Molecular Phylogenetics and Evolution 50 (2009) 197–203

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Review

Is there a molecular key to the level of "biological species" in eukaryotes? A DNA guide

Annette W. Coleman *

Division of Biology and Medicine, Brown University, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 18 March 2008
Revised 6 October 2008
Accepted 8 October 2008
Available online 21 October 2008

Keywords:
Biological species
Interbreeding limits
ITS-2
rRNA secondary structure
Speciation
Taxonomy

ABSTRACT

DNA sequences, powerful for phylogeny, have not yet proven as rewarding for taxonomic categorization purposes. However, further analyses of one locus, the second Internal Transcribed Spacer of the nuclear ribosomal gene cistron, has suggested a high degree of predictability across eukaryotes. Comparison of the secondary structure of ITS2 transcripts reveals its most conserved region, on the 5′-side of helix III. Comparison of this 5′ 30 bp highly conserved region with the extent of sexual compatibility in a clade of organisms produces two useful predictions: identity of this region predicts meaningful intercrossing ability, and, difference of even one CBC pairing in this region predicts total failure of crossing. Previous to the appearance of the first CBC in the highly conserved portion, all gametic compatibility has been lost, thanks to the parallel evolutionary changes in genes controlling mating. These two landmark events help to delimit the level of interbreeding taxa.

1. Introduction

The very basis of our taxonomic system and its growth over the past 250 years is the presumption that individuals very similar in morphology are those most closely related, constituting a species. Yet the category of morphological "species" is in the eye of the beholder. Nature does not produce species, but rather individuals, populations and clades, of continuing evolutionary change. The second major approach to classification states that individuals capable of interbreeding are each others' nearest relatives, as so importantly emphasized by Dobzhansky (1940) and his coworkers more than 50 years ago in their recognition of the "biological" species. They emphasized the critical significance of the loss of interbreeding potential, the major step that dooms clades to remain genetically separate forever after. Hence, in some ideal taxonomic world, one morphology = one interbreeding group = one species; such an entity would be morphologically uniform, and sexually isolated from other clades.

Such is not our world. Many individual populations/clades are indeed, at this point in time, uniform in morphology and capable of interbreeding, and are also unable to cross with individuals of other clades. For many other clades, however, morphological uniformity does not completely parallel interbreeding capacity, or breeding status of a morphological group may remain uncertain/undetermined, challenging the taxonomist to seek other critera.

Given the known pitfalls of our primarily morphology-based taxonomic system, how great then must have been the hopes of taxonomists, with the advent of DNA sequencing, to find at last a molecular surrogate for some level of the speciation process. Alas, this hope has been drowned in the accumulation of sequencing data on various candidate loci. Nucleotide change does occur, genera differ more than do their component pairs of species, but the nucleotide change is continuous, with no gap, no point of reference correlatable with some facet of speciation.

One approach would be to concentrate on understanding the genes involved in sexual behavior, particularly those controlling gamete approach and fusion; this has the advantage of focussing on a process common to all eukaryotes. Activities such as the response to mutual attraction by modification of gamete surface components, specific agglutination, and gamete fusion, require the coordinated action of a considerable number of genes. Such genes might seem to be the ideal object of study for evaluating the evolution of the speciation process. However, our understanding of their comparative molecular aspects is still in its infancy (Swanson and Vacquier, 2002; Hamaji et al., 2008).

Recently, one such DNA region (the Internal Transcribed Spacer 2 region, ITS2, of the nuclear ribosomal cistrons) has been found to vary in sequence and secondary structure in a way that correlates highly with taxonomic classification. The ITS2 region of nuclear DNA provides a powerful tool because not only can sequence be compared but also aspects of the secondary structure formed by the initial RNA transcript. Müller et al. (2007) compared the entire ITS2 of 1373 sequences of species to see how homogeneous were the sequences within species compared to their nearest congeners.

* Fax: +1 401 863 2421.
E-mail address: Annette_Coleman@brown.edu

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.10.008

They found that for 93% of cases where two taxa differed somewhere in their ITS2 by a CBC (Compensatory Base Change, an altered pairing in a helix of the secondary structure of the ITS2 RNA transcript—see Fig. 1), they were classified as different species. Thus, in the vast majority of organisms they surveyed, a discrete property of the ITS2 (no CBC anywhere) characterized members of a species.

This result suggests an outer limit to the sequence variation accompanying current taxonomic assignment of species names, whatever their criteria. However, it says nothing about the level of biological species. Here we use the ITS2 region to explore further its information content, concentrating on characteristics that accompany separation of clades into two or more sexually isolated subclades. We evaluate its predictive powers for discriminating interbreeders, and derive a simplified sequenced-based technique for its application to evaluation of the biological species level among eukaryotes.

2. Properties of ITS2

ITS2 sequence comparisons are perhaps the most common source of phylogenetic reconstructions at the species, genus and family level among all eukaryotes (Alverez and Wendel, 2003; Bailey et al., 2003). The locus of interest is in the nuclear ribosomal cistrons, a region lying between the 5.8S RNA gene and the LSU RNA gene. The cistron is transcribed initially into a single long transcript. During "processing" in the nucleolus to release the ribosomal RNAs, this spacer region is degraded to nucleotides.

It was early reconized that ITS2 shares common features of sequence and secondary structure, conserved even above the family level (Hershkovitz and Zimmer, 1996). Analysis of the secondary structure formed by the RNA transcript as it folds back on itself at transcription, has generally been less commonly done, but has proven extremely useful in aiding proper sequence alignment (Mai and Coleman, 1997). As shown in Fig. 1, the typical ITS2 of eukaryotes forms four helices (Coleman, 2003). This secondary structure can be derived by applying the RNA folding program Mulfold (http://www.bioinfo.rpi.edu/applications/mfold/rna/form1.cgi) to all the sequences of a group of closely related organisms, comparing the multiple folds produced by the program to find the single type common to all the taxa, and then checking positions of nucleotide variation. Essentially all variation in positions paired in helices (Fig. 1) should show only those substitutions that preserve the pairing, both one-sided (hemi-CBC) and two sided (CBC). These characteristics are considered "proof" of the proposed RNA secondary structure, as first applied to studies of the ribosomal RNAs (Gutell et al., 1994). The realization that the same four helix structure characterizes the ITS2 transcript in essentially all eukaryotes has only been recognized recently (Coleman, 2003, 2007; Müller et al., 2007). This universality has led to computer automation of

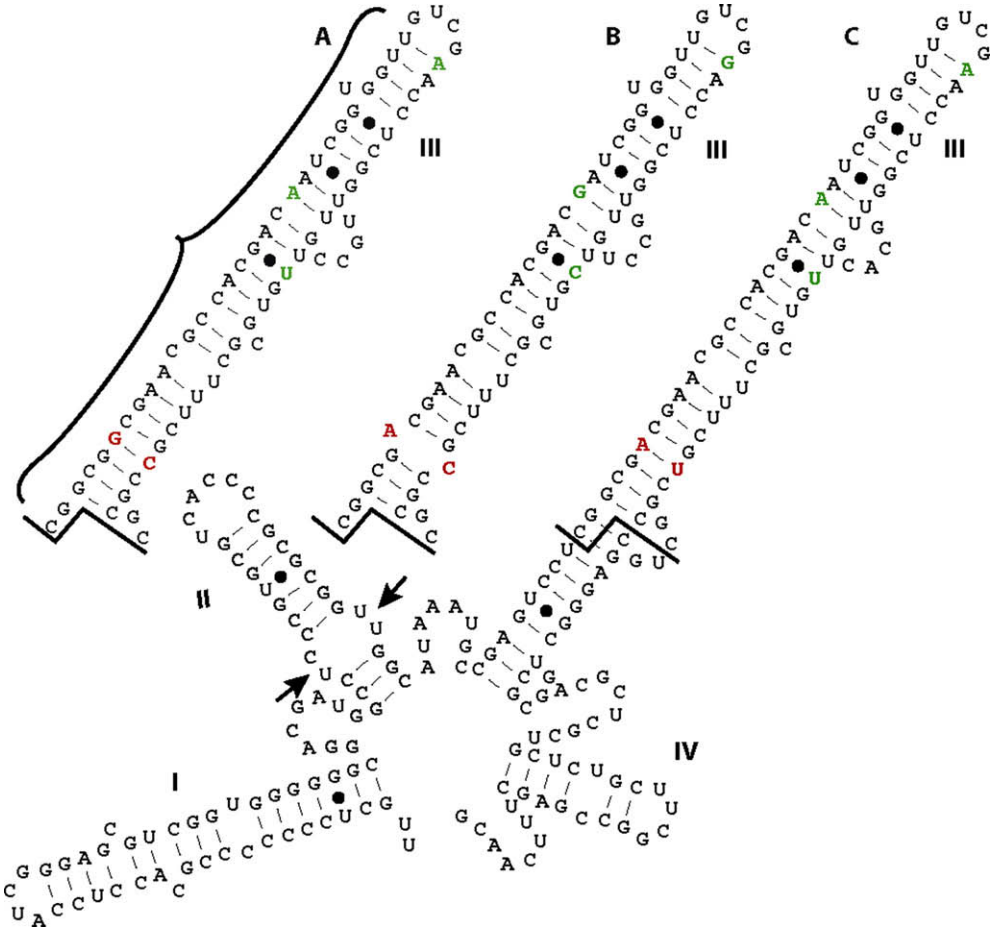

Fig. 1. Secondary structure of ITS2, and comparisons of helix III among examples from the apple subfamily Maloideae, (A) *Amelanchier*, (B) *Mespilus* and (C) *Photinia*. The entire ITS2 structure of *Photinia* is shown to illustrate the typical ITS2 secondary structure of eukaryotes; hallmark characteristics are marked in helix II by arrowheads (pyrimidine–pyrimidine bulge) and in helix III by a bracket (5′ 30 most conserved nucleotide positions). Helix III pairing variants among species of the family include one CBC (red) and three hemi-CBCs (green). *Amelanchier* fails to cross with any of the several genera that differ by the CBC (Robertson et al., 1991). (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

the search and folding computations with broadly applicable tools (Wolf et al., 2005).

The secondary structure of the ITS2 initial transcript contains further useful information that can be derived from analysis of the functional importance of particular nucleotide positions. From such studies of clades of related organisms, it has become clear that the first and fourth helices are the most evolutionarily variable, useful primarily for species and subspecies comparisons. Helices II and III are more conserved in sequence than helices I and IV, and recognizable among all eukaryote ITS2 structures by certain invariant features. The second helix typically is fairly short and has a pyrimidine-pyrimidine mismatch near the base; the third helix is longer, sometimes even bearing a branch, and has on the 5′-side near the tip the single most highly conserved stretch of nucleotides in the entire ITS2 (Fig. 1).

These conserved features, taken together, suggest that it is helices II and III that might contain the molecular signals necessary for successful RNA transcript processing. As summarized in Cote et al. (2002), there are one or two initial cleavages in processing of the ITS2 RNA transcript and they occur in the region extending along the 3′ side of helix II through the 5′-side of helix III. Thus the region thought to be most important functionally is also the region most conserved evolutionarily.

3. How does ITS2 sequence evolution parallel an aspect of biological species evolution?

This comparison requires a plethora of data, including complete ITS2 sequences for a group of related organisms, the derivation of their common RNA transcript secondary structure, and for the same group, experimental analyses of their intercrossing potential, both positive and negative. Furthermore, one cannot use the more typical definitions of "biological species" because such information is largely unavailable. Instead we have used for the pan-eukaryote comparisons a sexual characteristic, gamete compatibility sufficient to form a recognizable zygotic stage (members of the same "Z clade"). This definition is the most accessable to experimental analysis without extensive expenditure of time and effort; hence, more data are available.

The clades of organisms for which such data have been compared are found in Table 1. Examples from additional plants and green algae (Coleman, 2000) have been omitted because they are only repetitious; other groups are absent from the analysis for lack of the necessary data, most notably, various familiar mammalian groups. Here, it is the ITS2 sequence data that are lacking since, historically, phylogenetic studies in animals early concentrated on mitochondrial DNA regions, not ITS.

Table 1 summarizes the comparisons for 30 clades of protistans, plants and animals for which sufficient data are available. The analysis includes a total of >100 biological species, and more than 400 ITS2 sequences, including outgroups. Three different categories could be recognized, A, B and C, as cited in Table 1 and Fig. 2. Type B consists of only two cases, where one taxonomic species = one morphology = one Z clade (very broad "biological species" as defined here). Examples of Type C, single taxonomic species containing >1 Z clade, are confined to the microprotistan taxa examined here, but type A (a single Z clade encompassing several species or even genera) is present in both micro- and macroprotistans and is the near universal case for plants and animals. These three patterns of taxonomic classification are crudely diagrammed in Fig. 2, with a region of shading illustrating the decreasing intensity of sexual interaction from total interbreeding, through loss of sufficient interbreeding capacity to qualify as "biological species" in the conventional usage, to Z clade limit, and finally to the loss of any detectable gamete interaction.

Where multiple taxonomic species/genera are capable of gamete interaction, their ITS2 sequences are extremely similar, even identical for 30 or more nucleotides lying on the 5′-side of helix III, the most evolutionarily conserved region of ITS2. At the other extreme, where one species encompasses more than one interbreeding group, their ITS2 sequences contain considerable variation among the 30 relatively conserved helix III positions. Particularly striking examples help to illustrate these two extremes. For A, three Hawaiian genera of silverswords (*Argyroxipium, Dubautia, Wilkesia*), even those that hybridize poorly (Baldwin et al., 1991), display, in the 5′ ITS2 helix III, 38 identical nucleotides which encompass 26 helix pairings. At the other extreme cited, type C, the morphologically uniform algal species *Pandorina morum* encompasses multiple variant ITS2 sequences among the >25 cryptic biological species. Their longest common region of identity in ITS2 is only eight nucleotides, in the 5′ conserved region of helix III, and the taxa classified in the species encompass nine CBCs in helices II and III of ITS2 (Coleman, 2001).

4. Development of genetic isolation

Various clades/populations are at different stages of evolutionary divergence. Some biological species contain only highly compatible interbreeders, totally incapable of interacting with organisms from different species, while at least a few groups show, over a geographic range of tested isolates, some pairings that are highly interfertile, others that display lower levels of gamete interaction, some where a few zygotes are formed but never fully develop/germinate, and even cases where individuals interfertile with some strains of the clade will not themselves directly intermate. As ITS2 differences between potential mates increase, sexual compatibility and zygote productivity decrease. An example of this correlation has already been nicely illustrated for *Coreopsis* species by Archibald et al. (2005, see Fig. 2). The 25 *Gonium pectorale* strains serve to illustrate in more detail the parallels typical as reproductive isolation and ITS2 evolve. Overall, Southern hemisphere isolates interbreed successfully *inter se*, and so also do Northern hemisphere isolates, but only a few Southern × Northern crosses even produce zygotes, and they have no issue. In *G. pectorale*, the sole CBC nucleotide position in helix III distinguishes those strains of the Northern hemisphere totally incapable of crossing with those of the Southern hemisphere (Fabry et al., 1999).

5. Using helix III alone, is there a simple sequence-based way to predict potential crossing or total loss of interbreeding?

Deriving secondary structures of ITS2 is laborious, though many examples that can serve as guides for an order or class of eukaryotes are now in the literature and on line (Schultz et al., 2006). Furthermore, it would be naive to set guidelines to all eukaryotes on the basis of the whole helix III, or even a percentage of helix III, for in some organisms, ITS2 helix III may be quite long, or branched, with insertions in the stem or at the terminus (see Fig. 2 in Mai and Coleman, 1997). The additional nucleotides are not under such stringent selection. To take this situation into account for all eukaryotes, we have derived a generalization for the contiguous 5′-side 30 nucleotide positions of helix III encompassing the most highly conserved nucleotide sequence (thus avoiding confusions engendered by different ITS2 lengths). Interestingly, two simple rules apply to all the examples in Table 1.

1. *Organisms that differ by even one CBC in this region also are completely unable to cross.* There is no example yet of an exception to this rule. Organisms may have an altered nucleotide in a bulge (single stranded region) or even a hemi-CBC among these 30 and

Table 1
Comparison of ITS2 sequences versus level of biological species versus taxonomic category.

Type	Organisms	# Biological species	Identity of 30 nt motif	5′ helix III identity	All ITS2 identical	Reference
Algae/Protozoa						
A	*Chlamydomonas* *smithii* *incerta* *reinhardtii*	1 (8 isolates)	Yes	Yes	No	Pröschold et al. (2005)
A	*Chlamydomonas* *moewusii* *eugametos*	1 (4 isolates)	Yes	Yes	No	Mai and Coleman (1997)
C	*Chlamydomonas allensworthii*	5 (13 isolates)	Yes	Yes	Yes	Coleman (2001)
B	*Astrephomene* *gubernaculifera* *perforata*	2 (10 isolates)	Yes	Yes	No	Mai and Coleman (1997)
C	*Gonium pectorale*	1–2 (25 isolates)	Yes	Yes	No	Fabry et al. (1999)
B	*Blidingia minima*	1 (9 isolates)	Yes	Yes	No	Woolcott et al. (2000)
C	*Pandorina morum* *colemaniae* *smithii*	>25 (62 isolates)	Yes/no	No	No	Coleman (2001)
C	*Closterium ehrenbergianum*	11 (47 isolates)	Yes	No	No	Denboh et al. (2003)
C	*Sellaphora* *pupula* *laevissima*	>2 (16 isolates)	Yes	Yes	No	Behnke et al. (2004)
C	*Pseudonitzschia* *delicatissima* *pseudodelicatissima* *cuspidata* *calliantha*	5 (19 isolates)	Yes	Yes/no	No	Amato et al. (2007)
A	*Ectocarpus* *Kuckuckia*	1 (10 isolates)	Yes	Yes	No	Stache-Crain et al. (1997)
A	*Laminaria* *Kjellmannia*	1 (11 isolates)	Yes	Yes	No	Yotsukura et al. (1999)
A	*Alaria*	1 (5 spp.)	Yes	Yes	No	Lane et al. (2007)
A	*Macrocystis* *Pelagophycus* *Nereocystis*	1 (5 spp., 12 isolates)	Yes	Yes	No	Coyer et al. (2001)
A	Red algal parasites (Delessereaceae-3 spp.)	2 (8 isolates)	Yes	Yes	No	Goff et al. (1997)
C	*Paramecium aurelia* complex	15 (18 isolates)	Yes	Yes	Yes/no	Coleman (2005)
Fungi						
A	*Saccharomyces sensu stricto*	1 (7 spp.)	Yes	No	No	Naumov et al. (2000)
Plants						
B + A	*Coreopsis* W NAmer	11 (12 spp.)	Yes	Yes	No	Archibald et al. (2005)
A	E NAmer Palmatae	1 (4 spp.)	Yes	No	No	Kim et al. (1999)
A	Coreopsis	1 (6 spp.)	Yes	Yes	No	Kim et al. (1999)
A	Calliopsis	1 (3 spp.)	Yes	Yes	No	Kim et al. (1999)
A	*Corylus*	2 (7 spp.)	Yes	Yes	No	Erdogan and Mehlenbacher (2000)
A	*Notofagus* by pollen group	4 (22 spp.)	Yes	Yes	No	Manos (1997)
A	Silverswords/Tarweeds 3 genera	1 (35 isolates)	Yes	No	No	Baldwin et al. (1991)
A	Maloideae	1–2 (16 genera)	Yes/no	No	No	Robertson et al. (1991)
Animals						
A	*Haliotis*	2 (8 spp.)	Yes	Yes	No	Coleman and Vacquier (2002)
A	*Platygyra* 7 spp.	1 (9 isolates)	Yes	No	No	Miller and Babcock (1997)
A	*Drosophila*	1 (6 spp.)	Yes	Yes	No	Young and Coleman (2004)
A	European *Calopteryix*	1 (6 spp.)	Yes	Yes	No	Weekers et al. (2001)
A	*Ixodes dammini* *pacificus* *scapularis*	1 (17 isolates)	Yes	Yes/no	No	Wesson et al. (1993)

Type—see text. # Biological species—number of sexually isolated clades among the cited spp./genera. Parentheses indicate the number of isolates/species falling within these biological species. Identity of 30 nt motif—at least 30 contiguous nucleotides in the highly conserved region of the 5′-side of ITS2 helix III are identical among interbreeders. 5′ helix III identity—identity of all nucleotides on the 5′-side of ITS2 helix III among interbreeders. All ITS2 identical—entire sequence of ITS2 is identical in all examples of interbreeders examined. Reference—key to the literature on ITS and sexuality. Yes/no—identity in some interbreeding groups, but not in others.

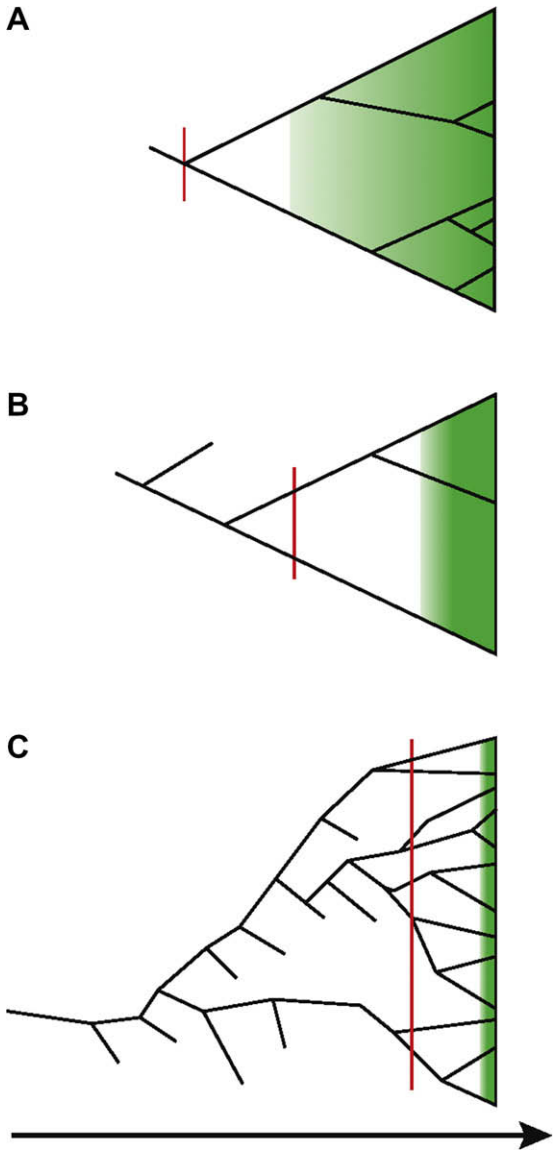

Fig. 2. Diagrammatic attempt to illustrate the relationship between interbreeding capability and current taxonomic classification. Termini represent taxonomic species (or in C, sometimes mating groups). There are three examplar types: (A) The situation seen in Fig. 1 where even some genera can hybridize; (B) where a species corresponds with the limits of sexual compatibility; and in (C), a single taxonomic species can contain several sexually isolated clades. Superimposed green coloration in A–C indicates the level of biological species (intense color) diminishing to the last vestige of gamete interaction (delimiting the Z clade). Red line suggests where a CBC is first found in the lineage. Black arrow indicates time. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

still show some weak degree of interbreeding, but the presence of a CBC in this landmark nucleotide stretch marks the end of any mating potential. In practical terms, one need not resolve the entire ITS2 secondary structure. One needs just this helix III portion of secondary structure, so one can fold the critical 30 nucleotides (the longest conserved sequence in the ITS2 among all the taxa) with the next 3′ ca. 40 nucleotides to see the germane portion of helix III.

2. *Identity for the entire ITS2 correlates with significant interbreeding potential.* Only two potential exceptions are known, where organisms with identical ITS2 sequence fail to cross. The first is one pair of syngens (mating types-v.Sonneborn, 1975) of *Paramecium aurelia* (syngens 8 and 12); here the whole ITS2 is very short (167 nt) compared to other organisms, but more importantly, Aury et al. (2006) have recently discovered that a whole genome duplication just preceded the evolution of the *P. aurelia* complex of syngens, and may be linked to this event. The second is a unicellular green alga, *Chlamydomonas allensworthii*, where two pairs of isolates (LCN, LCA) fail to interbreed (Coleman et al., 2001). Here there is some suspicion of the testing completeness, since even zygotes of apparently fertile crosses have never been germinated.

6. Explanation of the parallelism

Why should the ITS 2 conserved regions tell anything about potential interbreeding? First, there is no cause and effect for this observation; the explanation lies elsewhere. It is well established that individual genetic regions have different rates of evolutionary change. For example, actin genes are very slowly evolving (practically constant across vertebrates), while introns of genes are often exceedingly rapidly evolving (sometimes unalignable between species of the same genus). These widely differing rates of evolutionary change are maintained by selection for function.

Gamete compatibility is controlled by multiple, fairly rapidly evolving genes (Swanson and Vacquier, 2002; Hamaji et al., 2008). Presumably this is driven by selection. To carry the argument to the extreme, if any gamete of any organisms could fuse with any gamete of any other organism, gamete wastage would predominate because almost all fusions would involve two organisms having many, many differences in their genetic plan for development. The zygote would die. Thus the pressure of selection on gamete compatibility genes must be intermediate in strength: such genes must evolve fast enough to ensure that by the time developmental genes of two organisms have diverged sufficiently to cause total zygote wastage, the gamete compatibility genes of these organisms should have diverged sufficiently to prevent their gametes from making a zygote.

Meanwhile, all other regions of the genome are undergoing change; among them is the ITS. The empirical observation is that after the divergence time necessary for two lineages to fail even to interact to make zygote, the 5′ thirty most conserved nucleotides of helix III of the two organisms will show at least one difference in sequence engendering a CBC between them. In fact, by the time this first CBC has appeared, the clade may encompass two or even more Z clades (Coleman, 2000). Thus the reason the ITS2 can be used to predict whether two organisms can interact to make a zygote is that both mating genes and ITS are evolving, and that by coincidence, total loss of gamete compatibility is observed before the critical region of helix III contains a CBC (Fig. 2).

7. Hybridization and polyploidy

One must remember that there are multiple copies of the nuclear ribosomal cistrons in an organism, several hundred in a tandem array on each homolog of a diploid organism and sometimes also present elsewhere in a genome. Most organisms display a remarkably high degree of similarity among all the tandem repeats of RNA cistrons, although in no case have all the repeats been sequenced, for technical reasons. Perhaps the most exhaustive case is a haplont alga *Chlamydomonas reinhardtii*, where 200–300 repeats occur on at least two different chromosomes (Merchant et al., 2007). Fundamentally there are only two cistron variants in the ITS of this alga, differing by a 10 nucleotide indel at the tip of helix I. Extensive subcloning and sequencing of the standard strains (>150 subclones) have revealed only two other nucleotide sites in ITS2 that vary, both transitions and both in single stranded regions of the

secondary structure (Pröschold et al., 2005). This degree of homogenization seems fairly typical of most organisms, and has led to the postulation of some fairly rapid mechanism of homogenization of repeats, perhaps involving unequal crossing over or copy choice repair (Elder and Turner, 1995). It is this process that allows ITS2 in the vast majority of cases to be treated as effectively a single locus.

A number of investigators, particularly students of plant phylogenetics, have decried the use of ITS2 because of the potential for hybridization and polyploidy to mislead analysis (Alverez and Wendel, 2003; Bailey et al., 2003). In hybrids and allopolyploids, one can sometimes obtain an almost unreadable result upon sequencing the direct mixture of ITS2 PCR products. Subcloning of the PCR products instead produces the individual variant sequences present in the genome. In the cases where multiple ITS sequences of significant variance are found within an individual, it is clear that present or recent-past hybridization has occurred, and that some stage of homogenization of ITS is present. This, in itself, is very informative, and such exceptional cases only add to the value of ITS2 analyses (e.g., Feliner and Rossello, 2007).

8. What is the value of ITS2 analysis?

To the biochemist analyzing ribosomal RNA production, it should help to highlight the ITS2 regions of greatest importance to processing into the final rRNAs.

To the physiologist attempting to analyze interbreeding compatibility in a group of organisms, it provides a method of predicting which pairings are the most likely to succeed, hence most deserving of initial attention.

To the student of evolution, it can offer an opportunity for comparative analysis of the tempo of species evolution, in terms of how long is estimated to be required (based on ITS divergence) for the appearance of absolute species sexual incompatibility. This is possible perhaps only for groups with a reasonable geological record or its counterpart. For example, the estimated initiation of divergence of the silverswords on the Hawaiian islands is 5 mya (Baldwin and Sanderson, 1998).

Finally, to the taxonomist evaluating the significance of morphological or other traits to assigning taxonomic categories, it provides some guidance to sexual capabilities within the group in question. Where a CBC in the critical 30 nucleotide section of helix III has appeared, the clade of organisms defined by that CBC very likely contains at least one, and perhaps a very small number of Z clades, several "biological species", and one or more morphological species. This application has already been utilized in recent papers (Amato et al., 2007; Behnke et al., 2004; Müller et al., 2007).

The choice of what to include/exclude in a (taxonomic) species is still the decision of the taxonomist, but some sense of the "biological species" limits in a clade can help to guide such decisions. Conversely, interesting predictions can be made for related species lacking ITS2 data; for example, one can confidently predict that ITS2 of the horse family, Equidae, should be identical for essentially all of the nucleotide positions on the 5′-side of helix III, since they can interbreed. We will never alter our taxonomic system to conform to biological species analyses, but the recognition that a pan-eukaryote DNA indicator of the degree of potential interbreeding exists, whatever the applied taxonomy, should help our search for order in understanding the living world.

Acknowledgments

The author gratefully acknowledges numerous excellent suggestions from two referees.

References

Alverez, I., Wendel, J.R., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29, 417–434.
Amato, A., Kooistra, W.H.C.F., Ghiron, J.H., Mann, D.G., Pröschold, T., Montresor, M., 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. Protist 158, 193–207.
Archibald, J.K., Mort, M.E., Crawford, D.J., Kelly, J.K., 2005. Life history affects the evolution of reproductive isolation among species of *Coreopsis* (Asteraceae). Evolution 59, 2362–2369.
Aury, J-M., Jaillon, O., Duret, L., Wincker, P., 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. Nature 444, 171–178.
Bailey, C.D., Carr, T.G., Harris, S.A., Hughes, C.E., 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol. Phylogenet. Evol. 29, 435–455.
Baldwin, B.G., Kyhos, D.W., Dvorak, J., Carr, G.D., 1991. Chloroplast DNA evidence for a North American origin of the Hawaiian silversword alliance (Asteraceae). Proc. Natl. Acad. Sci. USA 88, 1840–1843.
Baldwin, B.G., Sanderson, M.J., 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc. Natl. Acad. Sci. USA 95, 9402–9406.
Behnke, A., Friedl, T., Chepurnov, V.A., Mann, D.G., 2004. Reproductive compatibility and rDNA sequences analyses in the *Sellaphora pupula* species complex (Bacillariophyta). J. Phycol. 40, 193–208.
Coleman, A.W., 2000. The significance of a coincidence between evolutionary landmarks found in mating affinity and a DNA sequence. Protist 151, 1–9.
Coleman, A.W., 2001. Biogeography and speciation in the *Pandorina/Volvulina* (Chlorophyta) superclade. J. Phycol. 37, 836–851.
Coleman, A.W., 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends Genet. 19, 370–375.
Coleman, A.W., 2005. *Paramecium aurelia* revisited. J. Eukaryotic Microbiol. 52, 68–77.
Coleman, A.W., 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucleic Acids Res. 35, 3322–3329.
Coleman, A.W., Jaenicke, L., Starr, R.C., 2001. Genetics and sexual behavior of the pheromone producer, *Chlamydomonas allensworthii* (Chlorophyceae). J. Phycol. 37, 345–349.
Coleman, A.W., Vacquier, V.D., 2002. Exploring the phylogenetic utility of ITS sequences for animals: a test case for abalone (*Haliotis*). J. Mol. Evol. 54, 246–257.
Cote, C.A., Greer, C.L., Peculis, B.A., 2002. Dynamic conformational model for the role of ITS2 in pre-rRNA processing in yeast. RNA 8, 786–797.
Coyer, J.A., Smih, G.J., Andersen, R.A., 2001. Evolution of *Macrocystis* spp. (Phaeophyceae) as determined by ITS1 and ITS2 sequences. J. Phycol 37, 574–585.
Denboh, T., Ichimura, T., Hendrayanti, D., Coleman, A.W., 2003. *Closterium moniliferum-ehrenbergii* (Charophyceae, Chlorophyta) species complex viewed from the 1506 group I intron and ITS2 of nuclear rDNA. J. Phycol. 39, 960–977.
Dobzhansky, T., 1940. Speciation as a stage in evolutionary divergence. Amer. Naturalist 74, 312–321.
Elder, J.R., Turner, B.J., 1995. Concerted evoution of repetitive DNA sequences in eukaryotes. Q. Rev. Biol. 70, 297–320.
Erdogan, V., Mehlenbacher, S.A., 2000. Phylogenetic relationships of *Corylus* species (Betulaceae) based on nuclear ribosomal DNA ITS region and chloroplast matK gene sequences. Syst. Bot. 25, 727–737.
Fabry, S., Kohler, A., Coleman, A.W., 1999. Intraspecies analysis: comparison of ITS sequence data and gene intron sequence data with breeding data for a worldwide collection of *Gonium pectorale*. J. Mol. Evol. 48, 94–101.
Feliner, G.N., Rossello, J.A., 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Mol. Phylogenet. Evol. 44, 911–919.
Goff, L.J., Ashen, J., Moon, D., 1997. The evolution of parasites from their hosts: a case study in the parasitic red algae. Evolution 51, 1068–1078.
Gutell, R.R., Larsen, N., Woese, C.R., 1994. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Rev. 58, 10–26.
Hamaji, T., Ferris, P.J., Coleman, A.W., Waffenschmidt, S., Takahashi, F., Nishii, I., Nozaki, H., 2008. Identification of the minus-dominance gene ortholog in the mating-type locus of *Gonium pectorale*. Genetics 178, 283–294.
Hershkovitz, M.A., Zimmer, E.A., 1996. Conservation patterns in angiosperm rDNA ITS2 sequences. Nucleic Acids Res. 24, 2857–2868.
Kim, S.-C., Crawford, D.J., Tadesse, M., Berbee, M., Ganders, F.R., Pirseyedi, M., Esselman, E.J., 1999. ITS sequences and phylogenetic relationships in *Bidens* and *Coreopsis* (Asteraceae). Syst. Bot. 24, 480–493.
Lane, C.E., Lindstrom, S.C., Saunders, G.W., 2007. A molecular assessment of northeast Pacific *Alaria* species (Laminariales, Phaeophyceae) with reference to the utility of DNA barcoding. Mol. Phylogenet. Evol. 44, 634–648.
Mai, J.C., Coleman, A.W., 1997. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J. Mol. Evol. 44, 258–271.
Manos, P.W., 1997. Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): taxonomic congruence with morphology and plastid sequences. Amer. J. Bot. 84, 1137–1155.
Merchant, S. et al., 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science 318, 245–250.
Miller, K., Babcock, R., 1997. Conflicting morphological and reproductive species boundaries in the coral genus *Platygyra*. Biol. Bull. 192, 98–110.

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

Naumov, G.I., James, S.A., Naumova, E.S., Louis, E.J., Roberts, I.N., 2000. Three new species in the *Saccharomyces sensu stricto* complex: *Saccharomyces cariocanus*, *Saccharomyces kudriavzevii* and *Saccharomyces mikatae*. Int. J. Syst. Evol. Microbiol. 50, 1931–1942.

Pröschold, T., Harris, E.H., Coleman, A.W., 2005. Portrait of a species: *Chlamydomonas reinhardtii*. Genetics 170, 1601–1610.

Robertson, K.R., Phipps, J.B., Rohrer, J.R., Smith, P.G., 1991. A synopsis of genera in Maloideae (Rosaceae). Syst. Bot. 16, 376–394.

Schultz, J., Müller, T., Achtziger, M., Seibel, P.N., Dandekar, T., Wolf, M., 2006. The internal transcribed spacer 2 database—a web server for (not only) low level phylogenetic analyses. Nucleic Acids Res. 34, W704–W707.

Sonneborn, T.M., 1975. The *Paramecium aurelia* complex of fourteen sibling species. Trans. Amer. Microsc. Soc. 94, 155–178.

Stache-Crain, B., Müller, D.G., Goff, L.J., 1997. Molecular systematics of *Ectocarpus* and *Kuckuckia* (Ectocarpales, Phaeophyceae) inferred from phylogenetic analysis of nuclear- and plastid-encoded DNA sequences. J. Phycol. 33, 152–168.

Swanson, W.J., Vacquier, V.D., 2002. The rapid evolution of reproductive proteins. Nat. Rev. Genet. 3, 137–144.

Weekers, P.H.H., DeJonckheere, J.F., Dumont, H.J., 2001. Phylogenetic relationships inferred from ribosomal ITS sequences and biogeographic patterns in representatives of the genus *Calopteryx* (Insecta: Odonata) of the West Mediterranean and adjacent West European Zone. Mol. Phylogenet. Evol. 20, 89–99.

Wesson, D.M., McLain, D.K., Oliver, J.H., Piesman, J., Collins, F.H., 1993. Investigation of the validity of species status of *Ixodes dammini* (Acari: Ixodidae) using rDNA. Proc. Natl. Acad. Sci. USA 90, 10221–10225.

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

Woolcott, G.W., Iima, M., King, R.J., 2000. Speciation within *Blidingia minima* (Chlorophyta) in Japan: evidence from morphology, ontogeny, and analyses of nuclear rDNA ITS sequence. J. Phycol. 36, 227–236.

Yotsukura, N., Denboh, T., Motomura, T., Horiguchi, T., Coleman, A.W., Ichimura, T., 1999. Little divergence in ribosomal DNA internal transcribed spacer-1 and -2 sequences among non-digitate species of *Laminaria* (Phaeophyceae) from Hokkaido. Japan. Phycol. Res. 47, 71–80.

Young, I., Coleman, A.W., 2004. The advantages of the ITS2 region of the nuclear rDNA cistron for analysis of phylogenetic relationships of insects: a *Drosophila* example. Mol. Phylogenet. Evol. 30, 236–242.