?GGTTG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGCCTGCCT	170 J GGTGTCACAA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC	[ 12 ????????? ??????????????????????????	2-145 ????????????????????????????????????	] 1 ?? ?? ?? ?? ?? ??	e/article/13/9/1276/99
12345678	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC	C???CAAGCC C???GAAGCC C???GAGGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GATGCC C???GAGCCC	TTCT?GGCCG ATTT?GGCCG ACTC?GGCCG ACTC?GGCCG -TC?GGCTG ATTA?GGTTG ACTA?GGCTG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCT	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 J GGTGTCACAA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC	[ 12 ????????? ????????? ?????????? ???????	2-145 ????????????????????????????????????	] 1 ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/9914
123456789	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC	C???CAAGCC C???GAAGCC C???GAGGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???TAAGCT	TTCT?GGCCG ATTT?GGCCG TTCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGTTG ACTA?GGCTG <u>TCC?GCCG?GGCCG</u>	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCT ????AGGGCT	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991468
1 2 3 4 5 6 7 8 9 10 11	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC	C???CAAGCC C???GAAGCC C???GAGGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGTG TCCG?GGCCG -?TC?GGCCG -?TC?GGCCG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCCACGC GGCGTCCAAGC GGCGTCGCAA	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?	e/article/13/9/1276/991469 b
1 2 3 4 5 6 7 8 9 10 11 12	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AAGTTG-GCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGC? C???GAAGCC C???GAAGCC C???GAAGCC CTCCG??GC? CTCCG??GCAGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGTG ACTA?GGCTG TCCG?GGCCG -?TC?CGCCA -?TC?CGCCA	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 J GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACAGC GGCGTCACAGC GGCGTCACAGC GGCGTCGCAA GGCGTCGCAA	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	25 25 25 25 25 25 25 25 25 25 25 25 25 2	e/article/13/9/1276/991469 by <u>c</u>
1 2 3 4 5 6 7 8 9 10 11 12 13	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AAGTTG-GCC AATTTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGC? CTCCG??GCAGCC C???GAGCC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCG ATCC?GGCTG ATTA?GGTG ACTA?GGCCG -?TC?GGCCG -?TC?GGCCG -?TC?GGTCG -?TC?GGTCG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 J GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCCACAGC GGCGTCCACAGC GGCGTCCGCAA GGCGTCGCAA GGCGTCGCAT	[ 12 ????????????????????????????????????	2-145 7777777777 7777777777 7777777777 777777	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?	e/article/13/9/1276/991469 by guŧ
1 2 3 4 5 6 7 8 9 10 11 12 13 14	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AAGTTG-GCC AATTTGCGCC AATTTGCGCC AATTTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGC? C???GAAGC? C???GAGC? C???GAGC? C???GAGC?	TTCT?GGCCG ATTT?GGCCG TTCT?GGCCG ACTC?GGCCG ATTC?GGCTG ATTA?GGTTG CCG?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGTCG -7TC?GGTCG -7TC?GGTCG	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA	[ 12 1 1 1 1 1 1 1 1 1 1 1 1 1	2-145 ????????????????????????????????????	] ;; ;; ;; ;; ;; ;; ;; ;; ;; ;	e/article/13/9/1276/991469 by guest
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS marsi osemir	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTGCGCC AATTGCGCC AATTGCCCC AATTGCCCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGCC C???GAGCC? C???GAGCC? C???GAGCC? C???GAGCC? C???GAGCC?	TTCT?GGCCG ATTT?GGCCG ACTC?GGCCG ATCC?GGCCG -TTC?GGCCG ATTA?GGTG ATTA?GGTG -7TC?GGCCA -7TC?GGCCG -7TC?GGTCG -7TC?GGTCG -7TC?GGTCG -7TC?GGTCG -7TC?GGTCG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170   GGTGTCACGA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAT GGCGTCGCAT GCCGTCCCAC?	[ 12 ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;	2-145 ????????????????????????????????????	] 7? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest or
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PICEA FICEA marsi osmur	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AATTGCGCC AATTGCGCC AATTGCGCC AATTG-GCC AAGTTG-GCC AAGTTG-GCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGCC C???GAGGC? C???GAGC? C???GAGC? C???GAGC? C???GAGCC? C???GAGCC? C???GAGCC? C???GAGCC? C???GAGCC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATC?GGCTG -TTC?GGCTG ATTA?GGTG ATTA?GGTG -?TC?GGCCG -?TC?GGCCG -?TC?GGTCG ?-TC?GGTCCG ?-TC?GTCCG ?-TC?GTCCG	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170   GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAT GGCGTCGCAT GGCGTCCCAC? AGCGTCCCC?	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ;; ;; ;; ;; ;; ;; ;; ;; ;; ;	e/article/13/9/1276/991469 by guest on 2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed ghetu PINUS PICEA TAXUS marsi osmun mnium CHLAM	101 J CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTGCGCC AATTGCGCC AAGTTG-GCC AAGTTG-GCC AAGTTGCGCC ATGTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGCC C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGTG TCCG?GGCG -7TC?GGCCG -7TC?GGCCG -7TC?GGTCG -7TC?GGTCG 7-TC?GTCCG -TC?GTCCG -TC?GTCCG -TC?GCCCA	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGCA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA ACGGTCCACC? AGCGTCCGCA	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL Potam ephed gnetu PINUS PICEA TAXUS marsi osmun mlum CHLAM GONIU	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AAGTTG-GCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AAGTTG-GCC AAGTTG-GCC ATATTGCGCT	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCG -TTC?GGCG ATTA?GGCG -TTC?GGCG -7TC?GGCG -7TC?GGCG -7TC?GGTCG -7TC?GGTCG -TC?GTCCG ?-TC?GTCCG -TC?GTCCG -TTC?GGCCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA	CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG CGTCTGCCTG TGCCTGCC	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCCGCAT GGCGTCCACC? AGCGTCCACC? AGCGTCCGCGT	[ 12 ] ?????????? ???????????????????????	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 Au
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS marsi osmun mnium CHLAM GONIU PANDO	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AATTTGCGCC AATTTGCGCC AAGTTG-GCC AAGTTG-GCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGCC? C???GAGCC? C???GAGCC? C???GAGCC? C???GAGCC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATC?GGCCG ATC?GGCTG ATTA?GGTG CCG?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG 7-TC?GTCCG -7TC?GTCCG -7TC?GCCCA -TC?GGCCA -TTC?GGCCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCGCAA AGCGTCCACC? AGCGTCACC? AGCGTCACCG AGCGTCACGGT AGCGTCACGGT	[ 12 1 1 1 1 1 1 1 1 1 1 1 1 1	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 Augu
1 2 3 4 5 6 7 8 9 101 112 13 14 15 16 17 18 19 20 22 22	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL potam ephed ephed PICEA TAXUS marsi Osmium CHLAM GONIU PANDO MIURH	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AAGTTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AAATTGCGCC AAATTGCGCC	C???CAAGCC C???GAGGCC C???GAGGCC C???GAGGCC C???GAGGCC C???GAGGCC C???GAGGCC C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCCG ATCC?GGCCG ATTA?GGTG ATTA?GGTG ATTA?GGCCG -?TC?GGCCA -?TC?GGCCG -?TC?GGCCG ?TC?GGCCG ?TC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA	????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGGCA ????AGGCA ????AGAGCA ????AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170   GGTGTCACGA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGCAT AGCGTCCGCAT AGCGTCCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT	[ 12 1 1 1 1 1 1 1 1 1 1 1 1 1	2-145 ????????????????????????????????????	] ;; ;; ;; ;; ;; ;; ;; ;; ;; ;	e/article/13/9/1276/991469 by guest on 21 August
1 2 3 4 5 6 7 8 9 10 11 12 13 4 15 16 17 18 19 20 22 23	ABABI LYCOP CUCUR triti ILACTU SILEN vicia CANEL potam ephed PINUS PICEA TAXUS marsi OSMUN CHLAM GONIU PANDO MIUEH MIURF SPERM	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AAGTTG-GCC AATTTGCGCC AATTGCGCC AATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGCC C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATC?GGCTG -TTC?GGCTG ATTA?GGCTG ATTA?GGCTG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG ?-TC?GTCCG -TC?GTCCG -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA	7777AGGGCA 7777AGGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAT GGCGTCGCAT GGCGTCGCAT AGCGTCGCAT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ;; ;; ;; ;; ;; ;; ;; ;; ;; ;	e/article/13/9/1276/991469 by guest on 21 August 20
1 2 3 4 5 6 7 8 9 10 11 12 13 4 15 6 7 8 9 10 11 12 13 4 15 16 17 18 19 20 22 23 24	ARABI LYCOP CUCUR triti ILACTU SILEN vicia CANEL PINUS PICEA TAXUS marsi OSMUN GONIU PANDO MIURH MIURF SPERM TETRA	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTTGCGCC AATTTGCGCC AATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGCC C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATC?GGCTG -TTC?GGCTG -TTC?GGCTG -TTC?GGCTG -TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -TC?GTCCG -TC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGCAT AGCGTCCGCAT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ] ] ] ] ] ] ] ] ] ] ] ] ] ] ] ] ] ]	e/article/13/9/1276/991469 by guest on 21 August 2022
$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$	ARABI LYCOP CUCUR triti ILACTU SILEN vicia CANEL potam ephed gnetu PINUS PICEA TAXUS marsi osmun mnium CHLAM GONIU PANDO MIURH MIURF	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTTGCGCC AATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGTG <u>TCC?GGCCG</u> -TTC?GGCCG -TC?GGCCG -TC?GGCCG -TC?GGCCG -TC?GGCCG -TC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGCAA AGCGTCCGCAA AGCGTCCGGGT AGCGTCGGCGT AGCGTCGGG	[ 12 ????????????????????????????????????	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?	e/article/13/9/1276/991469 by guest on 21 August 2022
$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$	ARABI LYCOP CUCUR triti LACTU SILEN vicia canel ophedu Gephed gnetu PINUS PICEA TAXUS marsi OSMUN GONIU PANDO MIURH MIURH SPERM TETRA CHLOR CHLOR CHLOR	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGCC C???GAGCC? C???GAGC? C???GAGC? C???GAGC? C???GAGGC?	TTCT?GGCCG ATTT?GGCCG ATC?GGCCG ATC?GGCCG ATC?GGCCG ATTA?GGTCG CCG?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCGCAC AGCGTCGCAC AGCGTCGCGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT	[ 12 1 1 1 1 1 1 1 1 1 1 1 1 1	2-145 ????????????????????????????????????	] ;; ;; ;; ;; ;; ;; ;; ;; ;; ;	e/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 10 11 2 13 4 15 6 7 8 9 10 11 2 13 4 15 6 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	ARABI LYCOP CUCUR triti LACTU SILEN vicia canel ophed qnetu PINUS PICEA TAXUS marsi OSNIU PICEA TAXUS marsi OSNIU PANDO MIURH MIURH MIURH MIURH MIURH MIURH MIURH MIURH MIURH	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AATTGCGCC AATTGCGCC AAGTTG-GCC AAGTTG-GCC AAGTTG-GCC AAGTTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCT ACATTGCGCT	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGCC C???GAGGC? C???GAGC? C???GAC? C???GAC? C???GAC? C???GAC? C????GAC? C???C? C???C?	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGCTG -7TC?GGCCA -7TC?GGCCA -7TC?GGCCA -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -TTC?GGCCA	2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170   GGTGTCACAA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGCAA AGCGTCCGCAT AGCGTCCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGT AGCGTCGGT AGCGTCGGT AGCGTCGGT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGGT AGCGTCGGGT AGCGTCGGCT AGCGTCGGGT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGT AGCGTCGGT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGGCT AGCGTCGCGCT AGCGTCGCGCT AGCGTCGCGCT AGCGTCGCGCT AGCGTCGCGCCGCT AGCGTCGCGC AGCGCTCGCA AGCGCTCGCA AGCGCTCGCA AGCGCTCGCA AGCGCTCGCA AGCGC	[ 12 777777777 777777777 777777777 777777	2-145 ????????????????????????????????????	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 10 11 12 13 4 15 6 7 8 9 10 11 12 13 4 15 6 7 2 2 2 4 2 5 6 7 2 2 9	ABABI LYCOP CUCUR triti ILACTU SILEN vicia cANEL potam ephed qnetu PICEA MIUS PICEA MIUS MATSI CHIAM GONIU PANDO MIURH MIURF SPERM TETRA CHLOR CLADO HETER VOLVA SCLER	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTG-GCC AATTGCGCC AATTGCGCC AATTGCGCC ATATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC ACCTTGCGCC ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAC? C??????????	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATC?GGCTG -TTC?GGCTG ATTA?GGCTG ATTA?GGCTG -7TC?GGCCA -7TC?GGCCA -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -TTC?GGCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCCCA -TTC?GCA -TTC?GCA -TTC?GCA -TTC?GCA -TTC?GCA -TTC?GCCA -TTC?G	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCGCAA AGCGTCGGCA AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGT AGCGTCGGT AGCGTCGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGTT7 AGCGTCGT77 AGCGTCGT77 AGCGTCAT77	[ 12 ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;	2-145 2777777777 2777777777 27777777777 277777777	] ;? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 101 112 13 14 15 6 17 18 19 0 21 22 23 24 25 26 77 28 9 30	ARABI LYCOP CUCUR triti ILACTU SILEN vicia cANEL PINUS PICEA TAXUS marsi OSMUL MAINER GONIU PANDO KILAM GONIURH MIURH MIURF SPERM TETRA CHLOR CLADO HETER VOLVA SCLER SACCH	101 CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTTGCGCC AATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC ACATTGCGCC ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C????GAGC? C????GAGC? C????GAGC? C????C????	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATC?GGCTG -TTC?GGCTG -TTC?GGCTG -TTC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -TC?GTCCG -TC?GTCCG -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGGCA 7777AGGCA 7777AGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAT GGCGTCGCAT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGT7 AGCGTCGCGT7 AGCGTCGGT7 AGCGTCGGT7 AGCGTCGCGT7 AGCGTCGGT7 AGCGTCGCGT7 AGCGTCGCGT7 AGCGTCGCGT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGGT7 AGCGTCGCATT7 AGCGTCGGT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCATT7 AGCGTCGCGCAT	[ 12 ????????????????????????????????????	2-145 2777777777 27777777777 27777777777 277777777	] ;; ;; ;; ;; ;; ;; ;; ;; ;; ;	e/article/13/9/1276/991469 by guest on 21 August 2022
123456789101121341561718902122342226789312	ARABI LYCOP CUCUR triti LACTU SILEN vicia CANEL POLAM ephed qnetu PINUS PICEA TAXUS marsi OSMUN GONIU PANDO MIURH MIURF SPERM TETRA CHLAM GONIU FATER SPERM TETRA CHLOR CLADO LETRO SILER SACCHOR SILER SACCHOR SILER SACCHOR SILER SACCHOR SILER SACCHOR SILER SACCHOR SILER SACCHOR SILEN SILE	101 I CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C???GAGGC? C????GAGGC? C????GAGC? C????C????	TTCT?GGCCG ATTT?GGCCG ATC?GGCCG ATC?GGCCG ATC?GGCTG -TTC?GGCTG -TTC?GGCTG -TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -TC?GGCCG -TC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TTC?GGCCA -TC?GCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GGCCA -TC?GCCA -TC	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGCAT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGT7 AGCGTCGT7? AGGGTCGGT? AGGGTCGCGT? AGGGTCGCGCGCGT? AGGGTCGCGT? AGGGTCGCGT?	[ 12 1 1 1 1 1 1 1 1 1 1 1 1 1	2-145 2777777777 2777777777 2777777777 277777777	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?	e/article/13/9/1276/991469 by guest on 21 August 2022
123456789111123415678920312 111123415678920332 222222222222222222222222222222222	ARABI LYCOP CUCUR triti LACTU SILEN vicia cantel gpetu Picea TAXUS marsi OSMUN FICEA TAXUS MAURH MIURH MIURH MIURH SPERM TETRA CHLOR CHLOR CHLOR CHLOR CHLOR SICER Sacch BIPOL GLOMU CVANT	101 I CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGCC C???GAGCC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C???GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C??????????	TTCT?GGCCG ATTT?GGCCG ATC?GGCCG ATC?GGCCG ATC?GGCCG ATTA?GGTG ATTA?GGTG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCA -7TC?GGCCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GGC	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGCA 7777AGGCA 7777AGGCA 7777AGGCA 7777AGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAAG GGCGTCGCAA GGCGTCGCAA GGCGTCGCAC AGCGTCGCAC AGCGTCGCAC AGCGTCGCAC AGCGTCGCGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT	[ 12 777777777 777777777 777777777 777777	2-145 2777777777 2777777777 27777777777 277777777	] ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022
$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$	ARABI LYCOP CUCUR triti LACTU SILEN vicia canel ophed qnetu PINUS PICEA TAXUS marsi OSMUN GONIU PANDO MIURH MIURH MIURH MIURH MIURH MIURH SPERM TETRA CHLOR CLADO HETER SACCH GLOMU CYANI SARCO	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTTGCGCC AAGTTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCT ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGC? C???GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C?????C??????????	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGCTG ATTA?GGCTG -7TC?GGCCA -7TC?GGCCA -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCA -TTC?GCCCA -TTC?GCCA -TTC?GCA -TTC?GCCA -	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777GGGCA 7777GGGCA 7777GGGCA 7777GGGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGCAT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCGTA	[ 12 1 1 1 1 1 1 1 1 1 1 1 1 1	2-145 2777777777 2777777777 27777777777 277777777	] ;? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022
1 2 3 4 5 6 7 8 9 10 11 2 3 4 5 6 7 8 9 10 11 2 3 4 5 6 7 8 9 10 11 2 3 4 15 6 7 8 9 3 10 11 2 3 4 15 6 7 8 9 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	ABABI LYCOP CUCUR triti LACTU SILEN vicia canel ophed ophed phetu PICEA PICEA MIUS PICEA MIUS PICEA MIUS MIUS MIUS MIUS MIUS MIUS MIUS MIUS	101 I CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC ATATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGC? C????GAGC? C????GAGC? C????GAGC? C????C????	TTCT?GGCCG ATTT?GGCCG ATCT?GGCCG ATCC?GGCTG -TTC?GGCTG ATTA?GGTG ATTA?GGTG -7TC?GGCCA -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCA -7TC?GGCCA -TTC?GCCCA -TTC?GCCA -TTC?GCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCA -TTC?GCA -TTC?GCA -TTC?GCCA -	2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGGGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777AGAGCA 2777GGGCA 2777GGAGCA 2777GGAGCA 2777GGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGA GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA AGCGTCCGGCT AGCGTCCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCATTT AGCGTCGTTA AGCGTCGTTA	[ 12 777777777 777777777 777777777 777777	2-145 ????????????????????????????????????	] ;? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022
$1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 11 \\ 11 \\ 11 \\ 11 \\ 11 $	ABABI LYCOP CUCUR triti LACTU SILEN vicia cANEL potam ephed qnetu PICEA MIUS PICEA MIUS MATSI CHIAM GONIU PANDO MIUSH MIUS SPERM TETRA CHLOR CLADO HETER SACCO BIPOL GLOMU CYANI SARCO PHYTO SCIVO	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC AATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC ACATTGCCCC ACCTCCCCC ACATTGCCCCCC ACATTGCCCC ACATTGCCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCC ACATTGCCCCCCCCCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGC? C???GAGCC C???GAGC? C???GAGC? C???GAGC? C???GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GAGC? C????GACGC? C????GACGC? C????GAC? C??????????	TTCT?GGCCG ATTT?GGCCG ATTT?GGCCG ATTC?GGCCG ATTC?GGCTG -TTC?GGCTG ATTA?GGTG -7TC?GGCCA -7TC?GGCCA -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCA -7TC?GGCCA -TTC?GCCCA -TTC?GCCA -TTC?GCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCCA -TTC?GCA -TTC?GCA -TTC?GCCA -TTC?GCCA -TTC?GCA -TTC?GCCA -TTC?GCA -TTC?	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAA GGCGTCGCAT AGCGTCGGCT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGTTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTTA AGCGTCGTCA AGCGTCGTATT AGCGTCGTCA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTCA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTA AGCGTCGTCA AGCGTCGTCA AGCGTCGTCA AGCGTCGTCA AGCGTCGTCG AGCGTCGCCG AGCGTCGCG AGCGTCGCG AGCGTCGCG AGCGTCGCG AGCGTCGCCG AGCGTCGCG AGCGTCGCG AGCGTCGCG AGCGTCGCG AGCGTCGCG AGCGTCGCG AGCGTCGCCG AGCGTCGCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTCC AGCGTC AGCGTCC AGCGTC	[ 12 777777777 777777777 777777777 777777	2-145 ????????????????????????????????????	] ;? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022
12345678910112134156789012222452678901323333356738	ABABI LYCOP CUCUR triti ILACTU SILEN vicia CANEL potam ephed PICEA MILEN PICEA MILEN MILEN CHLAM GONIU PANDO MIUEH MIURF SPERM TETRA CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CHLOR CCLADO HETER SCLER SACCH BIPOL GLOMU CYANI SARCO PHYTO SCYPt	101   CTTTGAACGC CTTTGAACGC CTTTGAACGC TTTTGAACGC TTTTGAACGC CTTTGAACGC	AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AAGTTGCGCC AATTTGCGCC AATTTGCGCC AATTGCGCC AATTGCGCC ATATTGCGCT ATATTGCGCT ATATTGCGCT ATATTGCGCT ACATTGCGCC	C???CAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAAGCC C???GAGGCC C???GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGGC? C????GAGC? C?????C??????????	TTCT?GGCCG ATTT?GGCCG ATTT?GGCCG ATTC?GGCCG ATTC?GGCTG -TTC?GGCTG ATTA?GGCTG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -7TC?GGCCG -TC?GGCCG -TC?GGCCA -TTC?GGCA -TTC?GCA -TTC?GGCA -TTC?GGCA -TTC?GGCA -TTC?GGCA -TTC?GGCA -TTC?GGCA	7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGGGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777AGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA 7777GGAGCA	CGTCTGCCTG CGTCTGCCTG CGCCTGCCTG CGCCTGCCT	170 GGTGTCACAG GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACGC GGCGTCACAGC GGCGTCGCAA GGCGTCGCAA GGCGTCGCAT GGCGTCGCAT GGCGTCGCAT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGGT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGGTT AGCGTCGTTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCGTTA AGCGTCCTTA AGCGTCCGTTA AGCGTCCGTTA AGCGTCGTTA AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCATTT AGCGTCCTTA AGCGTCCTTA AGCGTCCTTA AGCGTCCTTA AGCGTCCTTA AGCGTCCGTTA	[ 12 ????????????????????????????????????	2-145 2777777777 2777777777 2777777777 277777777	] ;? ?? ?? ?? ?? ?? ?? ?? ?? ??	e/article/13/9/1276/991469 by guest on 21 August 2022

a "?" Represents sites coded as missing; "-" are those positions coded as gaps.

<sup>b</sup> Abbreviations used and Genbank accession numbers are given in table 1.