Nuclear ribosomal spacer regions in plant phylogenetics : problems and prospects

Poczai, Péter

2010

submittedVersion

Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects

Péter Poczai · Jaakko Hyvönen

Abstract The nuclear ribosomal locus coding for the large subunit is represented in tandem arrays in the plant genome. These consecutive gene blocks, consisting of several regions, are widely applied in plant phylogenetics. The regions coding for the subunits of the rRNA have the lowest rate of evolution. Also the spacer regions like the internal transcribed spacers (ITS) and external transcribed spacers (ETS) are widely utilized in phylogenetics. The fact, that these regions are present in many copies in the plant genome is an advantage for laboratory practice but might be problem for phylogenetic analysis. Beside routine usage, the rDNA regions provide the great potential to study complex evolutionary mechanisms, such as reticulate events or array duplications. The understanding of these processes is based on the observation that the multiple copies of rDNA regions are homogenized through concerted evolution. This phenomenon results to paralogous copies, which can be misleading when incorporated in phylogenetic analyses. The fact that non-functional copies or pseudogenes can coexist with orthologues in a single individual certainly makes also the analysis difficult. This article summarizes the information about the structure and utility of the phylogenetically informative spacer regions of the rDNA, namely internal- and external transcribed spacer regions as well as the intergenic spacer (IGS).

P. Poczai (✉)
Department of Plant Sciences and Biotechnology, Georgikon
Faculty, University of Pannonia, Festetics 7, 8360 Keszthely,
Hungary
e-mail: guanine@ex1.georgikon.hu

J. Hyvönen
Plant Biology, University of Helsinki, P.O. Box 65,
FI-00014 Helsinki, Finland
e-mail: jaakko.hyvonen@helsinki.fi

Introduction

The ribosomal RNA (rRNA) genes and their spacer regions have become widely used as a source of phylogenetic information across the entire breadth of life [1]. The popularity of the rDNA locus for phylogenetics might be attributed to the phenomena that they serve the same function in all free-living organisms. They have the same or almost the same structure within a wide range of taxa. The coding regions, like the small- and large subunit gene, represent some of the most conservative sequences in eukaryotes [2, 3], which is a result of a strong selection against any loss-of-function mutation in components of the ribosome subunits [4]. The most conservative part appears to be the 3′ end of the 26S rDNA representing the α-sarcin/ricin (S/R) loop [5]. The information provided by the rDNA locus in phylogenetic research is significant, and it can be used at different taxonomic levels, since the specific regions of the rDNA loci are conserved differentially. The spacer regions of the rDNA locus possess information useful for plant systematics from species to generic level. They have also been used on studies of speciation and biogeography, due to the high sequence variability and divergence. There are three notable spacer regions: the external- and internal transcribed spacers (ETS, ITS) and the intergenic spacer (IGS). The general properties of these rDNA spacer regions will be reviewed in a phylogenetic context. Besides the general description, organization and

structure of each spacer, the recent advances made in the utilitization of each unit will also be discussed. Some of these are well summarized in other studies (like for ITS), while for ETS and IGS the relevant new findings have not been adequately reviewed. Thus, the aim of this study is to summarize the features of all rDNA spacer regions suitable for phylogenetic research.

The internal transcribed spacer as a phylogenetic marker

The internal transcribed spacer (ITS) is intercalated in the 16S-5.8S-26S region separating the elements of the rDNA locus (Fig. 1). The ITS region consists of three parts: the ITS1 and ITS2 and the highly conserved 5.8S rDNA exon located in between [6]. The total length of this region varies between 500 and 750 bp in angiosperms [7] while in other seed plants it can be much longer, up to 1,500–3,500 bp [8, 9]. Both spacers are incorporated into the mature ribosome, but undergo a specific cleavage during the maturation of the ribosomal RNAs [10–12]. It is now certain that ITS2 is sufficient for the formation of the large subunit (LSU) rRNA during the ribosome biogenesis [13]. The correct higher order structure of both spacers is important to direct endonucleolytic enzymes to proper cut sites [14]. Although, the sequence length of the ITS2 is highly variable between different organisms, Hadjiolova et al. [15] identified structurally homologous domains within mammals and *Saccharomyces cerevisiae*. In contrast to the coding regions, spacers evolve more quickly, like the internal transcribed spacer (ITS) region, which is extensively used as a marker for phylogenetic reconstruction at different levels. Since its first application by Porter and Collins [16] it has become widely used for phylogeny reconstruction. As a part of the transcriptional unit of rDNA, the ITS is present in virtually all organisms [11]. The advantages of this region are: (1) biparental inheritance, in comparison to the maternally inherited chloroplast and mitochondrial markers; (2) easy PCR amplification, with several universal primers available for a various kind of organisms; (3) multicopy structure; (4) moderate size allowing easy sequencing; and (5) based on published studies it shows variation at the level that makes it suitable for evolutionary studies at the species or generic level [7–9]. Álvarez and Wendel [17] and Baldwin et al. [7] summarize that this variability is due to frequently occurring nucleotide polymorphisms or to common insertions/deletions in the sequence. This high rate of divergence is also an important source to study population differentiation or phylogeography [18–21]. It has been widely utilized across the whole tree of life, including fungi [22–31], animals [32–36], different groups of 'algae' [14, 37–39] lichens, and bryophytes [40, 41]. In addition it is often used in the other two major domains of the tree of life Archaea and Bacteria [42–46] where RISSC, a novel database for ribosomal 16S–23S RNA genes and spacer regions is developed to provide easy access to information [47].

The high copy numbers allow for highly reproducible amplification and sequencing results, as well the potential
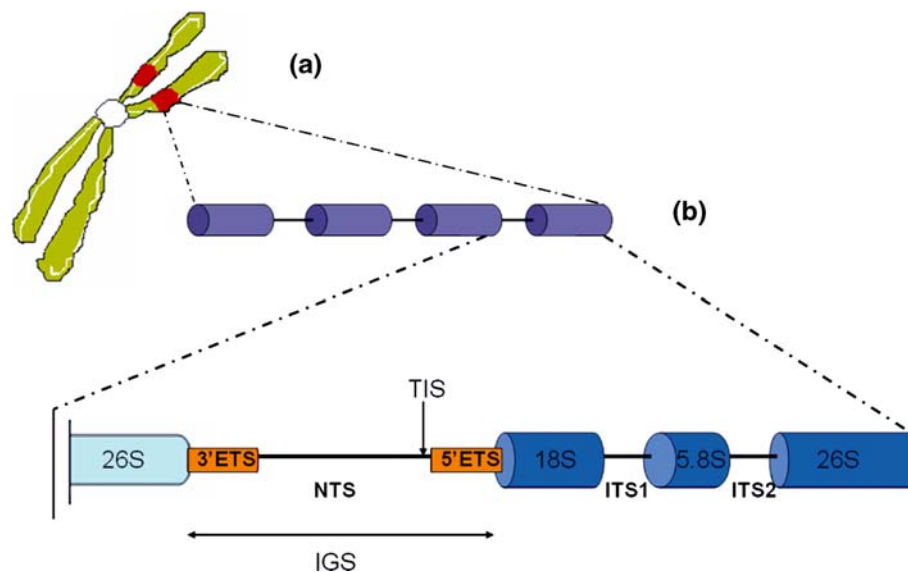
Fig. 1 Schematic presentation of the universal structure of the rDNA region in plants. (a) The chromosomal location of the rDNA regions. (b) Tandem arrays of the consecutive gene blocks (18S-5.8S-26S). In the tandem arrays each gene block is separated by an intergenic spacer (IGS) consisting of a 5′ end and 3′end external transcribed spacer (ETS). The two ETS regions are separated by a non-transcribed region (NTS). The transcription start site (TIS) labels the start position of the 5′ETS. The small subunit (18S) and large subunit genes (5.8S and 26S) are separated by the internal transcribed spacer 1 (ITS1) and internal transcribed spacer 2 (ITS2)

to study concerted and reticulate evolution. The number of studies utilizing ITS in phylogenetic studies is increasing, publicly available ITS sequences has tripled since 2003 [11]. The plant families most intensively studied are Asteraceae, Fabaceae, Orchideaceae, Poaceae, Brassicaceae, and Apiaceae. At the genus level there are for example more than 1,000 sequences available for different species of *Carex* (NCBI GenBank, nucleotide search preformed in 15.02.2009).

Besides several advantages there are many drawbacks for use of rDNA ITS data in evolutionary studies. There are hundreds or thousands of ITS copies in a typical plant genome [17]. Inferring phylogeny from multigene families like ITS can lead to erroneous results, because there is variation among the different repeats present in a single eukaryote genome [48]. Evidence now suggests that this variation among ITS sequences of an organism is found only within organisms that are hybrids or polyploids [49].

Multiple rDNA arrays and paralogy

Several ribosomal loci, both transcriptionally active and inactive, are usually present in plant genomes [50]. As ribosomes are the workhorses of the protein biosynthesis, translating mRNA to build polypeptide chains, they are extremely important structures in the cell. For this reason many copies are required to tend to the needs of an organism for this important process. These copies as well as their number and distribution in the plant genome are highly variable [51–56]. As both ITS regions are part of the cytoplasmic ribosome genes playing a role in the formation of the mature ribosome, there are hundreds, or in some cases thousands of tandem copies [57, 58]. Because of the high copy number this region is recognized as a multi-copy gene family, which provides easy amplification via PCR. This is an advantage, but on the other hand it can be a problem in phylogenetic analyses, if paralogous sequences are present. However, the general assumption for phylogenetic studies is that all ribosomal copies present within the genome have fairly identical sequences due to functional constraints. Orthologous genes and gene products found in different species are the basic requirement of phylogenetic inferences concerning common ancestry among species [59]. Unidentified paralogous relationships and infrequent recombination between paralogues can result in erroneous species phylogenies [60]. Paralogous sequences can occur at many levels: within an individual, among individuals within a species, and among species. To determine intra-individual paralogues among sequences of an individual and to find which are maintained and shared with other species is a potential problem in phylogenetic analysis. Another problem is PCR amplification, because

the ITS sequence amplified is a consensus of many targets sharing the same priming sites in one or several loci usually located in separate chromosomes. This consensus sequence used as a row of data in phylogenetic analysis is a molecular phenotype from which the genotype of the organism cannot always be inferred [50]. It is also impossible to determine the zygosity of the marker. There are two types of alternative copies which can be detected with PCR. First there are sequences having the same size as the others from different loci, but there are SNPs (single nucleotide polymorphisms) in different positions within their sequence. Sequences can differ also in size, because of permanent insertion/deletion events. Both types occur when different ITS repeats are merged within a single genome via hybridization (including allopolyploidy) or introgression. These processes are very common in plants; recent estimates suggest that 70% of all angiosperms have experienced one or more episodes of polyploidization [61].

Concerted evolution

In plants the ribosomal genes are present in several copies. For example in *Arabidopsis thaliana* more than 1,400 genes encode rRNAs, and occur on different chromosomes, with specific polymorphic alleles largely homogeneous in each rDNA array [62]. All copies within and among ribosomal loci are expected to be homogenized through genomic mechanism of turnover like gene conversion, the non-reciprocal transfer of genetic information between similar sequences, and unequal crossing over [63]. This phenomenon was first reported by Schlötterer and Tautz [64] and later by Polanco et al. [65] studying polymorphisms within the ITS in populations of *Drosophila*. They found that individual rDNA arrays are homogenized for different polymorphic alleles, which indicate that intra-chromosomal recombination events occur at rates much higher than those for recombination between homologous chromosomes at the rDNA locus. The intra-genomic rDNA diversity is generally low, and this low diversity results from concerted evolution within and between ribosomal loci [66]. The mechanism of concerted evolution completely, or almost completely, reduces the level of inter-repeat sequence variation between the multiple arrays of rDNA in every organism [17].

The fact that the ribosomal multigene family evolves through the process called concerted evolution certainly makes phylogenetic analysis much more difficult. It is important to recognize that concerted evolution is a complex process. According to various authors there are special stages during the process of concerted evolution, which lead to different classes in the plant genome [49, 67–69]. These stages can be important features in

phylogenetics, leading to questions: Are the several copies homogenized properly? Are there any heterogenic sequences? Which copy is the dominant sequence in the genome? Are there any variations between the sequences of an individual? However, concerted evolution does not act immediately after organismal processes such as hybridization or polyploidization, or after genomic changes like gene and chromosome segment duplication, and various forms of homologous and non-homologous recombination [11]. Thus, divergent rDNA copies could be present throughout the genome, disturbing phylogenetic analysis and sequencing. Because paralogous copies occur due to polyploidy or hybridization they can be utilized to study these processes. The presence of parental rDNA repeat types in a hybrid is determined by many forces affecting their molecular evolution [70]. The detection of these alternative copies depends on their number. If hybrids are recent, both parental types are almost always present [71]. Such hybridization can be easily reveled by direct sequencing, where an additive pattern of sequence variation is present. In such cases, the sites differing between species yield signals from two different nucleotides. According to Rauscher et al. [70] it is unclear how common a repeat type must be, relative to the other parental type. In the case of *Gossypium* spp. the homogenization process was complete, leaving no easily traceable evidence in the ITS region to track polyploidy [69]. However, in the *Glycine tomentella* complex the ITS region was successfully used to evaluate parental relationships and hybrid speciation [70]. In this study repeat-specific and exclusion PCR primers were designed to detect rare parental ITS types. In another study Koch et al. [72] clarified the multiple hybrid origin of natural populations of *Arabis divaricarpa*, the putative hybrid of *A. holboellii* and *A. drummondii*. They detected multiple intraindividual ITS copies in several *A. divaricarpa* accessions which were also present in the parental species. But concerted evolution in this case also resulted in different ITS types, in the hybrid *A. divaricarpa* and in the parental taxa, respectively. In other groups like *Potamogeton* [73], *Bromus* [74], *Nymphaea* [75], *Armeria* [76] and *Cardamine* [77] ITS was a valuable source to reveal complex reticulate events between putative hybrids [78–81]. Concerted evolution is sometimes incomplete and some copies of the tandem arrays became non-functional pseudogenes [50]. Mayol and Rosselló [82] reanalyzed datasets by two different and independent laboratory teams [83, 84] generated for the study of systematics of the genus *Quercus*. Their surprising result was that the divergent ITS alleles reported by one of the teams were non-functional paralogous copies (pseudogenes). It was also concluded that the incorporation of these ITS paralogues in evolutionary studies can lead to

erroneous hypotheses about phylogeny. Standard definitions of pseudogenes are hard to be apply to rDNA pseudogenes. In the context of phylogenetic reconstruction Bailey et al. [85] determined rDNA pseudogenes as sequences with nucleotide divergence pattern that has not been constrained by function irrespective of expression patterns.

Secondary structure modeling of the ITS region

The construction of the secondary structure model of the ITS RNA transcript was proposed as a novel tool for phylogenetics. These new methods have also made analyses easier in a user-friendly interface (e.g., online databases and programs). The importance of this recent advance enable inference of phylogenies not only based on sequence information, but also based on predicted secondary structures. The phenomena that rRNA single-stranded chains form a secondary structure which contain stemmed regions and different loops correlating with base pairing opened a new field to infer phylogenies. During phylogenetic analysis it is hard to determine whether a pseudogene or a paralogous sequence has interfered the results. ITS2 is a well suited marker with a broad use in low level phylogenetic analyses, as its sequence evolves quite fast. This feature, which made the region useful for analyses at generic and infrageneric level, is a 'hindrance' for the application of this marker for more general phylogenetic analyses [86]. The possibility to predict the folding structure has enhanced the role of ITS in phylogenetic studies, since this will enhance sequence alignment which can be based on secondary structures [87]. When comparing the structure of the ITS2 RNA transcript, it turned out that a conserved core is found in different species. Many methods have been applied to infer the secondary structure of the ITS2, like electron microscopy [88], chemical and structure probing [89], and site-directed mutagenesis [90, 91] Also different softwares have been developed for this purpose [86]. The surprising result of these studies has been that the examined eukaryote groups share the same general ITS2 secondary structure [92]. It was concluded that the secondary structure for the ITS2 consists of four helixes. Among plants, nucleotide sequence evolves most rapidly in region IV followed by helix I [48]. It was also described that 'helix II is more stabile, and characteristically has a pyrimidine-pyrimidine bulge while helix III contains on the 5′ side the single most conserved primary sequence, a region of approx. 20 bp encompassing the TGGT' [48]. The fundamental role of the helicoidal ring of ITS2 during the pre-RNA processes is to trigger the maturation of the 26S rRNA, because it was observed that the lack of its structure blocks the productions of the mature large

subunit, when specific structures of the ring model can not be formed [93–95]. These features led Schultz et al. [86] to construct an ITS2 database, a web-based tool for phylogenetic analysis. This new server is open for structure prediction and provides a way for utilizing more information from sequences. Recent studies have used a parsimony approach to predict the appropriate structure for ITS2 according to free energy measurements. The method developed by Schultz et al. [86] to predict ITS2 structures is based on the Needleman–Wunsch algorithm [96], but applies a BLAST search with the newly predicted structure in the database to compare it with others [86, 97, 98]. The phylogenetic estimates made from secondary structure characters are based on newly generated information that differs from the sequence level data.

When comparing the secondary structures of the ITS2 it turned out that it is also useful to distinguish species based on compensatory base changes (CBCs). CBCs occur in a paired region of a primary RNA transcript when both nucleotides of a paired site mutate, while the pairing itself is maintained [99]. In other words following the definition of Kimura [100] compensatory mutations are a pair of mutations at different loci (or nucleotide sites) that are individually deleterious but are neutral in appropriate combinations. Basically the RNA secondary structures comprises single-stranded and double-stranded regions, where the double-stranded stems are formed by Watson–Crick (WC) pairing of complementary bases.

In all eukaryote groups where a broad array of species has been compared for ITS2 sequence secondary structure and tested for any vestige of species sexual compatibility, an interesting correlation has been found: 'when sufficient evolutionary distance has accumulated to produce even one CBC in relatively conserved pairing positions of the ITS2 transcript secondary structure, taxa differing by the CBC are observed experimentally to be totally incapable intercrossing' [101]. Using the ITS2 database now consisting of 65,000 ITS2 sequences Müller et al. [99] concluded that CBCs in an ITS2 secondary structure are sufficient indicators to distinguish even closely related species. Secondary structures can be determined by an alignment against sequences with already known structures to depict common base-pairing patterns [102]. Secondary structures are particularly useful because they include information not found in the primary sequence [4]. The modeled structures are of importance in phylogenetic analysis, because they can be used to enhance alignments obtained by different methods [102, 103].

The secondary structure of ITS1

Structural prediction of the ITS2 is more common in phylogenetic studies than the rarely used ITS1, but several models exist also for ITS1. The successful prediction of the secondary structures of ITS1 and to compare these sequences among genera or families is more difficult than in the case of ITS2. ITS1 seems to evolve faster, and has less conserved sites than ITS2. As discussed above the prediction for ITS2 is easier since there are conserved motifs in the structure across a wide range of divergent lineages, like the four helixes or the 5′ UGGU motif and many others [86]. This is unfortunately not the case in ITS1. Liu and Schardl [104] identified a 20 bp region in the inferred secondary structure of ITS1, which is highly conserved among angiosperms, and this was also confirmed by Goertzen et al. [105] but there are no reports of other conserved motifs within the structure of ITS1. However, there are some studies that have utilized ITS1 structure prediction in phylogenetic analyses [106–108]. The reconstruction of the ITS1 RNA transcript secondary structure model is not always successful between different plant genera. The failure in some cases might be related to the features of ITS1 which are different from ITS2. The ITS1 is more variable than the ITS2. The two spacers evolve with different substitution rates, and the rates seem to vary from species to species [109]. Thus, a universal rate could not be applied to plants; it must be edited and determined for each genera. This difference between the two spacers is due to their function and role in the ribosome biogenesis. The sequence of the ITS2 is more conserved because it is more important in the formation of the mature ribosome than the ITS1. Thus, the higher order and secondary structure is more important for the appropriate ribosome formation, which is reflected in its sequence. It must be recognized that there is more conformational similarity between the aligned and predicted structures of ITS2 than in the ITS1. Baldwin et al. [7] also concluded that in contrast to ITS1, ITS2 displays more sequence similarity across plant families in the central region of the spacer. The conformational similarities in the higher order in the predicted structures of the RNA transcript might be attributed to stronger functional constrains of the ITS2. Baldwin et al. [7] also tried to predict the structure of ITS1 for five plant families, but the sequences showed insufficient retention of similarity during the alignment. The final conclusion was that the ITS1 lacks similarity in the structure across plant families, but there were minor differences in the free energy between the most parsimonious structures. It would be highly interesting to predict the ITS1 structures with the novel approach of Schultz et al. [86]. A tabulation based on the available sequences in the Genbank across all ITS1 sequences would give new insights to the theoretical and practical application of the secondary structure of the ITS1. With a proper algorithm it would be possible to predict the potential secondary structure models from the available sequences. Comparing them with each

other might elucidate the utility of ITS1 secondary structure models in phylogenetic analyses. In summary a more extensive study on the ITS1 structure should provide new data and information for evolutionary studies.

The intergenic spacer region

The IGS region, separating the rDNA tandem arrays (Fig. 1), consists of regulatory elements like promoters, enhancers and terminators, and the NTS (Non-transcribed spacer) as well as the external transcribed regions (discussed in the next section). For the last several years this region have been used for evolutionary biology and genetics rather than studies of phylogeny. The major findings of recent studies are the exact and descriptive structure of the rDNA IGS for many species. This region contains several kinds of repeating elements, also refered as subrepeats, various types of enchancers, and promoter regions which are often distributed throughout the whole region in duplicated forms. There are elements which form conserved secondary structures. The full detailed exposition of the structural elements of the IGS region is beyond the phylogenetic context of this review. For further details in this context see Weider et al. [110] and Gorokhova et al. [111]. In general it has to be noticed that these elements play an important and essential role in the control of the rRNA transcription and also in the processing of the transcript during the replication of the unit. Thus, IGS is an important functional region, because it contains the nucleotide sequences that trigger and/or terminate transcription [112–116]. The phylogenetic interpretation and utility of this region is naturally constrained by these features. Before the full detailed discussion of the IGS region some nomenclatural clarification must be made in order to refer precisely to each region. There are two non-transcribed (NTS) regions in plants. The 5S repeats in plants are located in different locus or loci than the nucleolar organizer region (NOR, 18S-5.8S-26S rDNA) locus or loci, although the 5S and NOR can also be located in the same chromosome [117]. This organization is unique in plants in contrast to other eukaryotes, where the 5S region repeats are intercalated in the IGS region. The 5S units of the rDNA are also present in repetitive tandem arrays and are separated by simple NTS regions frequently referred as the 5S NTS which are not equal to the regions found between the separately organized NOR arrays. Predominantly they have the same features and structure, but the term NTS in this review will attribute only to the spacer region found between the NOR tracts of rDNA locus, as the 5S NTS will not be discussed here.

High sequence variability

As some parts of the IGS are also variable or more variable than the widely utilized ITS, they have also been used as phylogenetic markers. The utility of IGS for phylogenetic studies is criticized by its major features, which are: (1) high sequence variability; (2) subrepeat tracks present in the region; and (3) the length of these subreapeats are highly variable disrupting the sequence alignment between taxa. It must be realized, that the IGS is a rapidly evolving region of the rDNA with several internal subrepeats present in its sequence, which evolve rapidly both in size and structure making comparative anlayses difficult in some cases. The length polymorphisms have been observed between populations, species or even in individuals [116]. The region seems to be extremely dynamic [118, 119]. The primer design for this region can be problematic too, because the rDNA IGS is known for gradual decrease in sequence conservation upstream from the 18S gene to the center of the rDNA IGS which consists of repetitive elements [120, 121]. These troublesome properties of the rDNA IGS region cause serious obstacles to the development of primers and in the alignability apart from the ETS, even at low taxonomic levels [11]. Most plant species show length variation in the IGS [116, 117]. IGS has not proven itself as a useful tool for phylogenies of species that are not very closely related, not only because the IGS has a large number of related subrepeats but also because the subrepeat length and primary sequences are too dissimilar to be aligned [120]. Several length variants of the IGS can be present in the same genome [122]. Species of the same genus often differ significantly in the length of the IGS, which can range from 1.7 to 6.4 kb in the case of *Trillium* [123]. Even in the case of individuals of the same population, the length of the IGS can differ, but there are examples for low variability or even uniformity among species of the same genus [124]. The upstream of the TIS accumulates base substitution with a high rate: obvious sequence similarity can be found here only among members of the same genus or closely related genera [5]. As summarized above there are several factors which affect the use of IGS as a potential phylogenetic marker. Amplification of subrepeats in the IGS have occurred several times during evolution, as it is known from the examples of the genus *Nicotiana*, where original duplication of the ancestral A-subrepeat sequence in the ETS took place before the divergence of the subgenera, and produced two subvariants, A1 and A2 [125]. Longer stretches of the A1/A2-subrepeats are formed independently during the later speciation of *Nicotiana*, while similarly several rounds of amplification/delations generated C-subrepeats in the upstream of the TIS [126, 127]. As it is seen in the example of *Nicotiana* this "chaotic-mix" of subrepeats

prohibits the routine use of easy amplification with PCR and the development of universal primers across plant families.

The phylogenetic utility of IGS subrepeats

Despite of the fact that there are many reiterated subrepeats present within the sequence of the IGS it has been successfully used to infer phylogenies. These studies were mainly carried out among closely related species. When closely related species are compared with each other, the possibility that the IGS region could be aligned properly greatly increases. Maughan et al. [128] analyzed the IGS region of different cultivars of Andean grain crop quinoa (Chenopodium quinoa) and a related ancestor C. berlandieri subsp. zschackei. During this analysis it turned out that the IGS regions of quinoa and its wild relative have strikingly similar subrepeat sequences which differ in their number and in the presence of species-specific motifs toward the 3′ end of the IGS. Sequence comparison indicated that the two allotetraploid species descended from at least one common diploid ancestor. In another study Fernández et al. [119] compared the IGS sequence of Lens culinaris with other species of Vicieae and Phaseoleae. The amplified spacer was also composed of nonrepetitive sequences and four tandem arrays of repeated sequences. According to the number and length of subrepeats, different repeat types were identified named A to D. Among the sequences conserved motifs were also found, which were attributed as functional sequences. The most noticeable result of this study is that despite of the rapid evolution of the IGS sequences within and between the two legume tribes, some motifs have been conserved in their sequence and relative position. Some of these motifs were found in other phylogenetically distant taxa. Although, there are some exceptions where the IGS can be successfully applied as a phylogenetic marker this can not done universally to all plant species. Additionally, IGS sequences could only be employed as phylogenetic markers when closely related species are the subjects of interest [118, 129–132]. Subrepeats appear in the IGS region of all plant species and in eukaryotes, except for one unambiguous case of Caenorabditis elegans, which has a simple organized structure for an unknown reason. Because of these subrepeats present in the well characterized region of IGS in plants, the length of this region varies dramatically. But is there any consequence in the duplication of these regions?

The repeats can be grouped into different classes in Oryza sativa [133], Vicia faba [134], Triticum aestivum [135] and several other species. Later, this has been confirmed also for Olea europea [136] and Quercus [137]. If we apply the widely accepted model of Dover and Tautz [138] to subrepeats; that is DNA sequences evolve through successive cycles of tandem duplications and perhaps, the divergence of an ancestral sequence, interesting conclusions can be made. Ryu et al. [139] presented that the evolution of the IGS can be thought to include duplications/delations through divergence processes, resulting in a dynamic change in the subrepeat composition. These duplications, often called as homopolymeric runs or mononucleotide microsatellites can be abbreviated with the term poly (N). Ryu et al. [139] concluded that if the repeats within the IGS are interpreted as poly (N) runs the nature of this fraction could be better understood. They characterized the IGS region of several species and developed an alignment algorithm, which can take into account the differences induced by the poly (N) runs and recover the underlying phylogenetic signals from the IGS subrepeat comparisons. The method is called dropout alignment. Thus, prior to the alignment of the sequences all consecutive bases in each poly (N) run are deleted (dropped out) except one base, because many difficulties arose from the differences in the expansion of the poly (N) repeat types. This new method allowed the proper alignment of different types of subrepeats within the same taxa, and enabled alignment and comparison of the subrepeats between different species. The method of Ryu et al. [139] suggests that most of the variation found within the IGS is manifested by the occurrence of poly (N) runs. The new method could be a useful tool for the further evolutionary investigation of the IGS and opens the opportunity for the better exploitation of IGS sequences for phylogenetic purposes.

The external transcribed spacer

The external transcribed spacer (ETS) lies in the intergenic spacer region separating the repetitive 18S-5.8S-26S ribosomal gene blocks from each other. There are two ETS sites: the 3′- and 5′ prime parts which are bordering the 18S and 26S exons (Fig. 1). There is a substantial difference between the 5′- and 3′ prime ETS. In some studies the 5′ ETS is also referred as ETS1, while the 3′ external transcribed spacer as ETS2. But what are the 3′- and 5′ prime ETS regions? There is a considerable confusion around the nomenclature of the ETS region. In the literature both spacers are named differently and the denomination of 5′- and 3′ parts are sometimes confusing depending on what were the bases of the nomenclature. In the first model the 18S-5.8S-26S is treated as a separate unit and the bordering sequences in both ends are considered to be the external transcribed spacers. Thus, when the rDNA locus is treated as a separate unit, the external transcribed spacer from the 18S part is labled as 5′ETS and the 3′ spacer region of the 26S exon is called as 3′ ETS (e.g., Hershkovitz et al. [140]). In the second model the IGS is taken

as the base labeling the ETS regions. The 5′ prime part of the IGS would then be the 5′ ETS and likewise the 3′ part of the IGS should be named as 3′ ETS (e.g., Calonje et al. [11]). In this review the terms 3′- and 5′ETS will be used to refer exactly and directly to the two parts of the rDNA region instead of ETS1 and ETS2. Because the tandem repeat sections of the rDNA locus begin with the 18S gene in order with the main direction of the DNA strand, the site upstream from the 18S exon, should be unambiguously labeled as the 5′ETS (5′end ETS or ETS1) and the site downstream from the 26S exon (the 3′end of the 26S gene) likewise 3′ETS (3′end ETS or ETS2). In this review nomenclatural terms will be treated as stated above. Another confusion surrounding the ETS is whether the 3′ETS exist or not? In some studies the 5′end part of the IGS is referred to as the 3′ETS. Since the transcription of this region was not always obvious, as well as the exact position of the transcription termination site, the nomenclatural term 3′ETS is not applied by every study. In summary it has been shown that the 5′end of the IGS contains non-repetitive sequences which are highly similar across different species [118, 130, 141, 142]. Recent studies have shown that this region is transcribed and plays a role in the ribosome transcription [12, 143, 144]. The transcription termination site in the 3′ETS—as the transcription initiation site in the 5′ETS—is highly variable in plants. In the recent years a great progress was made surrounding the external transcribed spacers, revealing interesting new features about the region. As this mainly concerns the 5′ETS; and much more data is available for this region since it is more widely utilized in phylogenetic studies, this review will mainly focus on the 5′ETS. Consequently, in the following the major characteristics of this region in context of the new findings will be discussed to provide a conclusive description about their potential use in evolutionary studies.

The phylogenetic signal of the ETS

Since its first application by Baldwin and Markos [145] several analyses successfully adopted this marker as a valuable phylogenetic tool. Sequence comparisons of the rDNA external transcribed spacer (ETS) indicated that it represents an even more valuable instrument for the phylogenetic analysis than ITS [130]. The ETS has been used in phylogenetic analysis of families Asteraceae [145, 146], Fabaceae [147] and Myrtaceae [148] only to mention some examples. The 5′ETS is more frequently used in phylogenetic studies, than the 3′ part. The length of the 5′end ETS range from 425 to 575 bp [149–152] making it easily sequenced. There are less sequences available for ETS compared to ITS. This might be attributed for the ambiguous amplification and primer design for this region.

Several studies utilized this marker prior to the prominent article of Baldwin and Markos [145] and Bena et al. [153]. The information about the molecular features of the ETS region has increased recently. According to a fast tabulation among submitted sequences in GenBank (search preformed in 04.06.2009 in NCBI, GenBank). A large portion of sequences is available mainly for crop plants.

The protocol provided by Baldwin and Markos [145] with the study of Bena et al. [153] for primer design made it easy to exploit the ETS for phylogenetics. The major principle of this method is the amplification of the total IGS region with primers starting from the flanking region of the 18S and 26S genes. The procedure requires a long-PCR protocol; because the total length of the IGS is classically longer then 4 kb [154] and can be up to 12 kb [58]. The next step is to design taxon specific internal primers for the conserved regions of the amplified intergenic spacer according to the nucleotide data of the reverse sequenced product. Problems may arise when duplications of promoter regions intercalate in the region. An ideal point would be to design primers to the conserved promoter motive flanking the RNA polymerase I transcription initiation site, but the position of this region in the 5′ETS is often variable in plants. Finding the [TATA(G)TA] motif with reverse sequencing may require additional time as compared to the routine use of the well developed and widely applied ITS primers. In those cases where additional work has been done in this context it is easy to utilize the ETS region. In plant lineages where such preliminary data is lacking it would require much additional work in the primer design. But is this additional laboratory work warranted? Does the ETS region contain enough information to be used in phylogenetic studies? These questions could be answered according to the several studies that have been published [155–159] making primers available for the scientific community. The major observation is that the ETS evolves at faster rate compared to ITS. However, the restriction site studies by Kim and Mabry [160] have found that variation within ETS is comparable to ITS. The mean rate was estimated as 2.86×10^{-9} subs/site/year according to the study of Kay et al. [109] which was based on 29 independent ITS substitution rates ranging from 0.38×10^{-9} (Hamamelis) to 8.34×10^{-9} (Soldanella). These kinds of estimates are not available for the ETS. According to Bena et al. [153] the 5′ETS evolves 1.5× faster than the internal transcribed spacer region. In another study Baldwin and Markos [145] calculated a 1.3–2.4× higher rate of sequence evolution, while Linder et al. [146] estimated 7× higher rate in Asteraceae and closely allied families. 5′ETS provided approximately 1.5× higher information content in terms of parsimony informative sites. A possible explanation for this difference in the substitution

rates would be the reduction of ribosomal maturation processes in the case of ETS.

Divergent repeat types of ETS

The homogenization process of concerted evolution is the operating force to eliminate the different repeat types of ETS found within the genome of a single individual. However, concerted evolution is a well known and specific feature of multigene families—such as the rDNA locus—the rate and even its accuracy is not well known. In general the whole process of concerted evolution enhances the sequence similarity between multiply arrays of ITS and ETS. But does concerted evolution operate at different rates between each region? In a theoretical model ETS also evolves in concert and therefore its utility is affected by the same 'drawbacks' as ITS. The inclusion of paralogous sequences in the analysis would also be a problem as well as the appearance of pseudogenes in the datasets which can greatly disturb the reconstruction of phylogenies if they are accepted as orthologs. In addition, the putative repeat types of the ETS would also provide data to study complex evolutionary processes as well as to reveal hybridization no matter if they have arisen in recent or ancient hybridization. Another approach would be to combine the results of the alternative copies of ITS and ETS, thus further information would be available for tracking polyploidy. However, concerted evolution does not seem to operate at the same level between ITS and ETS. High level of intragenomic similarity has been found between 5′ETS sequences in *Helianthus* [146]. It was concluded that concerted evolution does not always eliminate all intragenomic variation in ETSs of all rDNA repeats, but it proceeds rapidly enough to not obscure specific relationships. Vander Stappen et al. [161] found no evidence for the presence of multiple ETS sequence types—or in a previous analysis for multiple ITS types—within individuals, indicating that concerted evolution acted affectively in both regions in the allotetraploid species of *Stylosanthes* [162]. Furthermore, it was also reported that the homogenization covers all parental rDNA repeat types. It seems that ETS homogeneity is maintained within rDNA clusters and throughout genomes [11]. However, it was reported that several subrepeat types exist within different taxonomic groups. These subrepeats, which are repeated regions in the ETS sequence, are found in *Solanum* sect. *Petota* and also in many other taxa like *Arnica mollis* and *Hemizonia perennis* [145]. Volkov et al. [163] reports A, B and C variants of 5′ETS subrepeats within *Solanum*. Their results showed that during the evolution of sect. *Petota* at least two large rearrangements of ETS occurred, resulting in B and C structural variants. These variants now succesfully has been used as sources of phylogenetic information for potatoes and the major taxonomic groups can be separated based on this information. An interesting feature is that these repeat groups evolved through stepwise base substitutions allowing the additional discrimination of closely related species. The latter structural study of the 5′ETS also ambiguously supported the position of 'Lycopersicon' (tomato clade) in sect. *Petota* through the discovery of a new D subrepeat, among the members of the tomato clade [164]. Another interesting feature of these subrepeats is that variant D might have originated directly from the ancestral variant A found in ser. *Etuberosa*. Taken together, a high rate and accuracy of concerted evolution seem to operate in the 5′ETS region and the sequences of the ETS region can be successfully used in the reconstruction of phylogeny at different taxonomic levels and can be widely applied in plant phylogenetic research.

Combining ETS and ITS data

In several cases high genetic diversity have been reported in plants. This large diversity is attributed to reticulate evolution, hybrid speciation or polyploidy. Such events are common in angiosperms. To address changes or evolutionary patterns and/or complex events at the molecular level markers with robust phylogenetic signal are desirable for phylogenetic analyses. The ETS region can be used successfully in phylogenetic studies where ITS seems to have only a weak signal, such as in recently diverged lineages, because it shares the same favorable features of the ITS, and it is generally known to evolve faster and to contain more phylogenetically informative characters than the ITS in plants [145, 155, 161]. Since both ITS and ETS are part of the rDNA locus their application in a combined analysis seems obvious. Although, when the ITS and ETS regions are analyzed separately the results about the phylogenetic signal provided by each region is variable. This variability might depend on phylogenetic histories and might be displayed differently in various taxonomic levels. It has been found that the 5′ETS includes less parsimony informative sites than ITS in resolving higher level relationships. Oh and Potter [165] combined the data of ITS and 5′ETS in the study of tribe Neillieae (Rosaceae). They also report that 5′ETS region included less parsimony informative sites than the ITS. Markos and Baldwin [166] analyzed the ETS and ITS data separately and in combination to study the higher-level relationships and major lineages of *Lessingia*. In their analysis it turned out that the amplified ETS region possessed more variable sites and it was 1.4× more informative than the ITS site. In another case [167] the 5′ETS did not allow reliable alignment of the compared sequences. Tucci et al. [167] report that in *Cynara* and *Onopordon* the sequences showed only 27% similarity.

Clearly there are differences in the diversity of ETS and ITS as reported by various studies. In some case the ITS is more informative at the higher taxonomic levels and the ETS can be less informative or vice versa. Possibly the ETS can accumulate great diversity or repetitive elements (like promoters) in its sequence and it cannot be aligned properly even in closely related genera. In studies where both ETS and ITS data have been used, no matter if the phylogenetic signal of the ITS or of the ETS was stronger, the combined analysis of the datasets have resulted in better resolved and more robust trees. Bena et al. [153] also reports that combining the two regions, greatly improves the resolution and increases the bootstrap support values for clades. It can be concluded that ETS data, besides the well utilized ITS, should be included to improve the estimation of phylogenies of different plant groups.

Concluding remarks

It can be summarized that the spacer regions of the rDNA locus are useful phylogenetic markers. They share the small size, high sequence level variability, conserved flanking regions, and rapid concerted evolution under similar functional constraints. In all eukaryote groups phylogenetic relationships can be inferred from rDNA markers, because different parts of this region evolve with different rates. However, rDNA regions coding for the large- and small subunits of the rRNA display relatively little variation, and thus they will remain the major targets of studies inferring phylogeny at the higher taxonomic levels. Other parts of these valuable loci can also be easily used in plant systematics. On the other hand doubts about the correct use of ITS and ETS data are considerable. These sequences can be difficult to handle if polyploidization or other events disturb the phylogenetic signal. The detailed nature of these regions should be taken into account. Paralogy, if overlooked and mistakenly incorporated in the in the phylogenetic data, can be a problem and speaks against the utilization of ITS. However, with considerable investment in amplification and analysis pseudogenes and paralogs can be isolated. If they include enough information, or they are the targets of interest they can be included in phylogenetic analyses.

As the nature of the ITS region is well understood new advances in its utilization (e.g., RNA transcript secondary structure prediction based reconstructions) are welcome additions to be used for phylogeny reconstruction. As demonstrated by several studies and summarized here, besides the routine use of ITS, other variable regions of rDNA could provide new information about phylogeny. Besides the nucleotide sequences of the ITS the recently developed database of the predicted secondary structures can be utilized.

Although, concerted evolution and the repetitive nature of the ITS could prevent its routine usage, it still might have great potential to study more complex evolutionary relationships. The process of concerted evolution is a fundamental phenomenon operating in all eukaryote organisms. The incorporation and application of new developed protocols and methods to study the divergence among and within repeat types of multigene families, such as the rDNA locus, is a developing area providing new data about phylogeny in both higher and lower level evolutionary studies. Besides the routine use of the nuclear ribosomal spacer regions (ITS, ETS or IGS) searching for alternative repeat type sequences which have escaped the homogenization process of concerted evolution is a field that deserves much further attention in future.

The internal transcribed spacer regions are widely employed markers and more attention should be focused on external transcribed spacers. Based on previous studies these regions seem to include enough information to warrant their use as phylogenetic markers. The fact that these regions are more variable in length or in their sequence composition makes them underutilized, because the routine amplification with available universal PCR primers is not always successful. ETS sequences can be used instead of ITS sequences or in combination when the ITS provides relatively weak phylogenetic signal. In the case of recently evolved lineages it is recommended to use ITS and ETS data in combination, because they share the same features and it has been shown repeatedly [168–171] that combining datasets in simultaneous analysis provides robust information. As pointed out by Wheeler et al. [172] simultaneous analyses hold the key for the use of 'problematic' areas in phylogenetic analyses. If a true, divergent phylogeny exists the signal should get stronger when more characters and terminals are added to the analysis. On the other hand if clear pattern does *not* emerge this might indicate a genuine lack of divergent lineages and thus need for other approaches such as the use of methods capable to handle reticulations [173–176].

Álvarez and Wendel [17] urge the routine utilization of single-copy nuclear genes as an alternative to the use of ITS sequences, because these single-copy sequences become more easily accessible via whole genome databases. Nowadays they can be routinely applied to the same purposes as rDNA markers. Single-copy genes do not suffer from the 'drawbacks' of concerted evolution, paralogy and homoplasy. The modeling of reticulate evolution or horizontal gene transfer is a developing area in evolutionary genetics. In order to visualize these scenarios several new algorithms are being developed, e.g., phylogenetic networks [177–180]. As reticulating events are

common in plants there is a great need of nuclear gene markers capable to reveal these events. The rDNA spacer regions could have potential to be used because of their universality and simplicity. The complex nature of the rDNA locus in plants could also be a useful feature. Consequently, plant phylogenetic studies should supplement the routine use of these markers and combine this data with recently developed new methodologies available for these regions. Combining the data of the ITS and ETS with the appropriately predicted secondary structures by utilizing newly developed algorithms in large plant phylogenetic super-networks will in many cases most likely provide robust phylogenetic signal.

References

1. Dixon MT, Hillis DM (1993) Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. Mol Biol Evol 10(1):256–267
2. Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. Q Rev Biol 66:411–453
3. Lipscomb D, Farris JS, Källersjö M, Tehler A (1998) Support, ribosomal sequences and the phylogeny of the Eukaryotes. Cladistics 14(4):303–338
4. Caetano-Anollés G (2002) Tracing the evolution of RNA structures in ribosomes. Nucleic Acid Res 30:2575–2587
5. Volkov RA, Medina FJ, Zentgraf U, Hemleben V (2004) Molecular cell biology: organization and molecular evolution of rDNA, nucleolar dominance, and nucleus structure. In: Esser K, Luttge Darmstadt U, Lüttge U, Beyschlag W (eds) Progress in botany. Springer, Berlin, pp 108–109
6. Wheeler WC, Honeycutt RL (1988) Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. Mol Biol Evol 5:90–96
7. Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann Missouri Bot Gard 82:247–277
8. Maggini F, Marrocco R, Gelati TM, De Dominicis RI (1998) Length and nucleotide sequences of the internal spacers of nuclear ribosomal DNA in gymnosperms and pteridophytes. Plant Syst Evol 213(3–4):199–205
9. Liston A, Robinson WA, Oliphant JM, Alvarez-Buylla ER (1996) Length variation in the nuclear ribosomal DNA internal transcribed region of non-flowering seed plants. Syst Bot 21(2):109–120
10. Hadjiolova KV, Georgiev OI, Nosikov VV, Hadjiolov AA (1984) Localization and structure of endonuclease cleavage sites involved in the processing of the rat 32S precursor to ribosomal RNA. Biochem J 220:105–116
11. Calonje M, Martin-Bravo S, Dobeš C, Gong W, Jordon-Thaden I, Kiefer C, Kiefer M, Paule J, Schmickl R, Koch MA (2008) Non-coding nuclear DNA markers in phylogenetic reconstruction. Plant Syst Evol. doi:10.1007/s00606-008-0031-1
12. Houseley J, Kotivic K, Hage AE, Tollervey D (2007) Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. EMBO J 26:4996–5006
13. Baldridge GD, Dalton MW, Fallon AM (1992) Is higher-order structure conserved in eukaryotic ribosomal DNA intergenic spacers. J Mol Evol 35:514–523
14. Mai JC, Coleman AW (1997) The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J Mol Evol 44:258–271
15. Hadjiolova KV, Normann A, Cavaillé J, Soupéne E, Mazan S, Hadjiolov AA, Bachellerie JP (1994) Processing of truncated mouse or human rRNA transcribed from ribosomal minigenes transfected into mouse cells. Mol Cell Biol 14:4044–4056
16. Porter CH, Collins FH (1991) Species-diagnostic differences in the ribosomal DNA internal transcribed spacer from the sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera:Culicidae). Am J Trop Med Hyg 45:271–279
17. Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. Mol Phylogenet Evol 29:417–434
18. Vargas P, McAllister HA, Morton C, Jury SL, Wilkinson MJ (1998) Polyploid speciation in *Hedera* (Araliaceae): phylogenetic and biogeographic insights based on chromosome counts and ITS sequences. Plant Syst Evol 219:165–179
19. Vargas P, Morton C, Jury SL (1999) Biogeographic patterns in Mediterranean and Macaronesian species of *Saxifraga* (Saxifragaceae) inferred from phylogenetic analysis using ITS sequences. Am J Bot 86:724–734
20. Baldwin BG (1993) Molecular phylogenetics of *Calycadenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. Am J Bot 80:222–238
21. Yamaji H, Fukuda T, Yokoyama J, Pak J-H, Zhou C, Yang C, Kondo K, Morota T, Takeda S, Sasaki H, Maki M (2007) Reticulate evolution and phylogeography in *Asarum* sect. *Asiasarum* (Aristolochiaceae) documented in internal transcribed specer sequences (ITS) of nuclear ribosomal DNA. Mol Phylogenet Evol 44:863–884
22. Takamatsu S, Hirata T, Sato Y (1998) Phylogenetic analysis and predicted secondary structures of the rDNA internal transcribed spacers of the powdery mildew fungi (Erysiphaceae). Mycoscience 39:441–453
23. Kiss L, Nakasone KK (1999) Ribosomal DNA internal transcribed spacer sequences do not support the species status of *Ampelomyces quisqualis* a hyperparasite of powdery mildew fungi. Curr Genet 33:362–367
24. Saenz GS, Taylor JW (1999) Phylogeny of the Erysiphales (powdery mildews) inferred from internal transcribed spacer ribosomal DNA sequences. Can J Bot 77(1):150–168
25. Lee SB, Taylor JW (1992) Phylogeny of five fungus-like protoctistan *Phytophthora* species, inferred from the internal transcribed spacers of ribosomal DNA. Mol Biol Evol 9(4):636–653
26. Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. Mol Ecol 3:113–118
27. Sikaroodi M, Lawrey JD, Hawksworth DL, Depriest PT (2001) The phylogenetic position of selected lichenicolous fungi: *Hobsonia*, *Illosporium*, and *Marchandiomyces*. Mycol Res 105(4):453–460
28. Stenroos S, Hyvönen J, Myllys L, Thell A, Ahti T (2002) Phylogeny of the genus *Cladonia* s.lat. (Cladoniaceae, Ascomycetes) inferred from molecular, morphological, and chemical data. Cladistics 18:237–278
29. Stenroos S, Stocker-Wörgötter E, Yoshimura I, Myllys L, Thell A, Hyvönen J (2003) Culture experiments and DNA sequence data confirm the identity of *Lobaria* photomorphs. Can J Bot 81(9):232–247

30. Wu Q-X, Mueller GM, Lutzoni FM, Huang Y-Q, Guo S-Y (2000) Phylogenetic and biogeographic relationships of eastern Asian and North American disjunct *Suillus* species (Fungi) as inferred from ITS sequences of nuclear ribosomal RNA ITS sequencies. Mol Phylogenet Evol 17:37–47

31. Martín MP, LaGreaca S, Schmitt I, Lumbsch HT (2003) Molecular phylogeny of *Diploschistes* inferred from ITS sequence data. Lichenologist 35:27–32

32. Fritz GN, Conn J, Cockburn A, Seawright J (1994) Sequence analysis of the ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culucidae). Mol Biol Evol 11:406–416

33. Nadler SA, Adams BJ, Lyons ET, DeLong RL, Melin SR (2000) Molecular and morphometric evidence for separate species of *Uncinaria* (Nematoda: Ancylostomatidae) in California sea lions and northern fur seals: hypothesis testing supplants verification. J Parasitol 86(5):1099–1106

34. Arnheim N, Krystal M, Schnickel R, Wilson G, Ryder O, Zimmer E (1980) Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc Natl Acad Sci USA 77(12):7323–7327

35. Stothard JR, Hughes S, Rollinson D (1995) Variation within the internal transcribed spacer (ITS) of ribosomal DNA genes of intermediate snail hosts within the genus *Bulinus* (Gastropoda: Planorbidae). Acta Trop 61(1):19–29

36. Pleyte KA, Duncan SD, Phillips RB (1992) Evolutionary relationships of the salmonid fish genus *Salvelinus* inferred from DNA sequences of the fish internal transcribed spacer (ITS 1) of ribosomal DNA. Mol Phylogenet Evol 1(3):223–230

37. van Oppen MJ, Mieog JC, Sanchez CA, Fabricius KE (2005) Diversity of algal endosymbionts (zooxanthellae) in octocorals: the roles of geography and host relationships. Mol Ecol 14:2403–2417

38. Rocap G, Distel DL, Waterbury JB, Chisholm SW (2002) Resolution of Prochlorococcus and Synechococcus ecotypes by using 16S–23S ribosomal DNA internal transcribed spacer sequences. Appl Environ Microbiol 68(3):1180–1191

39. Leclerc MC, Barrel V, Lecointre G, de Reviers B (1998) Low divergence in rDNA ITS sequences among five species of *Fucus* (Phaeophyceae) suggests a very recent radiation. J Mol Evol 46:115–120

40. Miadlikowska J, McCune B, Lutzoni F (2002) *Pseudocyphellaria perpetua*, a new lichen from Western North America. Bryologist 105(1):1–10

41. Koskinen S, Hyvönen J (2004) *Poganatum* (Polytrichales, Briophyta) revisited. In: Goffinet B(Ed) Molecular systematics of Briophytes. Monographs in Systematic Botany 98:150–167

42. Zhang Z, Kudo T, Nakajima Y, Wang Y (2001) Clarification of the relationship between the members of the family Thermomonosporaceae on the basis of the rDNS, 16S–23S rRNA internal transcribed spacer and 23S rDNA sequences and chemotaxonomic analysis. Int J Syst Evol Microbiol 51:373–383

43. Leblond-Bourget N, Philippe H, Mangin I, Decaris B (1996) 16S rRNA and 16S to 23S internal transcribed spacer sequence analysis reveal inter- and intraspecific *Bifidobacterium* phylogeny. Int J Sys Bacteriol 46(1):102–111

44. Brown MV, Fuhrman JA (2005) Marine bacterial microdiversity as revealed by internal transcribed spacer analysis. Aquat Microbiol Ecol 41(11):15–23

45. Chen F, Wang K, Kan J, Suzuki MT, Wommack KE (2006) Diverse and unique Picocyanobacteria in Chespeake Bay, revealed by 16S–23S rRNA internal transcribed spacer sequences. Appl Environ Microbiol 72(3):2239–2243

46. Jan-Roblero J, Magos X, Fernández L, Hernández-Rodríguez C, Le Borgne S (2004) Phylogenetic analysis of bacterial populations in waters of the former Texoco Lake, Mexico. Can J Microbiol 50(12):1049–1059

47. García-Martínez J, Bescós I, Rodríguez-Sala JJ, Rodríguez-Valera F (2001) RISSC: a novel database for ribosomal 16S–23S RNA genes spacer regions. Nucleic Acids Res 29(1):178–180

48. Coleman AW (2003) ITS2 is a double-edged tool for eukaryote evolutionary comparisions. Trends Genet 19(7):370–375

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

50. Feliner NG, Rosselló JA (2007) Better the devil you know? Guidelines for insightful utilization of rDNA ITS in species-level evolutionary studies in plants. Mol Phylogenet Evol 44:911–919

51. Phillips RL, McMullen MD, Enomoto S, Rubenstein I (1988) Ribosomal DNA in maize. In: Gustafson JP, Appels R (eds) Chromosome structure and function; impact of new concepts. Plenum Press, New York, pp 201–214

52. Lapitan NLV (1992) Organization and evolution of higher plant nuclear genomes. Genome 35:171–181

53. Schrader O, Ahne R, Fuchs J, Schubert I (1997) Computer-aided karyotype analysis of *Helianthus annuus* after Giemsa banding and fluorescence in situ hybridization. Chromosome Res 5:451–456

54. Fuchs J, Kühne M, Schubert I (1998) Assignment of linkage groups to pea chromosomes after karyotyping and gene mapping by fluorescent in situ hybridization. Chromosoma 107:272–276

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

56. Danna KJ, Workman R, Coryell V, Keim P (1996) 5S rRNA genes in tribe Phaseoleae: array size, number, and dynamics. Genome 39:445–455

57. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo W-L, Chen C, Zhai Y, Dairkee SH, Ljung B, Gray JW, Albertson DG (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20:207–211

58. Rogers SO, Bendich AJ (1987) Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. Plant Mol Biol 9(5):509–520

59. Patterson C (1987) Ortology and paralogy. In: Molecules and morphology in evolution: conflict or compromise? Cambridge University Press, London, pp 12

60. Sandersomn J, Doyle JJ (1992) Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. Syst Biol 41:4–17

61. Masterson J (1994) Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264:421–423

62. Liao D (1999) Concerted evolution: molecular mechanism and biological implications. Am J Hum Genet 64:24–30

63. Dover G (1994) Concerted evolution, molecular drive and natural selection. Curr Biol 4:1165–1166

64. Schlötterer C, Tautz D (1994) Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intra chromosomal exchanges drive concerted evolution. Curr Biol 4:777–783

65. Polanco C, Gonzalez AI, de la Fuente A, Dover GA (1998) Multigene family of ribosomal DNA in *Drosophila melanogaster* reveals contrasting patterns of homogenization for IGS and ITS spacer regions: a possible mechanism to resolve this paradox. Genetics 149:243–256

66. Arnheim N (1983) Concerted evolution of multigene families. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer, Sunderland, pp 38–61

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

68. Ritland CE, Ritland RK, Straus NA (1993) Variation in the ribosomal internal transcribed spacers (ITS1 and ITS2) among eight taxa of the *Mimulus guttatus* species complex. Mol Biol Evol 10:1273–1288

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

70. Raucher JT, Doyle JJ, Brown HD (2002) Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. Mol Ecol 11:2691–2702

71. Zimmer EA, Jupe ER, Walbot V (1988) Ribosomal gene structure, variation and inheritance in maize and its ancestors. Genetics 120:1125–1136

72. Koch MA, Dobeš C, Mitchell-Olds T (2003) Multiple hybrid formation in natural populations: concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). Mol Biol Evol 20(3):338–350

73. Wang QD, Zhang T, Wang JB (2007) Phylogenetic relationships and hybrid origin of *Potamogeton* species (Potamogetonaceae) distributed in China: insights from the nuclear ribosomal internal transcribed spacer sequences (ITS). Plant Syst Evol 267(1–4):65–78

74. Ainouche ML, Bayer RJ (1997) On the origins of the tetraploid *Bromus* species (section *Bromus*, Poaceae): insights from internal transcribed spacer sequences of nuclear ribosomal DNA. Genome 40(5):730–743

75. Löhne C, Borsch T, Jacobs SWL, Hellquist CB, Wiersema JH (2008) Nuclear and plastid DNA sequences reveal complex reticulate patterns in Australian water-lilies (*Nymphaea* subgenus *Anecphya*, Nymphaeaceae). Aust Sys Bot 21:229–250

76. Fuertes Aguilar J, Rosselló JA, Nieto Feliner G (1999) Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Mol Ecol 8(8):1341–1346

77. Franzke A, Mummenhoff K (1999) Recen hybrid speciation in Cardamine (Brassicaceae)–conversion of nuclear ribosomal ITS sequences in statu nascendi. Theor Appl Genet 98(5):831–834

78. Hodkinson TR, Chase MW, Takahashi C (2002) The use of DNA sequencing (ITS and *trn*L-F), AFLP, and fluorescent in situ hybridization to study allopolyploid *Miscanthus* (Poaceae). Am J Bot 89:279–286

79. Whittall J, Liston A, Gisler S, Meinke RJ (2000) Detecting nucleotide additivity from direct sequences is a SNAP: an example from *Sidalcea* (Malvaceae). Plant Biol 2:211–217

80. Sang T, Crawford DJ, Stuessy TF (1995) Documentation of reticulate evolution in peonies (*Peonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proc Natl Acad Sci USA 92:6813–6817

81. Quijada A, Liston A, Robinson W, Alvarez Buylla E (1997) The ribosomal ITS region as a marker to detect hybridization in pines. Mol Ecol 6:995–996

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

83. Samuel R, Bachmair A, Jobst J, Ehrendorfer F (1998) ITS sequences from nuclear rDNA suggest phylogenetic relationships between Euro-Mediterranean, East Asiatic and north American taxa of *Quercus* (Fagaceae). Plant Syst Evol 211:129–139

84. Manos P, Doyle JJ, Nixon KC (1999) Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). Mol Phylogenet Evol 12:333–349

85. Bailey CD, Carr TG, Harris SA, Hughes CE (2003) Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol Phylogenet Evol 29:435–455

86. Schultz J, Müller T, Achtziger M, Seibel PN, Dandekar T, Wolf M (2006) The internal transcribed spacer 2 database-a web based server for (not only) low level phylogenetic analyses. Nucleic Acids Res 34:704–707

87. Michot B, Joseph N, Mazan S, Bachellerie J (1999) Evolutionarily conserved structuresal features in the ITS2 of mammalian pre-RNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. Nucleic Acids Res 27:2271–2282

88. Gonzales IL, Chambers C, Gorski JL, Stambolian D, Schmickel RD, Sylvester JE (1990) Sequence and structure correlation of human ribosomal transcribed spacers. J Mol Biol 212:27–35

89. Yeh LC, Lee JC (1990) Structural analysis of the internal transcribed spacer 2 of the prescursor ribosomal RNA from *Saccharomyces cerevisiae*. J Mol Biol 211:699–712

90. van der Sande CAFM, Kwa M, van Nues RW, van Heerikhuizen H, Raue HA, Planta RJ (1992) Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. J Mol Biol 223:899–910

91. Van Nues RW, Rientjes JMJ, Morre SA, Mollee E, Planta RJ, Venema J, Raue HA (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

92. Joseph N, Krauskopf E, Vera M, Michot B (1999) Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. Nucleic Acids Res 27:4533–4540

93. Cote CA, Greer CL, Peculis BA (2002) Dynamical conformational model for the role of the ITS2 in pre-RNA processing in yeast. RNA 8:786–797

94. Chen CA, Chang CC, Wei NV, Chen CH, Lein YT, Lin HE, Dai CF, Wallace CC (2004) Secondary structure and phylogenetic utility of the ribosomal internal transcribed spacer 2 (ITS2) in scleractinian corals. Zool Stud 43:759–771

95. Gottschling M, Plötner J (2004) Secondary structure models of the nuclear internal transcribed spacer regions and 5.8S rRNA in Calciodinelloideae (Peridiniaceae) and other dinoagellates. Nucleic Acids Res 32:307–315

96. Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48:443–453

97. Schultz J, Maisel S, Gerlach D, Müller T, Wolf M (2005) A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA 11:361–364

98. Wolf M, Friendrich J, Dandekar T, Müller T (2005) CBCAnalyzer: inferring phylogenies based on compensatory base changes in RNA secondary structures. In Silico Biol 5:291–294

99. Müller T, Philippi N, Dandekar T, Schultz J, Wolf M (2007) Distinguishing species. RNA 13:1469–1472

100. Kimura M (1985) The role of compensatory neutral mutation in molecular evolution. J Genet 64:7–19

101. Coleman AW, Vacquier VD (2002) Exploring the phylogenetic utility of ITS sequences for animals: a test case for abalone (*Haliotis*). J Mol Evol 54:246–257

102. Aguilar C, Sánchez JA (2007) Phylogenetic hypotheses of gorgoniid octocorals according to ITS2 and their predicted RNA secondary structures. Mol Phylogenet Evol 43:774–786

103. Sánchez JA, Lasker HR, Taylor DJ (2003) Phylogenetic analyses among octocorals (Cnidaria) according to mitochondrial and nuclear DNA sequences (lsu-rRNA 16S, and ssu-rRNA 18S) support two convergent clades of branching gorgonians. Mol Phylogenet Evol 29:31–42

104. Liu JS, Schardl CL (1994) A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. Plant Mol Biol 26:775–778

105. Goertzen LR, Cannone JJ, Gutell RR, Jansen RK (2003) ITS secondary structure derived from comparative analysis: implications for sequence alignment and phylogeny of the Asteraceae. Mol Phylogenet Evol 29:216–234

106. Van Nues RW, Rientejes JMJ, van der Sande CAFM, Zerp SF, Sluiter C, Venema J, Planta RJ, Raué HA (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

107. Schilthuizen M, Gittenberger E, Gultyaev AP (1995) Phylogenetic relationships inferred from the sequence and secondary structure of ITS1 rRNA in Albinaria and putative Isabellaria species (Gastropoda, Pulmonata, Clausiliidae). Mol Phylogenet Evol 4:457–462

108. Coleman AW, Preparata RM, Mehrotra B, Mai JC (1998) Derivation of the secondary structure of the ITS-1 transcript in Volvocales and its taxonomic correlation. Protist 149:135–146

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

110. Weider LJ, Elser JJ, Crease TJ, Mateos M, Cotner JB, Markow TA (2005) The functional significance of ribosomal (r) DNA variation: impacts on the evolutionary ecology of organisms. Annu Rev Ecol Evol Syst 36:219–242

111. Gorokhova E, Dowling TA, Crease T, Weider LJ, Elser JJ (2002) Functional and ecological significance of rDNA IGS variation in a clonal organism under divergent selection for production rate. Proc R Soc B 269:2373–2379

112. Reeder RH (1984) Enhancers and ribosomal gene spacers. Cell 38:349–351

113. Reeder RH (1990) rRNA synthesis in the nucleolus. Trends Genet 6:390–394

114. Alvares LE, Brison O, Ruiz IR (1998) Identification of enchancer-like elements in the ribosomal intergenic spacer of Odontophrynus americanus 2n and 4n (Amphibia, Anura). Genetica 104:41–44

115. Flavell RB (1993) Variation in structure and expression of ribosomal DNA loci in wheat. Genome 31:963–968

116. Paule MR, Lofquist AK (1996) Organization and expression of eukaryotic ribosomal RNA genes. In: Zimmerman RA, Dahlberg AE (eds) Ribosomal RNA: structure, evolution, processing, and function in protein biosynthesis. CRC Press, New York, pp 395–420

117. Fernnández M, Ruiz ML, Linares C, Fominaya A, Pérez de la Vega M (2005) 5S rDNA genome regions of Lens species. Genome 48:937–942

118. Bhatia S, Singh Negi M, Lakshmikumaran M (1996) Structural analysis of the rDNA intergenic spacer of Brassica nigra:evolutionary divergence of the spacers of the three diploid Brassica species. J Mol Evol 43:460–468

119. Fernández M, Polanco C, Ruiz ML, Perez de la Vega M (2000) A comparative study of the structure of the rDNA intergenic spacer of Lens culinaris Medik., and other legume species. Genome 43:597–603

120. McIntyre CL, Clarke BC, Appels R (1988) DNA sequence analyses of the ribosomal spacer regions in the Triticaceae. Plant Sys Evol 160:91–104

121. Perry KL, Palukaitis P (1990) Transcription of tomato ribosomal DNA and the organization of the intergenic spacer. Mol Gen Genet 221:102–112

122. Hemleben V, Ganal M, Grestner J, Schiebel K, Torres RA (1988) Organization and length heterogeneity of plant ribosomal RNA genes. In: Kahl G (ed) The architecture of eukaryotic gene. VHC, Weinheim, pp 371–384

123. Yakura K, Kato A, Tanifuji S (1983) Structural organization of ribosomal DNA in four Trillium species and Paris verticillata. Plant Cell Phys 24(7):1231–1240

124. Choumane W, Heizmann P (1988) Structure and variability of nuclear ribosomal genes in the genus Helianthus. Theor Appl Genet 76:481–489

125. Volkov R, Kostishin S, Ehrendorfer E, Schweizer D (1996) Molecular organization and evolution of the external transcribed rDNA spacer region in two diploid relatives of Nicotiana tabacum (Solanaceae). Plant Syst Evol 201:117–129

126. Volkov RA, Bachmair A, Panchuk II, Kostyshyn SS, Schweizer D (1999) 25S–18S rDNA intergenic spacer of Nicotiana sylvestris (Solanaceae): primary and secondary structure analysis. Plant Syst Evol 218:89–97

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

128. Maughan PJ, Kolano BA, Maluszynska J, Coles ND, Bonifacio A, Rojas J, Coleman CE, Stevens MR, Fairbanks DJ, Parkinson SE, Jellen EN (2006) Molecular and cytological characterization of ribosomal RNA genes in Chenopodium quinoa and Chenopodium berlandieri. Genome 49:825–839

129. Da Rocha PS, Bertrand H (1995) Structure and comparative analysis of the rDNA intergenic spacer of Brassica rapa. Implications for the function and evolution of the Cruciferae spacer. Eur J Biochem 229:550–557

130. King K, Torres RA, Zentgraf U, Hembleben V (1993) Molecular evolution of the intergeneric spacer in the nuclear ribosomal RNA genes of Cucurbitaceae. J Mol Evol 36:144–183

131. Tautz D, Tautz C, Webb D, Dover GA (1987) Evolutionary divergence of promoters and spacers in the rDNA family of four Drosophila species. Implications for molecular coevolution in multigene families. J Mol Biol 195:525–542

132. Ryu SH, Do YK, Hwang UW, Choe CP, Kim W (1999) Ribosomal DNA intergenic spacer of the swimming crab, Charybdis japonica. J Mol Evol 49:806–809

133. Takaiwa F, Kikuchi S, Oono K (1990) The complete nucleotide sequence of the intergenic spacer between 25S and 17S rDNAs in rice. Plant Mol Biol 15:933–935

134. Kato A, Yakura K, Tanifuji S (1984) Sequence analysis of Vicia faba repeated DNA, the FokI repeat element. Nucleic Acids Res 12:6415–6426

135. Barker RF, Harberd NP, Jarvis MG, Flavell RB (1988) Structure and evolution of the intergenic region in a ribosomal DNA repeat unit of wheat. J Mol Biol 201:1–17

136. Maggini F, Gelati MT, Spolverini M, Frediani M (2008) The intergenic spacer region of the rDNA in Olea europaea L. Tree Genet Genomes 4:293–298

137. Bauer N, Horvat T, Biruš I, Vičić V, Zoldoš V (2009) Nucleotide sequence structural organization and length heterogeneity of ribosomal DNA intergenic spacer in Quercus petraea (Matt.) Liebl. and Q. robur L. Mol Genet Genomics 281:207–221

138. Dover GA, Tautz D (1986) Conservation and divergence in multigene families: alternatives to selection and drift. Philos Trans R Soc Lond B Biol Sci 312:275–289

139. Ryu S, Do Y, Fitch DHA, Kim W, Mishra B (2009) Dropout alignment allows homology recognition and evolutionary

analysis if rDNA intergenic spacers. J Mol Evol. doi:10.1007/s00239-008-9090-8

140. Hershkovitz MA, Zimmer EA, Hahn WJ (1999) Ribosomal DNA sequences and angiosperm systematics. In: Hollingsworth PM, Bateman RM, Gornall RJ (eds) Molecular systematics and plant evolution. Taylor and Francis, London, pp 268–326

141. Starr JR, Harris SA, Simpson DA (2003) Potential of the 5′ and 3′ ends of the intergenic spacer (IGS) of rDNA in the Cyperaceae: new sequences for lower-level phylogenies in sedges with an example from *Uncinia* Pers. Int J Plant Sci 164:213–227

142. Borisjuk N, Hemleben V (1993) Nucleotide sequence of the potato rDNA intergenic spacer. Plant Mol Biol 21:381–384

143. Granneman S, Baserga SJ (2005) Crosstalk in gene expression: coupling and co-regulation of rDNA transcription, pre-ribosome assembley and pre-rRNA processing. Curr Op Cell Biol 17:281–286

144. Azuma M, Toyama R, Laver E, Dawid IB (2006) Perturbation of rRNA synthesis in the bap28 mutation leads to apoptosis mediated by p53 in zebrafish central nervous system. J Biol Chem 281(19):13309–13316

145. Baldwin BG, Markos S (1998) Phylogenetic utility of the external transcribed spacer (ETS) of 18S–26S rDNA: congruence of ETS and ITS trees of Calycadenia (Compositae). Mol Phylogenet Evol 10:449–463

146. Linder CR, Goertzen LR, Vanden Heuvel B, Francisco-Ortega J, Jansen RK (2000) The complete external transcribed spacer of 18S–26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. Mol Phylogenet Evol 14:285–303

147. Chandler GT, Bayer RJ, Crisp MD (2001) A molecular phylogeny of the endemic Australian genus *Gastrolobium* (Fabaceae: Mirbelieae) and allied genera using chloroplast and nuclear markers. Am J Bot 88:1675–1687

148. Wright SD, Yong CG, Wichman SR, Dawson JW, Gardner RC (2001) Stepping stones to Hawaii: a trans-equatorial dispersal pathway for *Metrosidos* (Myrtaceae) inferred from rDNA (ITS + ETS). J Biogeogr 28:769–774

149. Tremousaygue D, Laudie M, Grellet F, Delseny M (1992) The *Brassica oleracea* rDNA spacer revisited. Plant Mol Biol 18:1013–1018

150. Cordesse F, Cooke R, Tremousaygue D, Greellet F, Delseny M (1993) Fine structure and evolution of the rDNA intergenic spacer in rice and other cereals. J Mol Evol 36:369–379

151. McMullen MD, Hunter B, Phillips RL, Rubenstein I (1986) The structure of the maize ribosomal DNA spacer region. Nucleic Acids Res 14:4953–4968

152. Schiebel K, von Waldburg G, Gerstner J, Hemleben V (1989) Termination of transcription of ribosomal RNA genes of mung bean occurs within a 175 bp repetitive element of the spacer region. Mol Gen Genet 218:302–307

153. Bena G, Jubier MF, Olivieri I, Lejeune B (1998) Ribosomal external and internal transcribed spacers: combined use in the phylogenetic analysis of *Medicago* (Leguminosae). J Mol Evol 46:299–306

154. Jorgensen RA, Cluster PD (1988) Modes and tempos in the evolution of nuclear ribosomal DNA: new characters for evolutionary studies and new markers for genetic and population studies. Ann Missouri Bot Gard 75:1238–1247

155. Bena G (2001) Molecular phylogeny supports the morphologically based taxonomic transfer of the "medicagoid" *Trigonella* species to the genus *Medicago* L. Plant Syst Evol 229:217–236

156. Yamashiro T, Fukuda T, Yokoyama J, Maki M (2004) Molecular phylogeny of *Vincetoxicum* (Apocynaceae-Asclepladoideae) based on the nucleotide sequences of cpDNA and nrDNA. Mol Phylogenet Evol 31:689–700

157. Roalson EH, Friar EA (2004) Phylogenetic relationships and biogeographic patterns in North American members of *Carex* section *Acrocystis* (Cyperaceae) using nrDNA ITS and ETS sequence data. Plant Syst Evol 243:175–187

158. Kay KM, Reeves PA, Olmstead RG, Schemske DW (2005) Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. Am J Bot 92:1899–1910

159. Hipp AL, Reznicek AA, Rothrock PE, Weber JA (2006) Phylogeny and classification of *Carex* section *Ovales* (Cyperaceae). Int J Plant Sci 167:1029–1048

160. Kim K-J, Mabry TJ (1991) Phylogenetic and evolutionary implications of nuclear ribosomal DNA variation in dwarf dandelions (*Krigia*, Lactuceae, Asteraceae). Plant Syst Evol 177:53–69

161. Vander Stappen J, Marant S, Volckaert G (2003) Molecular characterization and phylogenetic utility of the rDNA external transcribed spacer region in *Stylosanthes* (Fabaceae). Theor Appl Genet 107:291–298

162. Vander Stappen J, De Laet J, Gama Lopez S, Van Campenhout S, Volckaert G (2002) Phylogenetic analysis of *Stylosanthes* (Fabaceae) based on the internal transcribed spacer region (ITS) of nuclear ribosomal DNA. Plant Syst Evol 234:27–51

163. Volkov RA, Kamarova NY, Panchuk II, Hemleben V (2003) Molecular evolution of rDNA external transcribed spacer and phylogeny of sect. *Petota* (genus *Solanum*). Mol Phylogenet Evol 29:187–202

164. Kamarova NY, Grimm GW, Hemleben V, Volkov RA (2009) Molecular evolution of 35S rDNA and taxonomic status of *Lycopersicon* within *Solanum* sect. *Petota*. Plant Syst Evol. doi:10.1007/s00606-008-0091-2

165. Oh S-H, Potter D (2005) Molecular phylogenetic systematics and biogeography of tribe Neillieae (Rosaceae) using DNA sequences of cpDNA, rDNA, and LEAFY. Am J Bot 92(1):179–192

166. Markos S, Baldwin BG (2002) Structure, molecular evolution, and phylogenetic utility of the 5′ region of the external transcribed spacer of 18S–26S rDNA in Lessingia (Compositae, Asteraceae). Mol Phylogenet Evol 23:214–228

167. Tucci GF, Simeone MC, Gregori C, Maggini F (1994) Intergenic spacers of rRNA genes in three species of the Cynareae (Asteraceae). Plant Syst Evol 190:187–193

168. Kluge AG (1989) A concern for evidence and phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst Zool 38:7–25

169. Nixon KC, Carpenter JM (1996) On simultaneous analysis. Cladistics 12:221–241

170. Grant T, Kluge AG (2003) Data exploration in phylogenetic inference: scientific, heuristic, or neither. Cladistics 19:379–418

171. Varón A, Vinh LS, Bomash I, Wheeler WC (2009) POY 4.1.1. American Museum of Natural History. http://research.amnh.org/scicomp/projects/poy.php

172. Wheeler WC, Cartwright P, Hayashi CY (1993) Arthropod phylogeny: combined approach. Cladistics 9:1–39

173. Huson DH, Steel M (2004) Phylogenetic trees based on gene content. Bioinformatics 20(13):2044–2049

174. Huson DH, Kloepper TH (2005) Computing recombination networks from binary sequences. Bioinformatics 21(2):159–165

175. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267

176. Bryant D, Moulton V (2004) NeighborNet: An agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 21(2):255–265

177. Than C, Ruths D, Nakhleh L (2008) PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinformatics 9:322

178. Ruths D, Nakhleh L, Ram PT (2008) Rapidly exploring structural and dynamic properties of signaling networks using PathwayOracle. BMC Syst Biol 2:76

179. Kanj I, Nakhleh L, Than C, Xia G (2008) Seeing the trees and their branches in the network is hard. Theor Comput Sci 401:153–164

180. Kanj I, Nakhleh L, Xia G (2008) The compatibility of binary characters on phylogenetic networks: complexity and parameterized algorithms. Algorithmica 51:99–128