CHAPTER 1

LITERATURE REVIEW:

USING PROTEIN-CODING GENES IN FUNGAL TAXONOMY

USING PROTEIN-CODING GENES IN FUNGAL TAXONOMY

1 INTRODUCTION

Since the onset of the application of molecular techniques, fungal taxonomy and phylogeny has been dominated by the use of the ribosomal RNA genes (*rrn*) 18S, 28S and 5.8S, as well as the internal transcribed spacers (ITS) separating these genes (see for example refs. 35, 119). Depending on the particular *rrn* gene used, taxonomic and phylogenetic questions at all levels have been addressed (see of example ref. 247). Unfortunately, phylogenetic trees inferred using these genes are often incongruent with fungal biology and they do not always provide sufficient resolution of the taxa being studied (see for example refs. 115, 170, 247, 254, 291, 293). The major reason for these irregularities and lack of resolution is that non-uniform evolutionary forces are potentially acting upon the *rrn* genes of closely related fungi. This problem is easily solved by including additional regions of the fungal genome in the analyses. For this purpose, fungal taxonomists and evolutionary biologists frequently use protein-coding genes (see for example refs. 212, 278).

Protein-coding genes can be applied to evolutionary questions at all taxonomic levels. They have, for example, been used to determine the root of the tree of life and to study the ancient eukaryotes (10, 77, 144, 149, 291). Protein-coding genes can also be applied successfully to lower (intra- and interspecies) and intermediate (intergenus or -order) taxonomic levels (18, 155, 170, 213). The use of protein-coding genes in phylogenetic studies has one major advantage over the use of *rrn* genes. Whereas only specific *rrn* genes can be used to address questions at certain levels, a single protein-coding gene, for example any one of the tubulin genes, can be used to address taxonomic questions at all levels (see for example refs. 11, 141, 214, 243).

Protein-coding genes are subjected to many different evolutionary forces, the effects of which can have profound implications on the interpretation of phylogenetic data. The purpose of this review is, partially, to discuss different forces acting on protein-coding genes and how they influence evolutionary reconstructions. This is done by providing background, firstly, on the structure of protein-coding genes and secondly, on the evolutionary forces that have shaped them. The discussion on protein-coding genes is mostly restricted to eukaryotic nuclear genes, but organellar and prokaryotic genes are briefly considered. The remainder of this review deals with the use of protein-coding genes in fungal taxonomy and provides some examples where they have been used successfully.

2 EXON-INTRON ORGANIZATION OF PROTEIN-CODING GENES

One of the most striking features of protein-coding genes is that they are generally organized into regions of coding sequences (exons) that are interrupted by intervening non-coding sequences (introns) (reviewed in ref. 31). Since these intervening sequences confer no apparent phenotype on the cell, and are thought to be without function, they are subjected to fewer evolutionary constraints than are exons (24, 65). Introns thus provide an attractive source of sequence variation in an otherwise highly conserved gene (6). In the following sections, introns are discussed with regard to their types, possible role in biology, origin and positional conservation. In addition phenomena specific to exons are reviewed. These include codon bias, G+C content, multiple overlapping substitutions or homoplasy and transversion/transition ratios. In all cases, special attention is given to issues pertaining to the evolutionary forces that not only gave rise to these regions, but also those that are currently acting upon them.

2.1 Introns

2.1.1 Types of introns

Introns are divided into different classes based on their mechanisms of splicing and in which genes they occur. Group I introns, for example, are found in the genomes of prokaryotes and organellar genomes. They have ribozymic activity and hence encode endonucleases that assist their splicing and trans-positioning (17, 72). Group II introns are also autocatalytic or self-splicing, but rather than encoding endonucleases, they generally encode reverse transcriptases (17). Group II introns are characteristic of prokaryotic and organellar genomes, where they are found in the genes encoding proteins, transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) (194). Group III introns are similar to group II intron, but differ in that their self-splicing mechanisms are somewhat defective (50, 194). Both group I and II introns are self-splicing and thus code for the 'machinery' necessary for their spread and/or removal.

Eukaryotic nuclear genes harbor a different type of intron that is generally referred to as a spliceosomal intron. Removal or splicing of this type of intron involves a multi-molecular RNA/protein complex or spliceosome that is formed through the interaction of a number of small nuclear ribonucleoproteins (snRNPs) (88, 104, 177). The structure and splicing mechanism of spliceosomal introns closely resembles that of group II introns, which suggests a common origin for spliceosomal and group II introns (43, 50, 65, 137, 160, 205, 236).

Spliceosomal introns are subdivided into two groups based on the consensus sequences at their splice sites (136, 200, 249, 289). The first and most common type is characterized by GT- and AG-dinucleotides at their 5' and 3' intron boundaries (Table 1) (289). Two types of spliceosomes can excise these GT-AG introns (Table 1). The most common type includes U2 snRNP as part of

the spliceosomal apparatus, whereas the other spliceosome include U12 snRNP (110, 136, 249, 289). The second type of spliceosomal intron has the dinucleotides AT and AC at its 5' and 3' boundaries (110, 200). The U2- and U12 types of spliceosomes also excise these AT-AC introns (Table 1). These AT-AC introns are, however, very rare. U12-type AT-AC introns occur at a frequency of less than one in five thousand, and U2-type AT-AC introns are even rarer, since there are only seven known examples (110, 111, 200, 269).

Spliceosomal introns are distinguished based on their phases. Intron phase refers to the placement of the intron relative to the reading frame (58). A phase 0 intron, for example, is situated exactly between two neighboring codons. Phase 1 and 2 introns split a codon after the first and second bases, respectively. Long and Deutsch (173) reported that most spliceosomal introns were of the phase 0-type. They further showed a strong correlation between intron phase and the degree of conservation of the splice signals in the exons surrounding an intron. In other words, the exonic sequences flanking phase 2 introns are most variable and are not in total agreement with the consensus sequences (Table 1). The exonic sequences surrounding a phase 0 intron usually match the consensus exactly. According to Long and Deutsch (173), the evolutionary forces determining the sequences surrounding a spliceosomal intron are strongly biased towards generating phase 0 introns. This is because variations in these splice signals produce phase 1 and phase 2 introns, which may lead to intron-loss associated with deleterious mutations. The relatively few phase 1 and 2 introns, therefore, reflect those cases where intron-losses were not associated with lethal mutations (173). This is illustrated in the β-tubulin genes of a diverse group of organisms, where more than half the introns are of the phase 0-type, while the remaining introns are phase 1 and 2 (Fig. 1).

2.1.2 The role of introns in biology

Broadly speaking introns have no obvious function (24). They can be removed without any phenotypic effect. There are, however, reports of certain introns performing regulatory functions such as enhancing or modulating expression of the genes harboring them (78, 255). Several authors have further indicated that introns might play a role in genetic recombination by creating so-called 'hot spots' for crossing-over (22, 23, 25, 65, 96, 97). This 'loosening' of genetic linkage between exons, forms the basis for one of the models explaining the evolution of introns. Nevertheless, if introns do play biologically important roles, they would be subjected to evolutionary forces that are most probably different to those acting on the exon regions.

2.1.3 Origin of introns

There are two opposing hypotheses for the origin of introns. These are known as the 'introns-early' and 'introns-late' theories (65). The 'introns-early' theory is also known as 'the exon

theory of genes' and suggests that the first exons were short (15-20 amino acid residues) and assembled by recombination within introns. According to this theory, the short exons (encoding functionally active polypeptides) were 'shuffled' to eventually form a gene consisting of many introns and exons. This suggests that primitive ancestral genes would have had many introns that separated various functionally active protein domains or exons. Divergence of prokaryotes and eukaryotes led to the loss of these introns through genomic streamlining in the case of prokaryotes, and retention associated with occasional loss of introns in the case of eukaryotes (65, 96, 98, 168, 185, 229, 248, 281).

Several lines of evidence are provided for the 'introns-early' theory (24, 64, 96, 98, 99, 168, 211). One of these is the relative position of introns in genes (168, 211). As is the case in many other genes (9, 79, 98, 185, 239, 281) the β-tubulin introns occur in clusters at regular intervals of 15-20 amino acid residues (Fig. 1) (168, 211). The positions of these introns also appear to be conserved over great evolutionary distances, suggesting that modern genes evolved via intron-loss from the ancestral intron-containing gene (Table 1).

The 'introns-late' theory suggests that prokaryotic genes most closely resemble the ancestral state (58, 59, 65, 172, 236). According to this model, spliceosomal introns evolved during eukaryotic evolution. These introns were then inserted into unsplit genes after the prokaryotic-eukaryotic 'transition' and have nothing to do with the development of genes (236). The 'introns-late' theory is supported by the presence of so-called proto-splice sites (172). These sites serve as recognition sequences for the insertion of introns (58). Proto-splice sites are characterized by the sequence KAG*R (K = A or C; R = A or G and * = splice junction), which closely resemble the exonic consensus surrounding intron splice junctions (Table 1). Such proto-splice sites are present at conserved positions in what is believed to be older intron-lacking versions of a gene. The 'introns-late' theory suggests that these proto-splice sites were present prior to evolutionary radiation (58, 172). The conserved positions of the proto-splice sites, therefore, determine the positions of introns (58, 172).

The origin of introns remains a controversial subject. Evidence for both the 'introns-late' and 'introns-early' models is inconclusive. Supporters of the 'introns-early' model interpret the presence of proto-splice sites as remnants of lost ancestral introns. Proponents of the 'introns-late' model find no statistically significant 'intron-splitting' of genes into functionally active domains (59, 277). They also argue that it is unlikely that large-scale intron-loss through genetic streamlining occurred twice in evolutionary history, once in the ancestor of all eubacteria and a second time in the archaebacterial ancestor (236). It is thus clear that neither the 'introns-late' nor the 'introns-early' theories can be discredited. Recent computer assisted analyses of large numbers of conserved protein-coding genes has provided evidence that supports both models (238, 277). Some introns are

evolutionarily old and were probably involved in ancestral exon shuffling to create genes (82, 248, 281), whereas other introns were recently gained or lost (82, 172).

The opposing theories on the origin of introns complicates their use in phylogenetic inferences. This is because evolutionary reconstructions using protein-coding genes reflect the phylogenies of both the exons and introns. If the introns in question originated prior to the divergence of the group of organisms included in the analyses, the intronic history would reflect the phylogeny of the exonic regions. However, if these introns were 'acquired' during or after the divergence of the organisms in question, the intronic and exonic phylogenies would be different.

2.1.4 Positional conservation

The nucleotide sequences of many genes display a higher degree of variability in the 5'-half, than in the 3'-half, of the gene. This is mainly due to the fact that introns are not distributed uniformly within genes, but appear to be more abundant in the first or 5'-half. Examples of genes where introns are preferentially situated in the 5'-half, are α- and β-tubulin (Fig. 1) (168), small G proteins (59), glyceraldehyde-3-phosphate dehydrogenase (185, 228), etc. This is either because of intron gain in the first portion of genes, or intron loss from the 3'- half of genes (58, 172, 236, 238).

Gain and loss of introns is explained using the 'introns reinsertion-homologous recombination' model. This model involves the spliceosome and proto-splice sites, mentioned earlier. Based on this model the loss or gain of an intron includes a four-step process (96, 168, 236), whereby the intron is (i) spliced from the premature messenger RNA (pre-mRNA) and (ii) reinserted into a nearby site of the pre-mRNA. This is followed by (iii) reverse transcription of the 'new' pre-mRNA into complimentary DNA (cDNA). The final step in this model is the (iv) reinsertion of this newly formed intron-containing cDNA into the genome via homologous recombination with the genomic copy. The result would be intron gain, but when step two is omitted, the result would be intron-loss. In other words, if the removed intron is not reinserted at a different position in the pre-mRNA, the pre-mRNA can be reverse transcribed and recombined back into the genome. This would generate a gene from which an intron has been deleted.

The involvement of a reverse transcription step in the 'intron reinsertion-homologous recombination' model provides an explanation for the polarization of genes with regard to intron position. According to this model, reverse transcription is always initiated at the 3'-end and seldom extends fully to the 5'-end of the pre-mRNA. Homologous recombination, therefore, results in 'replacement' of the 3'-portions of the gene with a reverse transcribed intron-less cDNA copy (85, 96, 185). In some genes, the involvement of reverse transcriptional errors at the 3' end of the gene has been used to explain why 5' intron positions are more conserved than the 3' intron positions (185, 228).

Intron positions in genes such as β-tubulin, are usually conserved across great evolutionary distances (168). Comparisons of intron positions in the β-tubulin genes from a diverse group of organisms (Fig. 1) has revealed that ascomycetous fungi are characterized by a unique intron (intron 5). The same is also true for the metazoan (intron 20) and plant lineages (intron 133) (Fig. 1). This conservation also extends to individual lineages, since all the fungi from the pyrenomycetous order Hypocreales have three specific introns in the first half of their β-tubulin genes (introns 5, 13 and 54). All the members of the lower plants also appear to harbor a unique intron (intron 57) in their β-tubulin genes. However, certain organisms have more than one β-tubulin gene per individual, each with unique intron positions. For example, the intron positions in the *tubC* and *benA* β-tubulin genes of *Aspergillus nidulans* are very different (Fig. 1).

Intron sequences are useful for answering phylogenetic questions at lower taxonomic levels. This is because it is possible to align homologous intron sequences from closely related taxa (6, 18, 63, 83, 212, 263, 285). However, 'homologous' introns of more divergent taxa usually only share the same position and little or no sequence homology (83, 172, 263, 280). For this reason the use of intron sequences for reconstructing deeper level phylogenies are not feasible (83, 158, 191, 310). To address this type of evolutionary question intronic regions can thus not be treated as nucleotide bases in phylogenetic analyses, but rather as 'presence' or 'absence' characters.

2.2 Exons

Exons are the coding regions of protein-coding genes. Their nucleotide base composition is, therefore, subjected to evolutionary forces that not only reflect lineage history, but also other constraints imposed at the translational and functional levels (30, 49, 131, 132, 292). Phenomena such as codon bias, G+C content, multiple overlapping substitutions and transversion/transition ratios serve as indicators of these selective forces, although they may in some cases also reflect gene and/or taxon phylogeny (1, 2, 30, 34, 55, 91, 130, 166, 171, 182, 235, 275, 298). These phenomena are not independent of one another and changes in one results in changes in the others. In the following sections, these interdependent factors are discussed in more detail.

2.2.1 Codon Bias

Codon bias is a phenomenon found in many protein-coding genes and is defined as the non-random use of synonymous codons (106). Leucine, for example, is encoded by six synonymous codons (CUU, CUC, CUA, CUG, UUA and UUG) and the preferential use of one during translation is referred to as codon bias. 'Non-biased' genes differ from biased genes, in that any of these codons can be used during translation. 'Non-biased' genes are further characterized by silent or synonymous substitutions rather than non-synonymous substitutions that would result in alteration of the amino acid sequence (36, 45, 117, 166, 292). For example, one or two substitutions in the

three nucleotide bases specifying an amino acid will generally be synonymous or silent. This is because the substitution results in a codon that still encodes the original amino acid. Non-synonymous substitutions that occur at the third nucleotide base are less frequent, since they will result in alteration of an amino acid residue that can lead to loss or decrease of functionality in the mature protein.

Codon bias is a prominent feature of highly expressed genes (45, 131, 251). It is determined and/or influenced by two main groups of factors. The first group of factors controls the efficiency of translation. The second group of factors control structural aspects of the gene without regard of translational efficiency (292). One of the factors that will influence translational efficiency is the abundance of a specific tRNA species (30, 49, 61, 103, 117, 130-132, 166, 225, 251, 289). In the yeast *Saccharomyces cerevisiae*, for example, the most abundant lysine tRNA species has the anticodon CUU, which will bias the codon usage of this amino acid towards AAG. Other tRNA species that will recognize the remaining lysine codon (AAA) are scarce and sometimes absent. The inclusion of this codon in the genes of the fungus will thus cause a reduction in translational efficiency. Such codons are, therefore, selected against and results in codon bias.

The second group of factors that will influence codon usage includes the requirements for gene and RNA secondary structure (224, 253). For example, portions of the downstream-untranslated regions of the alcohol dehydrogenase (*adh*) gene of *Drosophila*, interacts with nucleotides in the second exon of this gene. Synonymous substitutions (i.e. those that will not change the codon) in this exon, alter the secondary structure of the mRNA. This significantly reduces expression of the *adh* gene (224). This type of interaction, therefore, also contributes to the selection for specific codons.

Preferential use of certain codons is generally species and/or gene specific (1, 91, 159, 166), but is sometimes associated with phylogeny (106, 131, 132). Differences in the degree of codon bias can potentially have serious implications for determining evolutionary relationships. It is well documented that the use of many codon-biased genes distorts and obscures phylogenetic histories in many different eukaryotic and prokaryotic organisms (55, 81, 117, 125, 166, 167, 176, 182, 237, 292, 298).

2.2.2 G+C content

Variations in G+C content are usually located at the third bases of codons (30, 55, 166, 250, 292). This reduces the number of possible synonymous substitutions in a codon. For example, a bias towards high G+C content will reduce the number of codons (six) specifying a leucine residue to three codons. The codons CUU, CUA and UUA will be selected against and thus not occur. It is clear that G+C content and codon bias are very closely linked. For this reason the same

evolutionary forces that affect codon bias will generally also influence G+C content (36, 55, 81, 125, 166, 292).

2.2.3 Multiple overlapping substitutions

Apart from the restrictive effects of codon bias and G+C content, a multiplicity of silent or synonymous substitutions can occur at a specific position during the evolutionary history of a gene or species (167). For example, the third base in the codon specifying a leucine, can change from T to C and back to T. This is termed a 'reversal' and together with other phenomena such as convergence and parallelism, is referred to as homoplasy (268). Many of these overlapping substitutions or homoplastic events will obliterate the historical information at that position. They can thus lead to an underestimation of the degree of divergence or an overestimation of the degree of similarity between different taxa (100, 237, 294). The use of these characters in evolutionary inferences, therefore, results in lack of phylogenetic resolution and inconsistencies (125, 233, 237, 294).

2.2.4 Transversion/transition ratio

A transversion is defined as the substitution of a pyrimidine (C or T) for a purine (A or G) or vice versa. Transition is defined as substitution of a purine for a purine or a pyrimidine for a pyrimidine (167). According to DeSalle et al. (57), there is a general lack of transitional bias between distantly related taxa, because the record of transitional events is erased by transversions. This apparently results in an accumulation of transversions among more divergent genomes (34, 57, 124, 275). Transitions are, therefore, usually more abundant than transversions among closely related organisms (34, 55, 57, 117, 125, 167, 275).

Inference of phylogenetic relationships from protein-coding regions presupposes that evolutionary forces underlying nucleotide variation are common to all the taxa that are examined (167). However, many different selective forces and processes, other than those involved in the 'creation' of a lineage, are acting upon the exons and introns of protein-coding genes. These processes can distort phylogenetic information, thereby obscuring evolutionary histories and making it impossible to reconstruct genealogical relationships. Although these forces act on genes at all taxonomic levels, most problems are encountered at the deeper levels such as kingdom, family and order (10, 94, 123, 125, 166, 265, 292). Among closely related taxa at the species level a specific gene or part of a gene is generally subjected to comparable forces (55, 117, 166, 292).

3 SINGLE AND MULTI-COPY GENES

Evolutionary analyses are based on the assumption that the selected gene or region of the genome is orthologous (167). In other words, the evolutionary history for this region is similar to that of the individuals in which it is studied. Events that will create multiple copies of genes (duplication, hybridization and horizontal transfer) may result in non-orthology between genes. Protein coding genes that occur in multigene families and those that occur as single genes in the genome of an organism are further subjected to different evolutionary forces (217). The non-uniform evolutionary forces acting on single and multi-copy protein-coding genes greatly complicate their use in phylogenetic studies, since a gene phylogeny can be inaccurately interpreted as a species phylogeny.

3.1 Multi-copy genes

3.1.1 Types of multigene families

In evolutionary biology, the duplication events that gave rise to multigene families could have occurred very early or relatively recently. These ancient and recent duplication events are reflected in the degree of divergence from the ancestral gene. For example, modern gene families share a high degree of sequence homology, while ancient gene families show very little sequence homology. In many cases, the homology in ancient gene families will be restricted to a number of conserved domains, which is the result of functional constraint in the mature protein (28, 133). Examples of ancient multigene families are those that encode the different tubulin subunits of mature microtubules (174, 178) and those encoding the subunits of eukaryotic DNA-dependent RNA polymerases (133). Some of these ancient gene families are divided into subfamilies to form modern gene families.

Ancient multigene families are potentially of great use in evolutionary studies, especially for the inference of deep phylogenetic relationships. For this type of study, eukaryotic protein-coding genes are normally used. However, this approach is problematic, since suitable outgroups are not always available (107, 144). The problem can be overcome by using the gene sequence of another member of that ancient multigene family as an outgroup. This is, however, only possible when the duplication event that generated the gene family predated the divergence of the taxa of interest. This approach has been successfully employed by several research groups (107, 135, 144, 254).

Modern multigene families include those encoding β-tubulins (46), chitin synthases (29), actins (190), etc. The different members of some of these gene families are thought to be essential at different stages of the life cycles of organisms (38, 46, 80, 128, 129, 153, 190, 246, 259, 284). This is, however, not always the case, since disruption of genes in these families does not always

result in lethal mutations (38, 60, 266). It is suggested that the occurrence of more than one copy of a specific gene act as a form of multigene control of a specific trait, since another member of the family can 'replace' a defective copy (16, 301). The *rrn* genes also belong to the latter class of multigene families, despite the fact that they do not encode proteins.

Members of modern multigene families can occur clustered at a specific locus on a chromosome. For example, the genes encoding α- and β-tubulin in some plants and metazoans are organized as tandem repeats on a chromosome (46, 231). They can also occur at multiple loci on more than one chromosome (13, 16, 46, 190, 303). Multigene families can, however, also consist of individual genes scattered across the genome. Examples of these are the genes encoding chitin synthases in oomycetous fungi (199) and actins in mammals (190).

3.1.2 Concerted evolution

Large-scale sequence analyses of the repeated genes constituting multigene families, have revealed that the members of a repeat, share more similarities within a species, than between species. This species-specific homogeneity is generated by a process known as 'concerted evolution' or 'molecular drive' (66, 311). There are several mechanisms through which this process can take place e.g. unequal crossing-over, gene conversion, homologous recombination, transposition and replication slippage. Gene conversion and unequal crossing-over are considered the most important of these mechanisms.

Concerted evolution is best explained by Sanderson and Doyle (240) using a simple gene family consisting of two members (X and Z) in each of four species (1, 2, 3, and 4) (Fig. 2A). X and Z are paralogous genes that originated from a duplication event, prior to the radiation of species 1 - 4. The X genes in all the individuals are orthologous and trace their ancestry to a speciation event. The same is also true for the Z genes. Phylogenetic reconstructions using either the X or the Z genes will thus reflect the organismal evolution. However, in most cases it is impossible to differentiate between homologous genes of paralogous and orthologous origins. In the absence of, or prior to, homogenization (Fig. 2B), all the X genes will form a cohesive cluster, as is true for Z genes. The lineage or species history within each of these clades, however, still reflect the 'correct' phylogeny as depicted in the 'true genealogy' (Fig. 2A). After concerted evolution, or when it is highly effective, all the paralogues within an individual are homogenized (Fig. 2C). In these cases, interspecies variation by far exceeds intraspecies variation. Clearly, reconstructing evolutionary relationships from genes that are subjected to high levels of homogenization and those where concerted evolution is effectively absent, are relatively straightforward (52, 293). Either orthologues will group together (Fig. 2B) or paralogues will be homogenized, but in both cases it would be possible to infer the 'correct' phylogeny. However, intermediate levels of concerted

evolution introduces major complications to the inference of evolutionary relationships (183, 184, 240).

The mechanisms through which concerted evolution take place do not exclude the potential for 'horizontal' spread of a variant member (8, 66, 67, 120, 167). This is especially true when the variant caries a beneficial mutation. This mutation can then be spread to all the other members of that family, thus illustrating how a small selective advantage can become a great advantage via concerted evolution (167). Since this type of mutation is not acquired through descent from a common origin, many authors conclude that concerted evolution conceals true phylogenetic relationships (117, 234, 240, 293). The effect of concerted evolution can be summarized most appropriately in the words of Schimenti (244) who states that "concerted evolution can wipe out millions of years of divergence" or "introduce multiple sequence changes into a member of a gene family…in a single event".

The efficacy of concerted evolution to homogenize paralogous genes throughout in the genome varies greatly. In cotton (*Gossypium* spp.), for example, different gene families are subjected to different levels of concerted evolution. Cronn et al. (52) showed that all the members of the cotton 5S *rrn* family at one locus were very similar in sequence and different from the copies at another locus. These results were in contrast to those of Wendel et al. (293) using the cotton 18S-26S *rrn* multigene family. They showed that all the sequenced copies of the 18S-26S repeat, whether from a single or more than one locus, had almost identical sequences. Variations in the degree of homogenization thus occur not only among multigene families, but also among different clusters of the same gene family (19, 52, 69, 70, 90, 114, 127, 151, 189, 234, 274, 293). Although the majority of studies on the efficacy of concerted evolution focused on the homogenization of *rrn* genes, the homogenization of protein-coding gene families is well documented (1, 47, 81, 113, 114, 117, 127, 151, 234, 272, 274).

3.1.3 Example of a protein encoded by a multigene family: β-tubulin

Inspection of the sequences in nucleotide databases such as GenBank reveals that many different protein-coding genes are currently used to address evolutionary and taxonomic issues in diverse organisms (Table 2). The most entries for fungi are those encoding β-tubulin, translation elongation factor 1α, chitin synthases, actin and glyceraldehyde-3-phosphate dehydrogenase. The nucleotide databases for the genes encoding calmodulin, the mating type idiomorphs and histone H3 are also relatively large. Although the genes encoding other proteins, such as translation elongation factor 2, HSP70 and α-tubulin are less frequently used by fungal taxonomists, they have been successfully used in many other lower eukaryotes (Table 2). However, after the *rrn* genes, the use of those encoding β-tubulin in fungi is best documented.

β-Tubulin is one of the 50-kDa subunits of the heterodimeric protein, tubulin (46, 266). Tubulin is the primary component of microtubules, which are the cytoskeletal filaments of eukaryotic cells. For this reason, they are involved in determining the shape the cell and the nucleus, as well as, as well as in cell processes such as chromosome segregation, cell division, flagellar motility, etc. (46, 266).

From agricultural and veterinary perspectives, β-tubulin is an important protein. This is mainly due to the mode of action of benzimidazole containing fungicides and anthelmintics (54). These drugs specifically bind to β-tubulin, thereby preventing the assembly of mature microtubules and result in the inhibition of DNA synthesis (54). Single point mutations in the gene encoding β-tubulin have been shown to confer resistance to benzimidazoles (7, 16, 37, 150, 220, 266, 302, 304). Although only one of the β-tubulin-coding loci is usually associated with benzimidazole sensitivity or resistance, more loci can sometimes be involved (16, 266).

β-Tubulin is usually encoded by highly conserved multigene families or in some cases single genes. In the best-studied higher plants, multigene families consisting of five to nine different β-tubulin genes have been described (108, 128, 168, 181). Similar multigene families in animals have been reported (46, 266). Some fungi also appear to have more than one divergent copy of the β-tubulin gene. *Saccharomyces cerevisiae* (202), *Candida albicans* (257), *Neurospora crassa* (220), *Schizosaccharomyces pombe* (121), *Botrytis cinerea* (304) and *Fusarium* species in the *Gibberella fujikuroi* complex (212) all appear to have a single copy of this gene. On the other hand, fungi such as *Geotrichum candidum* (101), *Aspergillus nidulans* (187), *Colletotrichum gleosporioides* f. sp. *aeschynomene* (37), *C. graminicola* (222), *Erisiphe graminis* (252), *Acremonium coenophialum* (278) and species in the *F. solani* complex (212) have at least two different copies of the β-tubulin gene.

The β-tubulin gene sequence provides an excellent tool for studying phylogenetic relationships at all taxonomic levels. This protein-coding gene has been successfully used to determine both intra- and interspecific relationships. An example where β-tubulin gene sequences have been used at the intra-species level is in populations of the sheep gut parasite, *Haemonchus contortus* (16). A well-known example where it has been used at the interspecies level is for the molecular characterization of *Fusarium* species (5, 213, 216). β-Tubulin gene sequences have also been used to address deep phylogeny questions. The best example is probably where this gene, together with three other protein-coding genes, was used to demonstrate that fungi and animals are each other's closest relatives (11).

3.2 Single copy genes

Divergence of a single copy gene and speciation are two very close linked processes. This is because divergence of an ancestral gene would coincide with speciation. A single copy gene would be more 'resistant' to mutations than members of a multigene family. A lethal mutation in one of the members of multigene family can be 'corrected' through concerted evolution. If concerted evolution fails to 'correct' the mutation, another member of the multigene family can take on the role of the mutated gene (38, 301). A lethal mutation in a single copy gene, however, results in death of the individual (202, 218). For this reason, non-lethal nucleotide changes in a single copy gene will also cause changes in the individual, thus contributing to species evolution. Therefore, a single-copy gene would theoretically provide more reliable evolutionary reconstructions than multi-copy genes (163).

Most protein-coding genes used for reconstructing phylogenetic histories, occur as multigene families (Table 2). Some of them, however, also occur as single copy genes in lower eukaryotes such as actin genes in certain algae, protozoans and oomycetous fungi (19, 68, 71). The only protein-coding genes that apparently occur 'universally' as single copies, are those encoding the largest and the second largest subunits (*RPB1* and *RPB2*) of the DNA dependent RNA polymerase II complex (56, 170, 254). This may, however, be because of under-sampling, since the copy numbers of these genes are seldom determined.

3.2.1 Examples of proteins encoded by single genes: largest and second largest subunits of DNA-dependent RNA polymerase II

RNA polymerase is thought to be one of the earliest enzymes to have appeared (161). This is consistent with the idea that RNA preceded DNA as genetic material. This ancient RNA-dependent RNA polymerase then gave rise to the modern DNA-dependent RNA and DNA polymerases (161). The eukaryotic DNA-dependent RNA polymerases are large multi-subunit enzyme complexes that are divided into three groups, i.e. RNA polymerase I, II and III (289). RNA polymerase I is responsible for transcription of the 5.8S, 18S and 28S *rrn* genes, RNA polymerase II transcribes nuclear protein-coding genes into mRNA and RNA polymerase III produces tRNA and 5S ribosomal RNA (289).

The eukaryotic DNA-dependent RNA polymerases share a common origin with the eubacterial and archaebacterial RNA polymerases (133). Because of this, the genes encoding their protein-subunits closely resemble one another. For example, the genes encoding the largest subunits of the eukaryotic RNA polymerase I, II and III are homologous to the eubacterial β' subunit (133). This homology is reflected in the nine conserved domains (I-IX) present in these prokaryotic and eukaryotic genes.

Many researchers have indicated the potential use of the genes encoding DNA-dependent RNA polymerase subunits in evolutionary studies (133, 161, 227, 254). A number of sequences for the genes encoding the two largest subunits of RNA polymerase II are available (Table 2). However, few sequences encoding the subunits of RNA polymerase I and III are available, making them less suitable for evolutionary analyses.

The largest subunit of the RNA polymerase II is encoded by the gene, *RPB1*, and the second largest subunit is encoded by *RPB2* (306). The GenBank nucleotide database for both these genes is limited compared to those for other genes (Table 2), but several successes have recently been reported on using *RPB1* and *RPB2* gene sequences for evolutionary studies (51, 56, 123, 170). Liu et al. (170) showed that *RPB2* is more useful than 18S *rrn* to resolve the relationships among the different fungal orders. Croan et al. (51) showed that *RPB1* is useful for studying interspecific relationships among *Leishmania* species. Furthermore, both these genes provide good resolution at deeper phylogenetic levels. Denton et al. (56) reconstructed the possible phylogeny of the plant kingdom (viridiplantae) using *RPB2* gene sequence and Hirt et al. (123) placed the microsporidia within the fungal kingdom using *RPB1* gene sequence.

4 PROTEIN-CODING GENES AND FUNGAL TAXONOMY

In recent years, protein-coding genes have increasingly been used to address phylogenetic and taxonomic questions at all levels. These sequences have not only proven useful at deeper (Kingdom or Division) taxonomic levels, but also at the lower (inter- and intraspecies) levels (Table 2). Several protein-coding genes contain sufficiently variable and conserved regions to allow resolution at both deeper and lower taxonomic levels (Table 2). Most of the recent advances in fungal taxonomy have, therefore, been based on sequence for protein-coding genes.

4.1 Deep level fungal taxonomy: The microsporidia-fungi relationship

Microsporidia are spore forming obligate intracellular parasites of all major animal groups (41, 44). Although they represent an eukaryotic lineage, the microsporidia share a surprising number of features with prokaryotes. These include ribosomal features such as 70S rather than 80S ribosomes and fused 5.8S and large subunit *rrn* genes (287). The microsporidian genomes also correspond with those of bacteria, as they are small and rarely harbor introns (21, 83, 141). The microsporidia also lack eukaryotic organelles such as mitochondria (44, 141). Because of this resemblance to prokaryotes, they were thought to represent eukaryotic lineages that evolved prior to the acquisition of mitochondria. They were consequently classified as Archezoa (44, 286).

The archezoan status of the microsporidia has been supported by molecular data from the

rrn genes and those encoding the translation elongation factors, EF-1α and EF-2 (139, 286). In the microsporidian lineage, however, these genes are known to display features such as biased base composition, unique insertions and deletions and accelerated rates of substitution (145). Phylogenies based on these genes were thus not reliable (145, 226, 265), which resulted in the erroneous placement of the microsporidia at the base of the eukaryotic tree (20, 75, 82, 122, 123, 144, 145).

Phylogenies based on α-, β-, and γ-tubulin gene sequences indicated that the microsporidia are closely related to the fungi. Keeling et al. (145) further showed that the microsporidia evolved from within the fungal group, sometime after the divergence of the chitrids. The idea that the microsporidia are phylogenetically nested within the fungal kingdom is also supported by the recent discovery of functional spliceosomal introns (20, 83). Evidence that this group of organisms once contained mitochondria (122) also supported this finding. It thus appears that the microsporidia area a highly specialized fungal lineage that 'lost' many of their eukaryotic features during adaptation to the intracellular parasitic lifestyle (20, 83, 145).

The discovery of the fungal heritage of microsporidia serves as just one example where protein-coding sequences have been used to determine the phylogenetic position of an evolutionarily ancient group of organisms. There are several other examples where these sequence have been useful in reconstructing the evolutionary histories of ancient lineages (141-143). The microsporidial example further shows that not all protein-coding genes are equally suited to address phylogenetic questions, at all levels. In this fungal lineage, the genes encoding the translation elongation factors were apparently too variable, which resulted in distorted genealogies. These genes have, however, proven useful in other fungal lineages (214, 216).

4.2 Low-level fungal taxonomy

4.2.1 Interspecific relationships

One of the best examples of a protein-coding gene being used to elucidate the relationships among closely related species is found in the work of O'Donnell et al. (212, 213). They studied the relationships among *Fusarium* species in the *Gibberella fujikuroi* complex using β-tubulin gene sequences. The fungi in this complex include well-known pathogens of many important agricultural plants (164, 165). Their classification has been hampered by the fact that they are morphologically very similar (26, 93, 204, 209).

In the taxonomy of *Fusarium* species belonging to the *G. fujikuroi* complex, the use of ribosomal ITS regions has proven to be problematic (212, 213). This is because they harbor non-orthologous divergent homologues of the ITS2 region that appears to have escaped concerted evolution (212, 213). Apparently, these homologues were the result of an interspecific

hybridization (xenologous origin) or gene duplication (paralogous origin) event. This event occurred prior to the radiation of species in this complex.

The phylogenetic relationships among the *Fusarium* species in this complex have been resolved using β-tubulin gene sequences (212, 213). What was thought to be three to eight species based on morphology, turned out to be more than 30 distinct *Fusarium* spp. (212, 214, 216). These results have also been confirmed using other protein-coding gene sequences, as well as morphological characters (33, 126, 164, 209, 263). Additionally, these protein-coding sequences have formed the basis for diagnostic techniques to identify members of this economically important group of plant pathogens (263).

4.2.2 Intraspecific relationships

Characterization of the intraspecific relationships among different groups of fungi is extremely important to fungal taxonomists. Because many fungi are asexual, their populations often constitute clones. The classification of these clones becomes increasingly important when they are associated with the production of mycotoxins or when they are serious plant and human pathogens (5, 42, 92, 154, 214, 270). The classification of the clonal and recombining lineages in the aflatoxin producing fungus, *Aspergillus flavus*, is one such an example. Using various protein-coding genes, Geiser et al. (92) showed that this apparently asexual fungus is separated into groups that correspond with their ability to produce toxin. Protein-coding gene sequences are thus valuable tools for identifying and classifying clonal or asexual, as well as recombining fungal lineages (42, 74, 92).

The review of the use of protein-coding genes to study the intraspecific relationships among populations of fungi would not be complete without reference to their value in detecting interpopulation-recombination events. A simple method, known as 'gene-gene concordance', was suggested to detect these recombination events (73, 74, 188, 299). Gene-gene concordance assesses the congruence between the phylogenetic trees constructed using several different genes. If concordant trees are obtained from all the genes tested, it is concluded that recombination among the individuals tested is rare. They thus represent clonal populations. If the gene trees for the group of individuals in question differ, recent recombination events among these individuals will have occurred. Although 'gene-gene concordance' is a relatively recent introduction to fungal taxonomy and population genetics, several authors have been able to detect sexual and asexual fungal lineages using this approach (39, 92, 154, 270, 271).

5 CONCLUSIONS

Fungal taxonomy has entered an exciting era, especially when taking into account that it is possible to reconstruct the evolutionary history of any group of individuals by using many different genes. In this way, many problems associated with traditional classification (for example morphological crypsis) have been or are in the process of being resolved using DNA sequence information from protein-coding genes. This is especially true in cases where the *rrn* genes and ITS regions display insufficient variability or where lineages are in the process of divergence. Consequently, the available nucleotide information on many different protein-coding genes in public domain databases is expanding continuously. Already, considerable collections of sequence data for proteins are available for important fungal lineages such as *Fusarium* and *Aspergillus*. In the future, these sequences will undoubtedly form the basis for DNA-based identification techniques and classification systems.

6 REFERENCES

1. **Alonso, S., A. Minty, Y. Bourlet, and M. Buckingham.** 1986. Comparison of three actin-coding sequences in the mouse; evolutionary relationships between the actin genes of warm-blooded vertebrates. J. Mol. Evol. **23:**11-22.
2. **Alvarez, F., M. N. Cortinas, and H. Musto.** 1996. The analysis of protein coding genes suggests monophyly of *Trypanosoma*. Mol. Phylogenet. Evol. **5:**333-343.
3. **Amegadzie, B. Y., R. S. Zitomer, and C. P. Hollenberg.** 1990. Characterization of the cytochrome c gene from the starch-fermenting yeast *Schwanniomyces occidentalis* and its expression in baker's yeast. Yeast. **6:**429-440.
4. **An, S. S., B. Mopps, K. Weber, and D. Bhattacharya.** 1999. The origin and evolution of green algal and plant actins. Mol. Biol. Evol. **16:**275-285.
5. **Aoki, T., and K. O'Donnell.** 1999. Morphological and molecular characterization of *Fusarium pseudograminearum* sp. nov., formerly recognized as the Group 1 population of *F. graminearum*. Mycologia. **91:**597-609.
6. **Apoil, P.-A., and A. Blancher.** 2000. *Rh* gene evolution in primates: study of intron sequences. Mol. Biol. Evol. **17:**127-136.
7. **Armstrong, J. L., and D. L. Harris.** 1993. Biased DNA integration in *Colletotrichum gloeosporioides* f. sp. *aeschynomene* transformants with benomyl resistance. Phytopathology. **83:**328-332.
8. **Arnheim, N.** 1983. Concerted evolution of multigene families, p. 38-61. *In* M. Nei and R. K. Koehn (ed.), Evolution of genes and proteins. Sinauer Associates, Inc., Sunderland, Massachusetts.
9. **Bagavathi, S., and R. Malathi.** 1996. Introns and protein evolution - an analysis of the exon/intron organization of actin genes. FEBS Lett. **392:**63-65.
10. **Baldauf, S. L., and W. F. Doolittle.** 1997. Origin and evolution of the slime molds (Mycetozoa). Proc. Natl. Acad. Sci. USA. **94:**12007-12012.
11. **Baldauf, S. L., and J. D. Palmer.** 1993. Animals and fungi each other's closest relatives: congruent evidence from multiple proteins. Proc. Natl. Acad. Sci. USA. **90:**11558-11562.
12. **Baldo, A. M., D. H. Les, and L. D. Strausbaugh.** 1999. Potentials and limitations of histone repeat sequences for phylogenetic reconstruction of *Sophophora*. Mol. Biol. Evol. **16:**1511-1520.
13. **Bannon, G. A., J. K. Bowen, M.-C. Yoa, and M. A. Gorovsky.** 1984. *Tetrahymena* H4 genes: structure, evolution and organization in macro- and micronuclei. Nucleic Acids Res.

14. **Barderi, P., O. Campetella, A. C. C. Frasch, J. Santome, U. Hellman, U. Pettersson, and J. J. Cazzulo.** 1998. The NADP+-linked glutamate dehydrogenase from *Trypanosoma cruzi*: sequence, genomic organization and expression. Biochem. J. **330:**951-958.

15. **Barron-Casella, E., and J. L. Corden.** 1992. Conservation of the mammalian RNA polymerase II largest-subunit C-terminal domain. J. Mol. Evol. **35:**405-410.

16. **Beech, R. N., R. K. Prichard, and M. E. Scott.** 1994. Genetic variability of the β-tubulin genes in benzimidazole-susceptible and -resistant strains of *Haemonchus contortus*. Genetics. **138:**103-110.

17. **Belfort, M., and P. S. Perlman.** 1995. Mechanism of intron mobility. J. Biol. Chem. **270:**30237-30240.

18. **Berbee, M. L., M. Pirseyedi, and S. Hubbard.** 1999. *Cochliobolus* phylogenetics and the origin of known, highly virulent pathogens, inferred from ITS and glyceraldehyde-3-phosphate dehydrogenase gene sequences. Mycologia. **91:**964-977.

19. **Bhattacharya, D., S. K. Stickel, and M. L. Sogin.** 1991. Molecular phylogenetic analysis of actin gene regions for *Achlya bisexualis* (Oomycota) and *Costaria costata* (Chromophyta). J. Mol. Evol. **33:**525-536.

20. **Biderre, C., G. Méténier, and C. P. Vivarès.** 1998. A small spliceosomal-type intron occurs in a ribosomal protein gene of the microsporidian *Encephalitozoon cunculi*. Mol. Biochem. Parasitol. **94:**283-286.

21. **Biderre, C., M. Pagès, G. Méténier, E. U. Canning, and C. P. Vivarès.** 1995. Evidence for the smallest nuclear genome (2.9 Mb) in the microsporidium *Encephalitzoon cuniculi*. Mol. Biochem. Parasitol. **74:**229-231.

22. **Blake, C. C. F.** 1978. Do genes-in-pieces imply proteins-in-pieces? Nature. **273:**267.

23. **Blake, C. C. F.** 1983. Exons - present from the beginning? Nature. **306:**535-537.

24. **Blake, C. C. F.** 1985. Exons and the evolution of proteins. Int. Rev. Cytol. **93:**149-185.

25. **Blake, C. C. F.** 1979. Exons encode protein functional units. Nature. **277:**598.

26. **Booth, C.** 1971. The genus *Fusarium*. Commonwealth Agricultural Bureau, Bucks, England.

27. **Borchiellini, C., N. Boury-Esnault, J. Vacelet, and Y. Le Parco.** 1998. Phylogenetic analysis of the Hsp70 sequences reveals the monophyly of metozoa and specific phylogenetic relationships between animals and fungi. Mol. Biol. Evol. **15:**647-655.

28. **Bork, P., C. Sander, and A. Valencia.** 1992. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. Proc. Natl. Acad. Sci. USA. **89:**7290-7294.

29. **Bowen, A. R., J. L. Chen-Wu, M. Momany, R. Young, P. J. Szaniszlo, and P. W.**

Robbins. 1992. Classification of fungal chitin synthase. Proc. Natl. Acad. Sci. USA. **89:**519-523.

30. **Bradnam, K. R., C. Seoighe, P. M. Sharp, and K. H. Wolfe.** 1999. G+C content variation along and among *Saccharomyces cerevisiae* chromosomes. Mol. Biol. Evol. **16:**666-675.
31. **Breathnach, R., and R. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50:**349-383.
32. **Brinkman, H., P. Martinez, F. Quigley, W. Martin, and R. Cerff.** 1987. Endosymbiotic origin and codon bias of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. J. Mol. Evol. **26:**320-328.
33. **Britz, H., T. A. Coutinho, M. J. Wingfield, W. F. O. Marasas, T. R. Gordon, and J. F. Leslie.** 1999. *Fusarium subglutinans* f. sp. *pini* represents a distinct mating population in the *Gibberella fujikuroi* species complex. Appl. Environ. Microbiol. **65:**1198-1201.
34. **Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson.** 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18:**225-239.
35. **Bruns, T. D., T. J. White, and J. W. Taylor.** 1991. Fungal molecular systematics. Annu. Rev. Ecol. Syst. **22:**525-564.
36. **Brynnel, E. U., C. G. Kurland, N. A. Moran, and S. G. E. Andersson.** 1998. Evolutionary rates for *tuf* genes and endosymbionts of aphids. Mol. Biol. Evol. **15:**574-582.
37. **Buhr, T. L., and M. B. Dickman.** 1994. Isolation, characterization, and expression of a second β-tubulin-encoding gene from *Colletotrichum gloeosporioides* f.sp. *aeschynomene.* Appl. Environ. Microbiol. **60:**4155-4159.
38. **Bulawa, C. E.** 1993. Genetics and molecular biology of chitin synthesis in fungi. Annu. Rev. Microbiol. **47:**505-534.
39. **Burt, A., D. A. Carter, G. L. Koenig, T. J. White, and J. W. Taylor.** 1996. Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis.* Proc. Natl. Acad. Sci. USA. **93:**770-773.
40. **Bussink, H. J., F. P. Buxton, B. A. Fraaye, L. H. de Graaff, and J. Visser.** 1992. The polygalacturonases of *Aspergillus niger* are encoded by a family of diverged genes. Eur. J. Biochem. **208:**83-90.
41. **Canning, E. U.** 1990. Phylum Microsporia, p. 53-72. *In* L. Margulis, J. O. Corliss, M. Melkonian, and D. J. Chapman (ed.), Handbook of Protoctista. Jones and Bartlett Publishers, Boston.
42. **Carbone, I., J. B. Anderson, and L. M. Kohn.** 1999. Patterns of descent in clonal lineages and their multilocus fingerprints are resolved with combined gene genealogies. Evolution. **53:**11-21.

43. **Cavalier-Smith, T.** 1991. Intron phylogeny: a new hypothesis. Trends Genet. **7:**145-148.

44. **Cavalier-Smith, T.** 1993. Kingdom Protozoa and its 18 Phyla. Microbiol. Rev. **57:**953-994.

45. **Christiansen, S. K., A. F. Justensen, and H. Giese.** 1997. Disparate sequence characteristics of the *Erysiphe graminis* f. sp. *hordei* glyceraldehyde-3-phosphate dehydrogenase gene. Curr. Genet. **31:**525-529.

46. **Cleveland, D. W., and K. F. Sullivan.** 1985. Molecular biology and genetics of tubulin. Annu. Rev. Biochem. **54:**331-365.

47. **Coen, E., T. Strachan, and G. Dover.** 1982. Dynamics of concerted evolution of ribosomal DNA and histone gene families in the *melanogaster* species subgroup of *Drosophila*. J. Mol. Biol. **158:**17-35.

48. **Colgan, D. J.** 1999. Phylogenetic studies of marsupials based on phosphoglycerate kinase DNA sequences. Mol. Phylogenet. Evol. **11:**13-26.

49. **Comeron, J. M., and M. Kreitman.** 1998. The correlation between synonymous and nonsynonymous substitutions in *Drosophila*: mutation, selection or relaxed constraints? Genetics. **150:**767-775.

50. **Copertino, D. W., and R. B. Hallick.** 1993. Group II and group III introns of twintrons: potential relationships with nuclear pre-mRNA introns. Trends Biochem. Sci. **18:**467-471.

51. **Croan, D. G., D. A. Morrison, and J. T. Ellis.** 1997. Evolution of the genus *Leishmania* revealed by comparison of DNA and RNA polymerase gene sequences. Mol. Biochem. Parasitol. **89:**149-159.

52. **Cronn, R. C., X. Zhao, A. H. Paterson, and J. F. Wendel.** 1996. Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. J. Mol. Evol. **42:**685-705.

53. **Danforth, B. N., H. Sauqet, and L. Packer.** 1999. Phylogeny of the bee genus *Halictus* (Hymenoptera:Halictidae) based on parsimony and likelihood analyses of nuclear elongation factor-1α sequence data. Mol. Phylogenet. Evol. **13:**605-618.

54. **Davidse, L. C.** 1986. Benzimidazole fungicides: mechanism of action and biological impact. Annu. Rev. Phytopathol. **24:**43-65.

55. **De los Angeles Barriga Sosa, I., K. Beckenbach, B. Hartwick, and M. J. Smith.** 1995. The molecular phylogeny of five Eastern and North Pacific octopus species. Mol. Phylogenet. Evol. **4:**163-174.

56. **Denton, A. L., B. L. McConaughy, and B. D. Hall.** 1998. Usefulness of RNA polymerase II coding sequences for estimation of green plant phylogeny. Mol. Biol. Evol. **15:**1082-1085.

57. **DeSalle, R., T. Freedman, E. M. Prager, and A. C. Wilson.** 1987. Tempo and mode of

sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol. **26:**157-164.

58. **Dibb, N. J., and A. J. Newman.** 1989. Evidence that introns arose at proto-splice sites. EMBO J. **8:**2015-2021.

59. **Dietmaier, W., and S. Fabry.** 1994. Analysis of the introns in genes encoding small G proteins. Curr. Genet. **26:**497-505.

60. **Din, A. B., and O. Yarden.** 1994. The *Neurospora crassa chs-2* gene encodes a non-essential chitin synthase. Microbiology. **140:**2189-2197.

61. **Dix, D. B., and R. C. Thompson.** 1989. Codon choice and gene expression: synonymous codons differ in translational accuracy. Proc. Natl. Acad. Sci. USA. **86:**6888-6892.

62. **Dohmen, R. J., A. W. M. Strasser, U. M. Dahlems, and C. P. Hollenberg.** 1990. Cloning of the *Schwanniomyces occidentalis* glucoamylase gene (*GAM1*) and its expression in *Saccharomyces cerevisiae*. Gene. **95:**111-121.

63. **Donaldson, G. C., L. A. Ball, P. E. Axelrood, and N. L. Glass.** 1995. Primer sets developed to amplify conserved genes from filamentous ascomycetes are useful in differentiating *Fusarium* species associated with conifers. Appl. Environ. Microbiol. **61:**1331-13340.

64. **Doolittle, R. F.** 1995. The multiplicity of domains in proteins. Annu. Rev. Biochem. **64:**287-314.

65. **Doolittle, W. F.** 1987. The origin and function of intervening sequences in DNA: a review. Am. Nat. **130:**915-928.

66. **Dover, G.** 1982. Molecular drive: a cohesive mode of species evolution. Nature. **299:**111-117.

67. **Dowling, T. E., C. Moritz, and J. D. Palmer.** 1990. Nucleic Acids II. Restriction site analysis, p. 250-317. *In* D. M. Hillis and C. Moritz (ed.), Molecular Systematics. Sinauer Associates, Inc., Sunderland, Massachusetts.

68. **Drouin, G., M. M. De Sa, and M. Zuker.** 1995. The *Giardia lambia* actin gene and the phylogeny of eukaryotes. J. Mol. Evol. **41:**841-849.

69. **Drouin, G., and G. A. Dover.** 1990. Independent gene evolution in the potato actin gene family demonstrated by phylogenetic procedures for resolving gene conversions and the phylogeny of angiosperm actin genes. J. Mol. Evol. **31:**132-150.

70. **Drouin, G., F. Prat, M. Ell, and G. D. P. Clarke.** 1999. Detecting and characterizing gene conversion between multigene family members. Mol. Biol. Evol. **16:**1369-1390.

71. **Dudler, R.** 1990. The single-copy actin gene of *Phytophthora megasperma* encodes a protein considerably diverged from any other actin. Plant Mol. Biol. **14:**415-422.

72. **Dujon, B.** 1989. Group I introns as mobile genetic elements: facts and mechanistic speculations - a review. Gene. **82:**91-114.

73. **Dykhuizen, D. E., and L. Green.** 1991. Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. **173:**7257-7268.

74. **Dykhuizen, D. E., D. S. Polin, J. J. Dunn, B. Wilske, V. Preac-Mursic, R. J. Dattwyler, and B. J. Luft.** 1993. *Borrelia burgdorfi* is clonal: Implications for taxonomy and vaccine development. Proc. Natl. Acad. Sci. USA. **90:**10163-10167.

75. **Edlind, T. D., J. Li, G. S. Visvesvara, M. H. Vodkin, G. L. McLaughlin, and S. K. Katiyar.** 1996. Phylogenetic analysis of β-tubulin sequences from amitochondrial protozoa. Mol. Phylogenet. Evol. **5:**359-367.

76. **Eggert, C., P. LaFayette, U. Temp, K.-E. L. Eriksson, and J. F. D. Dean.** 1998. Molecular analysis of a laccase gene from the white rot fungus *Pycnoporus cinnabarinus*. Appl. Environ. Microbiol. **64:**1766-1772.

77. **Embley, T. M., and R. P. H. Hirt.** 1998. Early branching eukaryotes? Curr. Opin. Genet. Dev. **8:**624-629.

78. **Engelberg-Kulka, H., I. Benhar, and R. Schoulaker-Schwarz.** 1993. Translational introns: an additional regulatory element in gene expression. Trends Biochem. Sci. **18:**294-296.

79. **Fabry, S., K. Muller, A. Lindauer, P. B. Park, T. Cornelius, and R. Schmitt.** 1995. The organization structure and regulatory elements of *Chlamydomonas* histone genes reveal features linking plant and animal genes. Curr. Genet. **28:**333-345.

80. **Fackenthal, J. D., F. R. Turner, and E. Raff.** 1993. Tissue-specific microtubule functions in *Drosophila* spermatogenesis require the B2-tubulin isotype-specific carboxy terminus. Dev. Biol. **158:**213-227.

81. **Fang, H., and B. P. Brandhorst.** 1994. Evolution of actin gene families of sea urchins. J. Mol. Evol. **39:**347-356.

82. **Fast, N. M., J. M. Logsdon Jr, and W. F. Doolittle.** 1999. Phylogenetic analysis of the TATA box binding protein (TBP) gene from *Nosema locustae*: evidence for a Microsporidia-fungal relationship and spliceosomal intron loss. Mol. Biol. Evol. **16:**1415-1419.

83. **Fast, N. M., A. J. Roger, C. A. Richardson, and W. F. Doolittle.** 1998. U2 and U6 snRNA genes in the microsporidian *Nosema locustae*: evidence for a functional spliceosome. Nucleic Acids Res. **26:**3202-3027.

84. **Figge, R. M., M. Schubert, H. Brinkmann, and R. Cerff.** 1999. Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra-

and inter-kingdom gene transfer. Mol. Biol. Evol. **16:**429-440.

85. **Fink, G. R.** 1987. Pseudogenes in Yeast? Cell. **49:**5-6.

86. **Fitch, W. M.** 1976. The molecular evolution of cytochrome c in eukaryotes. J. Mol. Evol. **8:**13-40.

87. **Fletcher, L., J. M. McDowell, R. R. Tidwell, R. B. Meagher, and C. C. Dykstra.** 1994. Structure, expression and phylogenetic analysis of the gene encoding actin I in *Pneumocystis carinii*. Genetics. **137:**743-750.

88. **Frendewey, D., and W. Keller.** 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell. **42:**355-367.

89. **Fukami-Kobayashi, K., M. Nosaka, A. Nakazawa, and M. Gõ.** 1996. Ancient divergence of long and short isoforms of adenylate kinase: molecular evolution of the nucleoside monophosphate kinase family. FEBS Lett. **385:**214-220.

90. **Ganley, A. R. D., and B. Scott.** 1998. Extraordinary ribosomal spacer length heterogeneity in a Neotyphodium endophyte hybrid: Implications for concerted evolution. Genetics. **150:**1625-1637.

91. **Garcia-Valve, S., J. Palau, and A. Romeu.** 1999. Horizontal gene transfer in glycosyl hydrolases inferred from codon usage in *Escherichia coli* and *Bacillus subtilis*. Mol. Biol. Evol. **16:**1125-1134.

92. **Geiser, D. M., J. I. Pitt, and J. W. Taylor.** 1998. Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. Proc. Natl. Acad. Sci. USA. **95:**388-393.

93. **Gerlach, W., and H. I. Nirenberg.** 1982. The genus *Fusarium* - A pictorial atlas. Mitt. Biol. Bundesanst. Land-Forstw. Berlin-Dahlem. **209:**1-406.

94. **Germot, A., H. Philippe, and H. Le Guyader.** 1997. Evidence for loss of mitochondria in Microsporidia from a mitochondrial-type HSP70 in *Nosema locustae*. Mol. Biochem. Parasitol. **87:**159-168.

95. **Giardina, P., G. Palmieri, A. Scaloni, B. Fontanell, V. Faraco, G. Cennamo, and G. Sannia.** 1999. Protein and gene structure of a blue laccase from *Pleurotus ostreatus*. Biochem. J. **341:**655-663.

96. **Gilbert, W.** 1987. The exon theory of genes. Cold Spring Harbor Symp. Quant. Biol. **52:**901-905.

97. **Gilbert, W.** 1978. Why genes in pieces? Nature. **271:**501.

98. **Gilbert, W., M. Marchionni, and G. McKnight.** 1986. On the antiquity of introns. Cell. **46:**151-154.

99. **Gõ, M., and M. Nosaka.** 1987. Protein architecture and the origin of introns. Cold Spring Harbor Symp. Quant. Biol. **52:**915-924.

100. **Gojobori, T., E. N. Moriyama, and M. Kimura.** 1990. Statistical methods for estimating sequence divergence. Meth. Enzymol. **183:**331-351.

101. **Gold, S. E., W. L. Casale, and N. T. Keen.** 1991. Characterization of two β-tubulin genes from *Geotrichum candidum*. Mol. Gen. Genet. **230:**104-112.

102. **Gonzales, P., G. Barroso, and J. Labarere.** 1998. Molecular analysis of the split *cox1* gene from the Basidiomycota *Agrocybe aegerita*: relationships of its introns with homologous Ascomycota introns and divergence levels from common ancestral copies. Gene. **220:**45-53.

103. **Gouy, M., and C. Gautier.** 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. **10:**7055-7074.

104. **Grabowski, P. J., S. R. Seiler, and P. A. Sharp.** 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell. **42:**345-353.

105. **Grandy, D. K., and J. B. Dodgson.** 1987. Structure and organization of the chicken H2B histone gene family. Nucleic Acids Res. **15:**1063-1080.

106. **Grantham, R., C. Gautier, M. Gouy, R. Mercier, and A. Pave.** 1980. Codon catalog usage and the genome hypothesis. Nucleic Acids Res. **8:**r49-r62.

107. **Gribaldo, S., and P. Cammarano.** 1998. The root of the universal tree of life inferred from anciently duplicated genes encoding components of the protein-targeting machinery. J. Mol. Evol. **47:**508-516.

108. **Guiltinan, M. J., D.-P. Ma, R. F. Barker, M. M. Bustos, R. J. Cyr, R. Yadegari, and D. E. Fosket.** 1987. The isolation, characterization and sequence of two divergent β-tubulin genes from soybean (*Glycine max* L.). Plant Mol. Biol. **10:**171-184.

109. **Gwo-Shu Lee, M., S. A. Lewis, C. D. Wilde, and N. J. Cowan.** 1983. Evolutionary history of a multigene family: an expressed human β-tubulin gene and three processed pseudogenes. Cell. **33:**477-487.

110. **Hall, S. L., and R. A. Padgett.** 1994. Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. J. Mol. Biol. **239:**357-365.

111. **Hall, S. L., and R. A. Padgett.** 1996. Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. Science. **271:**1716-1718.

112. **Han, I.-S., I. Jongewaard, and D. E. Fosket.** 1991. Limited expression of a diverged β-tubulin gene during soybean (*Glycine max* [L.] Merr.) development. Plant Mol. Biol. **16:**225-235.

113. **Hankeln, T., H.-G. Keyl, R. Ross, and E. R. Schmidt.** 1993. Evolution of histone gene loci in chironomid midges. Genome. **36:**852-862.

114. **Hankeln, T., and E. R. Schmidt.** 1993. Divergent evolution of an "orphan" histone gene cluster in *Chironomus*. J. Mol. Biol. **234:**1301-1307.

115. **Hasegawa, M., and T. Hashimoto.** 1993. Ribosomal RNA tree misleading? Nature. **361:**23.

116. **Hashimoto, T., Y. Nakamura, T. Kamaishi, F. Nakamura, J. Adachi, K.-I. Okamoto, and M. Hasegawa.** 1995. Phylogenetic place of mitochondrial-lacking protozoan, *Giardia lamblia*, inferred from amino acid sequences of elongation factor 2. Mol. Biol. Evol. **12:**782-793.

117. **He, M., and D. S. Haymer.** 1995. Codon bias in actin multigene families and effects on the reconstruction of phylogenetic relationships. J. Mol. Evol. **41:**141-149.

118. **Hensel, R., P. Zwickl, S. Fabry, J. Lang, and P. Palm.** 1989. Sequence comparison of glyceraldehyde-3-phosphate dehydrogenase from the three urkingdoms: evolutionary implications. Can. J. Microbiol. **35:**81-85.

119. **Hibbett, D. S.** 1992. Ribosomal RNA and fungal systematics. Trans. Mycol. Soc. Jpn. **33:**533-556.

120. **Hillis, D. M., C. Moritz, C. A. Porter, and R. J. Baker.** 1991. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science. **18:**308-310.

121. **Hiraoka, Y., T. Toda, and M. Yanagida.** 1984. The *NDA3* gene of fission yeast encodes β-tubulin: a cold-sensitive *nda3* mutation reversibly blocks spindle formation and chromosome movement in mitosis. Cell. **39:**349-358.

122. **Hirt, R. P., B. Healy, C. R. Vossbrinck, E. U. Canning, and T. M. Embley.** 1997. A mitochondrial *Hsp70* orthologue in *Vairimorpha necatrix*: molecular evidence that microsporidia once contained mitochondria. Curr. Microbiol. **7:**995-998.

123. **Hirt, R. P., J. M. Logsdon Jr, B. Healy, M. W. Dorey, W. F. Doolittle, and T. M. Embley.** 1999. Microsporidia are related to fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA. **96:**580-585.

124. **Holmquist, R.** 1983. Transitions and transversions in evolutionary descent: an approach to understanding. J. Mol. Evol. **19:**134-144.

125. **Honeycutt, R. L., M. A. Nedbal, R. M. Adkins, and L. L. Janecek.** 1995. Mammalian mitochondrial DNA evolution: a comparison of the cytochrome *b* and cytochrome *c* oxidase genes. J. Mol. Evol. **40:**260-272.

126. **Hsieh, W. H., S. N. Smith, and W. C. Snyder.** 1977. Mating groups in *Fusarium moniliforme*. Phytopathology. **67:**1041-1043.

127. **Hughes, A. L.** 1999. Concerted evolution of exons and introns in the MHC-linked tenascin-X gene of mammals. Mol. Biol. Evol. **16:**1558-1567.

128. **Hussey, P. J., N. Haas, J. Hunsperger, J. Larkin, D. P. Snustad, and C. D. Silflow.** 1990. The β-tubulin gene family in *Zea mays*: two differentially expressed β-tubulin genes. Plant Mol. Biol. **15:**957-972.

129. **Hussey, P. J., C. W. Lloyd, and K. Gull.** 1988. Differential and developmental expression of β-tubulin in higher plants. J. Biol. Chem. **263:**5474-5479.

130. **Ikemura, T.** 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. **2:**13-34.

131. **Ikemura, T.** 1982. Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. J. Mol. Biol. **158:**573-597.

132. **Ikemura, T., and H. Ozeki.** 1982. Codon usage and transfer RNA contents: Organism-specific codon-choice patterns in reference to isoacceptor contents. Cold Spring Harbor Symp. Quant. Biol. **47:**1087-1097.

133. **Iwabe, N., K. Kuma, H. Kishino, M. Hasegawa, and T. Miyata.** 1991. Evolution of RNA polymerase and branching patterns of the three major groups of Archaebacteria. J. Mol. Evol. **32:**70-78.

134. **Iwabe, N., K.-E. Kuma, H. Kishino, M. Hasegawa, and T. Miyata.** 1990. Compartmentalized isozyme genes and the origin of introns. J. Mol. Evol. **31:**205-210.

135. **Iwabe, N., K.-I. Kuma, M. Hasegawa, S. Osawa, and T. Miyata.** 1989. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA. **86:**9355-9359.

136. **Jackson, I. J.** 1991. A reappraisal of non-consensus mRNA splice sites. Nucleic Acids Res. **19:**3795-3798.

137. **Jacquier, A.** 1990. Self-splicing group II and nuclear pre-mRNA introns: how similar are they? Trends Biochem. Sci. **15:**351-254.

138. **Janbon, G., E. P. Rustchenko, S. Klug, S. Scherer, and F. Sherman.** 1997. Phylogenetic relationships of fungal cytochromes *c*. Yeast. **13:**985-990.

139. **Kamaishi, T., T. Hashimoto, Y. Nakamura, Y. Masuda, F. Nakamura, K. Okamoto, M. Shimizu, and M. Hasegawa.** 1996. Complete nucleotide sequences of the genes encoding translation elongation factors 1α and 2 from a microsporidian parasite, *Glugea plecoglossi*: Implications for the deepest branching of eukaryotes. J. Biochem. **120:**1095-1103.

140. **Kano, R., Y. Nakamura, T. Watari, S. Watanabe, H. Takahashi, H. Tsujimoto, and A. Hasagawe.** 1998. Molecular analysis of chitin synthase I (*CHS1*) gene sequences of *Trichophyton mentagrophytes* complex and *T. rubrum*. Curr. Microbiol. **37:**236-239.

141. **Keeling, P. J.** 1998. A kingdom's progress: Archaezoa and the origin of eukaryotes. BioEssays. **20:**87-95.

142. **Keeling, P. J., J. A. Deane, C. Hink-Schauer, S. E. Douglas, U.-G. Maier, and G. I. McFadden.** 1999. The secondary endosymbiont of the Cryptomonad *Guillardia theta* contains alpha-, beta- and gamma-tubulin genes. Mol. Biol. Evol. **16:**1308-1313.

143. **Keeling, P. J., J. A. Deane, and G. I. McFadden.** 1998. The phylogenetic position of alpha- and beta-tubulins from the *Chlorarachnion* host and *Cercomonas* (Cercozoa). J. Eukaryot. Microbiol. **45:**561-570.

144. **Keeling, P. J., and W. F. Doolittle.** 1996. Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin gene family. Mol. Biol. Evol. **13:**1297-1305.

145. **Keeling, P. J., M. A. Luker, and J. D. Palmer.** 2000. Evidence from beta-tubulin phylogeny that Microsporidia evolved from within the fungi. Mol. Biol. Evol. **17:**21-31.

146. **Kemmerer, E. C., M. Lei, and R. Wu.** 1991. Structure and molecular evolutionary analysis of a plant cytochrome c gene: surprising implications for *Arabidopsis thaliana*. J. Mol. Evol. **32:**227-234.

147. **Kita, K., T. Fukura, K.-I. Nakase, K. Okamoto, H. Yanase, M. Kataoka, and S. Shimizu.** 1999. Cloning, overexpression, and mutagenesis of the *Sporobolomyces salmonicolor* AKU4429 gene encoding a new aldehyde reductase, which catalyzes the stereoselective reduction of ethyl 4-chloro-3-oxobutanoate to ethyl (*S*)-4-chloro-3-hydroxybutanoate. Appl. Environ. Microbiol. **65:**5207-5211.

148. **Kmiecik, D., M. Couppez, D. Belaiche, and P. Sautiere.** 1983. Primary structure of histone H2A from nucleated erythrocyte of the marine worm *Sipunculus nudus*. Eur. J. Biochem. **135:**113-121.

149. **Kobayashi, M., H. Wada, and N. Satoh.** 1996. Early evolution of the metazoa and phylogenetic status of diploblasts as inferred form amino acid sequence of elongation factor-1α. Mol. Phylogenet. Evol. **5:**414-422.

150. **Koenraadt, H., and A. L. Jones.** 1993. Resistance to benomyl conferred by mutations in condon 198 or 200 of the beta-tubulin gene of *Neurospora crassa* and sensitivity to diethofencarb conferred by codon 198. Phytopathology. **83:**850-854.

151. **Kogan, G. L., V. N. Epstein, A. A. Aravin, and V. A. Gvosdev.** 2000. Molecular evolution of two paralogous tandemly repeated heterochromatic gene clusters linked to the X and Y chromosomes of *Drosophila melanogaster*. Mol. Biol. Evol. **17:**697-702.

152. **Kojima, S.** 1998. Paraphyletic status of Polychaeta suggested by phylogenetic analysis based on the amino acid sequences of elongation factor-1α. Mol. Phylogenet. Evol. **9:**255-261.

153. **Kopczak, S. D., N. A. Haas, P. J. Hussey, C. D. Silflow, and P. D. Snustad.** 1992. The small genome of *Arabidopsis* contains at least six expressed α-tubulin genes. Plant Cell.

4:539-547.

154. **Koufopanou, V., A. Burt, and W. J. Taylor.** 1997. Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. Proc. Natl. Acad. Sci. USA. **94:**5478-5482.

155. **Kretzer, A. M., and T. D. Bruns.** 1999. Use of *atp6* in fungal phylogenetics: an example from the Boletales. Mol. Phylogenet. Evol. **13:**483-492.

156. **Kruse, M., V. Gamulin, H. Cetkovic, Z. Pancer, I. M. Muller, and W. E. G. Muller.** 1996. Molecular evolution of the metazoan protein kinase C multigene family. J. Mol. Evol. **43:**374-383.

157. **Kuma, K., N. Nikoh, N. Iwabe, and T. Miyata.** 1995. Phylogenetic position of *Dictyostelium* inferred from multiple protein data sets. J. Mol. Evol. **41:**238-246.

158. **Kusakabe, T., I. Araki, N. Satoh, and W. R. Jeffery.** 1997. Evolution of chordate actin genes: evidence from genomic organization and amino acid sequences. J. Mol. Evol. **44:**289-298.

159. **Lafay, B., and P. M. Sharp.** 1999. Synonymous codon usage variation among *Giardia lamblia* genes and isolates. Mol. Biol. Evol. **16:**1484-1495.

160. **Lambowitz, A. M., and M. Belfort.** 1993. Introns as mobile genetic elements. Annu. Rev. Biochem. **62:**587-622.

161. **Lazcano, A., J. Fastag, P. Gariglio, C. Ramirez, and J. Oro.** 1988. On the early evolution of RNA polymerase. J. Mol. Evol. **27:**365-376.

162. **Lee, S. H., J. C. Kim, M. S. Lee, W. D. Heo, H. Y. Seo, H. W. Yoon, J. C. Hong, S. Y. Lee, J. D. Bahk, I. Hwang, and M. J. Cho.** 1995. Identification of a novel divergent calmodulin isoform from soybean which has differential ability to activate calmodulin-dependent enzymes. J. Biol. Chem. **270:**21806-21812.

163. **Lee, Y. H., X.-Y. Huang, D. Hirsh, G. E. Fox, and R. M. Hecht.** 1992. Conservation of gene organization and *trans*-splicing in the glyceraldehyde-3-phosphate dehydrogenase-encoding genes of *Caenorhabditis briggsae*. Gene. **121:**227-235.

164. **Leslie, J. F.** 1995. *Gibberella fujikuroi*: Available populations and variable traits. Can. J. Bot. **73:**S282-S291.

165. **Leslie, J. F.** 1991. Mating populations in *Gibberella fujikuroi* (*Fusarium* section *Liseola*). Phytopathology. **81:**1058-1060.

166. **Lewis, L. A., B. D. Mishler, and R. Vilgalys.** 1997. Phylogenetic relationships of the liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene *rbc*L. Mol. Phylogenet. Evol. **7:**377-393.

167. **Li, W.-H., and D. Graur.** 1991. Fundamentals of molecular evolution. Sinauer Associates,

Inc., Sunderland, Massachusetts.

168. **Liaud, M.-F., H. Brinkmann, and R. Cerff.** 1992. The β-tubulin gene family of pea: Primary structures, genomic organization and intron-depèndent evolution of genes. Plant Mol. Biol. **18:**639-651.

169. **Liaud, M.-F., C. Lichtle, K. Apt, W. Martin, and R. Cerff.** 2000. Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of eukaryotic glycolytic pathway. Mol. Biol. Evol. **17:**213-223.

170. **Liu, Y. J., S. Whelen, and B. D. Hall.** 1999. Phylogenetic relationships among Ascomycetes: Evidence from an RNA polymerase II subunit. Mol. Biol. Evol. **16:**1799-1808.

171. **Lloyd, A. T., and P. M. Sharp.** 1992. Evolution of codon usage patterns: the extent and nature of divergence between *Candida albicans* and *Saccharomyces cerevisiae*. Nucleic Acids Res. **20:**5289-5295.

172. **Logsdon Jr, J. M., A. Stoltzfus, and W. F. Doolittle.** 1998. Molecular evolution: Recent cases of spliceosomal intron gain? Curr. Biol. **8:**R560-R563.

173. **Long, M., and M. Deutsch.** 1999. Association of intron phases with conservation at splice site sequences and evolution of spliceosomal introns. Mol. Biol. Evol. **16:**1528-1534.

174. **Lutkenhaus, J.** 1993. FtsZ ring in bacterial cytokinesis. Mol. Microbiol. **9:**403-409.

175. **Lyddiatt, A., D. Peacock, and D. Boulter.** 1978. Evolutionary change in invertebrate cytochrome c. J. Mol. Evol. **11:**35-45.

176. **Lyons-Weiler, J., and G. A. Hoelzer.** 1999. Null model selection, compositional bias, character state bias, and the limits of phylogenetic information. Mol. Biol. Evol. **16:**1400-1405.

177. **Madhani, H. D., and C. Guthrie.** 1994. Dynamic RNA-RNA interactions in the spliceosome. Annu. Rev. Genetics. **28:**1-26.

178. **Margolin, W., R. Wang, and M. Kumar.** 1996. Isolation of an *ftsZ* homologue from the archaebacterium *Halobacterium salinarium*: implications for the evolution of ftsZ and tubulin. J. Bacteriol. **178:**1320-1327.

179. **Marinets, A., M. Muller, P. J. Johnson, J. Kulda, O. Scheiner, G. Wierdermann, and M. Duchene.** 1996. The sequence and organization of the core histone H3 and H4 genes in the early branching amitochondriate protist *Trichomonas vaginalis*. J. Mol. Evol. **43:**563-571.

180. **Markos, A., A. Morris, C. Rozario, and M. Muller.** 1996. Primary structure of a cytosolic malate dehydrogenase of the amitochondriate eukaryote, *Trichomonas vaginalis*. FEMS

Microbiol. Lett. **135:**259-264.

181. **Marks, M. D., J. West, and D. P. Weeks.** 1987. The relatively large beta-tubulin gene family of *Arabidopsis* contains a member with an unusual transcribed 5' noncoding sequence. Plant Mol. Biol. **10:**91-104.

182. **Martin, A. P.** 1999. Substitution rates of organelle and nuclear genes in sharks: implicating metabolic rate (again). Mol. Biol. Evol. **16:**996-1002.

183. **Martin, P. G., D. Boulter, and D. Penny.** 1985. Angiosperm phylogeny studied using sequences of five macromolecules. Taxon. **34:**393-400.

184. **Martin, P. G., J. M. Dowd, and S. J. L. Stone.** 1983. The study of plant phylogeny using amino acid sequences of ribulose-1,5-bisphosphate carboxylase. II The analysis of small subunit data to form phylogenetic trees. Aust. J. Bot. **31:**411-419.

185. **Martinez, P., W. Martin, and R. Cerff.** 1989. Structure, evolution and anaerobic regulation of a nuclear gene encoding cytosolic glyceraldehyde-3-phosphate dehydrogenase from maize. J. Mol. Biol. **208:**551-565.

186. **Maxson, R., T. Mohun, G. Gormezano, and L. Kedes.** 1987. Evolution of late H2A, H2B and H4 histone genes of the sea urchin, *Strongylocentrotus purpuratus.* Nucleic Acids Res. **15**.

187. **May, G. S., M. L.-S. Tsang, H. Smith, S. Fidel, and N. R. Morris.** 1987. *Aspergillus nidulans* β-tubulin genes are unusually divergent. Gene. **55:**231-243.

188. **Maynard Smith, J., and N. H. Smith.** 1998. Detecting recombination from gene trees. Mol. Biol. Evol. **15:**590-599.

189. **McDowell, J. M., S. Huang, E. C. McKinney, Y.-Q. An, and R. B. Meagher.** 1996. Structure and evolution of the actin gene family in *Arabidopsis thaliana*. Genetics. **142:**587-602.

190. **Meagher, R. D.** 1991. Divergence and differential expression of actin gene families in higher plants. Int. Rev. Cytol. **125:**139-163.

191. **Mehmann, B., I. Brunner, and G. H. Braus.** 1994. Nucleotide sequence variation of chitin synthase genes among ectomycorrhizal fungi and its potential use in taxonomy. Appl. Environ. Microbiol. **60:**3105-3111.

192. **Mellado, E., A. Aufauvre-Brown, N. A. R. Gow, and D. W. Holden.** 1996. The *Aspergillus fumigatus chsC* and *chsG* genes encode class III chitin synthases with different functions. Mol. Microbiol. **20:**667-679.

193. **Mellado, E., A. Aufauvre-Brown, C. A. Specht, P. W. Robbins, and D. W. Holden.** 1995. A multigene family related to chitin synthase genes of yeast in the opportunistic pathogen *Aspergillus fumigatus*. Mol. Gen. Genet. **246:**353-359.

194. **Michel, F., and J.-L. Ferat.** 1995. Structure and activities of group II introns. Annu. Rev. Biochem. **64:**435-461.

195. **Michels, P. A. M., M. Marchand, L. Kohl, S. Allert, R. K. Wierenga, and F. R. Opperdoes.** 1991. The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate dehydrogenase in *Trypanosoma brucei* have a distant evolutionary relationship. Eur. J. Biochem. **198:**421-428.

196. **Mikuni, J., and N. Morohoshi.** 1997. Cloning and sequencing of a second laccase gene from the white-rot fungus *Coriolus versicolor*. FEMS Microbiol. Lett. **155:**79-84.

197. **Miller, D. J., P. L. Harrison, T. J. Mahony, J. P. McMillan, A. Miles, D. M. Odorica, and M. R. ten Lohuis.** 1993. Nucleotide sequence of the histone gene cluster in the coral *Acropora formosa* (Cnidaria; Scleractinia): Features of histone gene structure and organization are common to diploblastic and triploblastic metazoans. J. Mol. Evol. **37:**245-253.

198. **Mio, T., T. Yabe, M. Sudoh, Y. Satoh, T. Nakajima, M. Arisawa, and H. Yamada-Okabe.** 1996. Role of three chitin synthase genes in the growth of *Candida albicans*. J. Bacteriol. **178:**2416-2419.

199. **Mort-Bontemps, M., L. Gay, and M. Fevre.** 1997. *CHS2*, a chitin synthase gene from the oomycete *Saprolegnia monoica*. Microbiology. **143:**2009-2020.

200. **Mount, S. M.** 1996. AT-AC introns: an ATtACk on dogma. Science. **271:**1690-1692.

201. **Nakamura, Y., T. Hashimoto, T. Kamaishi, J. Adachi, F. Nakamura, K. Okamoto, and M. Hasegawa.** 1996. Phylogenetic position of kinetoplastid protozoa inferred from the protein phylogenies of elongation factors 1α and 2. J. Biochem. **119:**70-79.

202. **Neff, N. F., J. H. Thomas, P. Grisafi, and D. Botstein.** 1983. Isolation of the β-tubulin gene from yeast and demonstration of its essential function in vivo. Cell. **33:**211-219.

203. **Nelson, D. R.** 1999. Cytochrome P450 and the individuality of species. Arch. Biochem. Biophys. **369:**1-10.

204. **Nelson, P. E., T. A. Toussoun, and W. F. O. Marasas.** 1983. *Fusarium* species: an illustrated manual of identification. Pennsylvania State University Press, University Park.

205. **Newman, A. J.** 1997. The role of U5 snRNP in pre-mRNA splicing. EMBO J. **19:**5797-5800.

206. **Nielsen, P. S., A. Kleinshofs, and O.-A. Olsen.** 1997. Barley elongation factor 1α: genomic organization, DNA sequence, and phylogenetic implications. Genome. **40:**559-565.

207. **Nikoh, N., N. Hayase, N. Iwabe, K.-I. Kuma, and T. Miyata.** 1994. Phylogenetic relationship of the kingdoms Animalia, Plantae, and Fungi inferred from 23 different protein sequences. Mol. Biol. Evol. **11:**762-768.

208. **Nikoh, N., N. K. Iwabe, K.-E., M. Ohno, T. Sugiyama, Y. Watanabe, K. Yasui, Z. Shicui, K. Hori, Y. Shimura, and T. Miyata.** 1997. An estimate of divergence of time of Parazoa and Eumetazoa and that of Cephalochordata and Vertebrata by aldolase and triose phosphate isomerase clocks. J. Mol. Evol. **45:**97-106.

209. **Nirenberg, H. I., and K. O'Donnell.** 1998. New *Fusarium* species and combinations within the *Gibberella fujikuroi* species complex. Mycologia. **90:**434-458.

210. **Normark, B. B.** 2000. Molecular systematics and evolution of the aphid family Lachnidae. Mol. Phylogenet. Evol. **14:**131-140.

211. **Nyrberg, A. M., and M. B. Cronhjort.** 1992. Intron evolution: a statistical comparison of two models. J. Theor. Biol. **157:**175-190.

212. **O'Donnell, K., and E. Cigelnik.** 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. Mol. Phylogenet. Evol. **7:**103-116.

213. **O'Donnell, K., E. Cigelnik, and H. I. Nirenberg.** 1998. Molecular systematics and phylogeography of the *Gibberella fujikuroi* species complex. Mycologia. **90:**465-493.

214. **O'Donnell, K., H. C. Kistler, E. Cigelnik, and R. C. Ploetz.** 1998. Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear, mitochondrial gene sequences. Proc. Natl. Acad. Sci. USA. **95:**2044-2049.

215. **O'Donnell, K., H. C. Kistler, B. K. Tacke, and H. H. Casper.** 2000. Gene genealogies reveal global phylogenetic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. Proc. Natl. Acad. Sci. USA. **97:**7905-7910.

216. **O'Donnell, K., H. I. Nirenberg, T. Aoki, and E. Cigelnik.** 2000. A multigene phylogeny of the *Gibberella fujikuroi* species complex: Detection of additional phylogenetically distinct species. Mycoscience. **41:**61-78.

217. **Ohno, S.** 1970. Why gene duplication? *In* S. Ohno (ed.), Evolution by gene duplication. Springer-Verlag, New York.

218. **Ohya, Y., and D. Botstein.** 1994. Structure-based systematic isolation of conditional-lethal mutations in the single yeast calmodulin gene. Genetics. **138:**1041-1054.

219. **Old, R. W., and H. R. Woodland.** 1984. Histone genes: not so simple after all. Cell. **38:**624-626.

220. **Orbach, M. J., E. B. Porro, and C. Yanofsky.** 1986. Cloning and characterization of the gene for β-tubulin from a benomyl-resistant mutant of *Neurospora crassa* and its use as a dominant selectable marker. Mol. Cell. Biol. **6:**2452-2461.

221. **Palumbi, S. R., and F. Cipriano.** 1998. Species identification using genetic tools: the value of nuclear and mitochondrial gene sequences in whale conservation. J. Hered. **89:**459-464.

222. **Panaccione, D. G., and R. M. Hanau.** 1990. Characterization of two divergent β-tubulin genes from *Colletotrichum graminicola*. Gene. **86:**163-170.

223. **Paquin, B., M.-J. Laforest, and F. Lang.** 1994. Interspecific transfer of mitochondrial genes in fungi and creation of a homologous hybrid gene. Proc. Natl. Acad. Sci. USA. **91:**11807-11810.

224. **Parsch, J., S. Tanda, and W. Stephan.** 1997. Site-directed mutations reveal long-range compensatory interactions in the *Adh* gene of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA. **94:**928-933.

225. **Percudani, R., and S. Ottenello.** 1999. Selection at the wobble position of codons read by the same tRNA in *Saccharomyces cerevisiae*. Mol. Biol. Evol. **16:**1752-1762.

226. **Philippe, H., and A. Germot.** 2000. Phylogeny of eukaryotes based on ribosomal RNA: long-branch attraction and models of sequence evolution. Mol. Biol. Evol. **17:**830-834.

227. **Puhler, G., H. Leffers, F. Gropp, P. Palm, H.-P. Klenk, F. Lottspeich, R. A. Garrett, and W. Zillig.** 1989. Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. Proc. Natl. Acad. Sci. USA. **86:**4569-4573.

228. **Quigley, F., H. Brinkmann, W. F. Martin, and R. Cerff.** 1989. Strong functional GC pressure in a light-regulated maize gene encoding subunit GAPA of chloroplast glyceraldehyde-3-phosphate dehydrogenase: implications for the evolution of GAPA pseudogenes. J. Mol. Evol. **29:**412-421.

229. **Quigley, F., W. F. Martin, and R. Cerff.** 1988. Intron conservation across the prokaryote-eukaryote boundary: Structure of the nuclear gene chloroplast glyceraldehyde-3-phosphate dehydrogenase form maize. Proc. Natl. Acad. Sci. USA. **85:**2672-2676.

230. **Radford, A.** 1993. A fungal phylogeny based upon orotidine 5'-monophosphate decarboxylase. J. Mol. Evol. **36:**389-395.

231. **Raha, D., K. Sen, and B. B. Biswas.** 1987. cDNA cloning of β-tubulin gene and organization of tubulin genes in *Vigna radiata* (mung bean) genome. Plant Mol. Biol. **9:**656-671.

232. **Recipon, H., R. Perasso, A. Addoutte, and F. Quetier.** 1992. ATP synthase subunit c/III/9 gene sequence as a tool for interkingdom and metaphytes molecular phylogeny. J. Mol. Evol. **34:**292-303.

233. **Reed, R. D., and F. A. H. Sperling.** 1999. Interaction of process partitions in phylogenetic analysis: an example from the swallowtail butterfly genus *Papilio*. Mol. Biol. Evol. **16:**286-297.

234. **Regier, J. C., B. M. Wiegman, R. F. Leclerc, and T. P. Friedlander.** 1994. Loss of phylogenetic information in chorion gene families of *Bombyx mori* by gene conversion.

Mol. Biol. Evol. **11:**72-87.

235. **Reiger, J. C., Q. Q. Fang, C. Mitter, R. S. Peigler, T. P. Friedlander, and M. A. Solis.** 1998. Evolution and phylogenetic utility of the *period* gene in Lepidoptera. Mol. Biol. Evol. **15:**1172-1182.

236. **Roger, A. J., P. J. Keeling, and W. F. Doolittle.** 1994. Introns, the broken transposons. Soc. Gen. Physiol. Ser. **49:**27-37.

237. **Roger, A. J., O. Sandblom, W. F. Doolittle, and H. Philippe.** 1999. An evolution of elongation factor 1α as a phylogenetic marker for eukaryotes. Mol. Biol. Evol. **16:**218-233.

238. **Roy, S. W., M. Nosaka, S. J. de Souza, and W. Gilbert.** 1999. Centripetal modules and ancient introns. Gene. **238:**85-91.

239. **Saloheimo, M., M.-L. Niku-Paavola, and J. K. C. Knowles.** 1991. Isolation and structural analysis of the laccase gene from the lignin-degrading fungus *Phlebia radiata.* J. Gen. Microbiol. **137:**1537-1544.

240. **Sanderson, M. J., and J. J. Doyle.** 1992. Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. Syst. Biol. **41:**4-17.

241. **Sanson, G. F. O. B., M.R.S.** 2000. Typing of *Candida glabrata* in clinical isolates by comparative sequence analysis of the cytochrome *c* oxidase subunit 2 gene distinguishes two clusters of strains associated with geographical sequence polymorphisms. J. Clin. Microbiol. **38:**227-235.

242. **Sburlati, A., and E. Cabib.** 1986. Chitin synthase 2, a presumptive participant in septum formation in *Saccharomyces cerevisiae*. J. Biol. Chem. **261:**15147-15152.

243. **Schardl, C. L., A. Leuchtmann, H.-F. Tsai, M. A. Collett, D. M. Watt, and D. B. Scott.** 1994. Origin of a fungal symbiont of perennial ryegrass by interspecific hybridization of a mutualist with ryegrass choke pathogen, *Epichloë typhina.* Genetics. **136:**1307-1317.

244. **Schimenti, J. C.** 1999. Mice and the role of unequal recombination in gene-family evolution. Am. J. Hum. Genet. **64:**40-45.

245. **Schutze, J., A. Krasko, M. R. Custodio, S. M. Efremova, I. M. Muller, and W. E. G. Muller.** 1999. Evolutionary relationships of metazoa within the eukaryotes based on molecular data from Porifera. Proc. R. Soc. Lon. B. **266:**63-73.

246. **Schwob, E., and R. P. Martin.** 1992. New yeast actin-like gene required late in the cell cycle. Nature. **355:**179-182.

247. **Seiffert, K. A., B. D. Wingfield, and M. J. Wingfield.** 1995. A critique of DNA sequence analysis in the taxonomy of filamentous Ascomycetes and ascomycetous anamorphs. Can. J. Bot. **73:**S760-S767.

248. **Senapathy, P.** 1986. Origin of eukaryotic introns: a hypothesis, based on codon distribution statistics in genes, and its implications. Proc. Natl. Acad. Sci. USA. **83:**2133-2137.

249. **Sharp, P. A., and C. B. Burge.** 1997. Classification of introns: U2-type or U12-type. Cell. **91:**875-879.

250. **Sharp, P. M., and A. T. Lloyd.** 1993. Regional base composition variation along yeast chromosome III: evolution of chromosome primary structure. Nucleic Acids Res. **21:**179-183.

251. **Shearer Jr, G.** 1995. Cloning and analysis of cDNA encoding an elongation factor 1α from the dimorphic fungus *Histoplasma capsulatum.* Gene. **161:**119-123.

252. **Sherwood, J. E., and S. C. Sommerville.** 1989. Molecular characterization of a β-tubulin gene from *Erysiphe graminis*. Phytopathology. **79:**1205.

253. **Shpaer, E. G.** 1985. The secondary structure of mRNAs form *Escherichia coli*: its possible role in increasing the accuracy of translation. Nucleic Acids Res. **13:**275-288.

254. **Sidow, A., and W. K. Thomas.** 1994. A molecular evolutionary framework for eukaryotic model organisms. Curr. Biol. **4:**596-603.

255. **Simpson, G. G., and W. Filipowicz.** 1996. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organization of the spliceosomal machinery. Plant Mol. Biol. **32:**1-41.

256. **Singh, G., P. S. Dyer, and A. M. Ashby.** 1999. Intra-specific and inter-specific conservation of mating-type genes from the discomycete plant-pathogenic fungi *Pyrenopeziza brassicae* and *Tapesia yallundae*. Curr. Genet. **36:**290-300.

257. **Smith, H. A., H. S. Allaudeen, H. Whitman, Y. Koltin, and J. A. Gorman.** 1988. Isolation and characterization of a β-tubulin gene from *Candida albicans*. Gene. **63:**53-63.

258. **Smith, T. L.** 1989. Disparate evolution of yeast and filamentous fungi indicated by phylogenetic analysis of glyceraldehyde-3-phosphate dehydrogenase genes. Proc. Natl. Acad. Sci. USA. **86:**7063-7066.

259. **Snustad, D. P., N. A. Haas, D. S. Kopczak, and C. D. Silflow.** 1992. The small genome of *Arabidopsis* contains at least nine expressed β-tubulin genes. Plant Cell. **4:**549-556.

260. **Sogin, M. L.** 1997. History assignment: when was the mitochondrion founded? Curr. Opin. Genet. Dev. **7:**792-799.

261. **Stedman, T. T., D. R. Butler, and G. A. Buck.** 1998. The HSP70 gene family in *Pneumocystis carinii*: molecular and phylogenetic characterization of cytoplasmic members. J. Eukaryot. Microbiol. **45:**589-599.

262. **Steenkamp, E., H. Britz, T. Coutinho, B. Wingfield, W. Marasas, and M. Wingfield.**

2000. Molecular characterization of *Fusarium subglutinans* associated with mango malformation. Mol. Plant Pathol. **1:**187-193.

263. **Steenkamp, E. T., B. D. Wingfield, T. A. Coutinho, M. J. Wingfield, and W. F. O. Marasas.** 1999. Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. Appl. Environ. Microbiol. **65:**3401-3406.

264. **Steenkamp, E. T., B. D. Wingfield, T. A. Coutinho, K. A. Zeller, M. J. Wingfield, W. F. O. Marasas, and J. F. Leslie.** 2000. PCR-based identification of *MAT-1* and *MAT-2* in the *Gibberella fujikuroi* species complex. Appl. Environ. Microbiol. **66:**4378-4382.

265. **Stiller, J. W., and B. D. Hall.** 1999. Long-branch attraction and the rDNA model of early eukaryotic evolution. Mol. Biol. Evol. **16:**1270-1279.

266. **Sullivan, K. F.** 1988. Structure and utilization of tubulin genes. Annu. Rev. Cell Biol. **4:**687-716.

267. **Swofford, D. L.** 1998. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0*. Sinauer Associates, Sunderland, Massachusetts.

268. **Swofford, D. L., and G. J. Olsen.** 1990. Phylogeny reconstruction, p. 411-501. *In* D. M. Hillis and G. Moritz (ed.), Molecular Systematics. Sinaur Associates, Inc., Sunderland, Massachusetts, USA.

269. **Tarn, W.-Y., and J. A. Steitz.** 1997. Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. Trends Biochem. Sci. **22:**132-137.

270. **Taylor, J. W., D. M. Geiser, A. Burt, and V. Koufopanou.** 1999. The evolutionary biology and population genetics underlying fungal strain typing. Clin. Microbiol. Rev. **12:**126-146.

271. **Taylor, J. W., D. J. Jacobson, and M. C. Fisher.** 1999. The evolution of asexual fungi: Reproduction, speciation and classification. Annu. Rev. Phytopathol. **37:**197-246.

272. **Thatcher, T. H., and M. A. Gorocsky.** 1994. Phylogenetic analysis of the core histone H2A, H2B, H3 and H4. Nucleic Acids Res. **22:**174-179.

273. **Thatcher, T. H., J. MacGaffey, J. Bowen, S. Horowitz, D. L. Shapiro, and M. A. Gorovsky.** 1994. Independent evolutionary origin of histone H3.3-like variants of animals and *Tetrahymena*. Nucleic Acids Res. **22:**180-186.

274. **Thomas, G. H., E. C. Newbern, C. C. Korte, M. A. Bales, S. V. Muse, A. G. Clark, and D. P. Kiehart.** 1997. Intragenic duplication and divergence in the spectrin superfamily of proteins. Mol. Biol. Evol. **14:**1285-1295.

275. **Thomas, W. K., and A. C. Wilson.** 1991. Mode and tempo of molecular evolution in the nematode *Caenorhabditis*: cytochrome oxidase II and calmodulin sequences. Genetics. **128:**269-279.

276. **Tourasse, N. J., and M. Gouy.** 1999. Accounting for evolutionary rate variation among sequence sites consistently changes universal phylogenies deduced from rRNA and protein-coding genes. Mol. Phylogenet. Evol. **13:**159-168.

277. **Trotman, C. N. A.** 1998. Introns-early: slipping lately? Trends Genet. **14:**132-134.

278. **Tsai, H.-F., J.-S. Liu, C. Staben, M. J. Christensen, G. C. M. Latch, M. R. Siegel, and C. L. Schardl.** 1994. Evolutionary diversification of fungal endophytes of tall fescue grass by hybridization with *Epichloë* species. Proc. Natl. Acad. Sci. USA. **91:**2542-2546.

279. **Turgeon, B. G.** 1998. Application of mating type gene technology to problems in fungal biology. Annu. Rev. Phytopathol. **36:**115-137.

280. **van Daal, A., E. M. White, S. C. R. Elgin, and M. A. Gorovsky.** 1990. Conservation of intron position indicates separation of major and variant H2As is an early event in the evolution of eukaryotes. J. Mol. Evol. **30:**449-455.

281. **van den Hoff, M., A. Jonker, J. J. Beintema, and W. H. Lamers.** 1995. Evolutionary relationships of the carbamoylphosphate synthase genes. J. Mol. Evol. **41:**813-832.

282. **Verdoes, J. C., J. Wery, T. Boekhout, and A. J. J. Van Ooyen.** 1997. Molecular characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of *Phaffia rhodozyma*. Yeast. **13:**1231-1242.

283. **Viaud, M., Y. Couteaudier, and G. Riba.** 1998. Molecular analysis of hypervirulent somatic hybrids of the entomopathogenic fungi *Beauveria bassiana* and *Beauveria sulfurescens*. Appl. Environ. Microbiol. **64:**88-93.

284. **Villemur, R., N. Haas, C. M. Joyce, D. P. Snustad, and C. D. Silflow.** 1993. Characterization of four new β-tubulin genes and their expression during male flower development in maize (*Zea mays* L.). Plant Mol. Biol. **24:**295-315.

285. **Von Dorum, M., and M. Ruvolo.** 1999. Phylogenetic relationships of the new world monkeys (primates, Platyrrhini) based on nuclear G6PD DNA sequences. Mol. Phylogenet. Evol. **11:**459-476.

286. **Vossbrinck, C. R., J. V. Maddox, S. Friedman, B. A. Debrunner-Vossbrinck, and C. R. Woese.** 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature. **326:**411-414.

287. **Vossbrinck, C. R., and C. R. Woese.** 1986. Eukaryotic ribosomes that lack a 5.8S RNA. Nature. **320:**287-288.

288. **Wang, L., K. Yokoyama, M. Miyaji, and K. Nishimura.** 1998. The identification and phylogenetic relationship of pathogenic species of *Aspergillus* based on the mitochondrial cytochrome b gene. Med. Mycol. **36:**153-164.

289. **Watson, J. D., N. M. Hopkins, J. W. Roberts, J. A. Steitz, and A. M. Weiner.** 1987.

Molecular biology of the gene, fourth ed. Benjamin/Cummings, Menlo Park, California.

290. **Weerakoon, N. D., J. K. Roberts, L. P. Lehnen, J. M. Wilkinson, J. S. Marshall, and A. R. Hardham.** 1998. Isolation and characterization of the single β-tubulin gene in *Phytophthora cinnamomi*. Mycologia. **90:**85-95.

291. **Weiss, L. M., T. D. Edlind, C. R. Vossbrinck, and T. Hashimoto.** 1999. Microsporidian molecular phylogeny: the fungal connection. J. Eukaryot. Microbiol. **46:**17S-18S.

292. **Wells, D., W. Bains, and L. Kedes.** 1986. Codon usage in histone gene families of higher eukaryotes reflect functional rather that phylogenetic relationships. J. Mol. Evol. **23:**224-241.

293. **Wendel, J. F., A. Schnabel, and T. Seelanan.** 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. USA. **92:**280-284.

294. **Werman, S. D., M. S. Springer, and R. J. Britten.** 1990. Nucleic acids I: DNA-DNA hybridization, p. 204-249. *In* D. M. Hillis and C. Moritz (ed.), Molecular Systematics. Sinauer Associates, Inc., Sunderland, Massachusetts.

295. **Wery, J., M. J. M. Dalderum, J. Ter Linde, T. Boekhouts, and A. J. J. Van Ooyen.** 1996. Structural and phylogenetic analysis of the actin gene from the yeast *Phaffia rhodozyma*. Yeast. **12:**641-651.

296. **Wesseling, J. G., M. A. Smits, and J. G. G. Schoenmakers.** 1988. Extremely diverged actin proteins in *Plasmodium falciparum*. Mol. Biochem. Parasitol. **30:**143-154.

297. **Whitehead, M. P., M. T. Shieh, T. E. Cleveland, J. W. Cary, and R. A. Dean.** 1995. Isolation and characterization of polygalacturonase genes (*pecA* and *pecB*) from *Aspergillus flavus*. Appl. Environ. Microbiol. **61:**3316-3322.

298. **Wirth, T., R. Le Guellec, and M. Veuille.** 1999. Directional substitution and evolution of nucleotide content in the cytochrome oxidase II gene in earwigs (Dermapteran insects). Mol. Biol. Evol. **16:**1645-1653.

299. **Woese, C. R., J. Gibson, and G. E. Fox.** 1980. Do genealogical patterns in purple photosynthetic bacteria reflect interspecific gene transfer? Nature. **283:**212-214.

300. **Wojtas, K. M., L. Von Kalm, J. R. Weaver, and D. T. Sullivan.** 1992. The evolution of duplicate glyceraldehyde-3-phosphate dehydrogenase genes in *Drosophila*. Genetics. **132:**789-797.

301. **Xoconostle-Cazares, B., C. A. Specht, P. W. Robbins, Y. Liu, C. Leon, and J. Ruiz-Herrera.** 1997. *Umchs5*, a gene encoding for a class IV chitin synthase in *Ustilago maydis*. Fungal Genet. Biol. **22:**199-208.

302. **Yan, K., and M. B. Dickman.** 1996. Isolation of a β-tubulin gene from *Fusarium*

moniliforme that confers cold-sensitive benomyl resistance. Appl. Environ. Microbiol. **62:**3053-3056.

303. **Yang, T., G. Segal, S. Abbo, M. Feldman, and H. Fromm.** 1996. Characterization of the calmodulin gene family in wheat: structure, chromosomal location, and evolutionary aspects. Mol. Gen. Genet. **252:**648-694.

304. **Yarden, O., and T. Katan.** 1993. Mutations leading to substitutions at amino acids 198 and 200 of beta-tubulin that correlate with benomyl resistance phenotypes of field strains of *Botrytis cinerea*. Phytopathology. **83:**1478-1483.

305. **Yaver, D. S., F. Zu, E. J. Golightly, K. M. Brown, S. H. Brown, M. W. Rey, P. Schneider, T. Halkier, K. Mondorf, and H. Dalboge.** 1996. Purification, characterization, molecular cloning, and expression of two laccase genes from the white rot Basidiomycete *Trametes villosa.* Appl. Environ. Microbiol. **62:**834-841.

306. **Young, R. A., and R. W. Davis.** 1983. Yeast RNA polymerase II genes: isolation with antibody probes. Science. **222:**778-782.

307. **Yun, S.-H., M. L. Berbee, O. C. Yoder, and B. G. Turgeon.** 1999. Evolution of the self-fertile reproductive life style from self-sterile ancestors. Proc. Natl. Acad. Sci. USA. **96:**5592-5597.

308. **Zhao, J., and H. S. Kwan.** 1999. Characterization, molecular cloning, and differential expression analysis of laccase genes from the edible mushroom *Lentinula edodes*. Appl. Environ. Microbiol. **65:**4908-4913.

309. **Zhou, J., A. Kilian, R. L. Warner, and A. Kleinhofs.** 1995. Variation of nitrate reductase genes in selected grass species. Genome. **38:**919-927.

310. **Zhou, J., and A. Kleinhofs.** 1996. Molecular evolution of nitrate reductase genes. J. Mol. Evol. **42:**432-442.

311. **Zimmer, E. A., S. L. Martin, S. M. Beverley, Y. W. Kan, and A. C. Wilson.** 1980. Rapid duplication and loss of genes coding for the α–chains of hemoglobin. Proc. Natl. Acad. Sci. USA. **77:**2158-2162.

7 TABLES

Table 1. Consensus sequences for the 5'- and 3'-splice sites, as well as the putative branch sites for the different types of spliceosomal introns (249).

Intron type	5'-Splice site[1]	Branch site[1]	3'-splice site[1]	Dist.[2]	Freq.[3]
U2-type GT-AG intron	KAG|GTRAGTooo $^{71}_{60}{}^{81}$|●●$^{94}_{71}{}^{84}_{46}$---	ooooCTRAYooo ----$^{76}_{91}{}^{81}_{99}{}^{83}$---	YYYYYYYYYYYYoTAG|Goo $^{71}_{76}{}^{78}_{79}{}^{80}_{80}{}^{80}_{80}{}^{80}_{82}{}^{88}_{87}$-^{96}●●|$_{52}$--	10-50+	Common
U12-type GT-AG intron	ooT|GTATCCTTT --$_{54}$|●●●●●●$^{93}_{96}{}^{70}$	TTCCTTAACYoY $^{52}_{74}{}^{81}_{96}{}^{89}$●$^{89}_{86}{}^{96}_{68}$-^{71}	ooooooooooooYAG|Aoo ----------^{96}●●|^{54}--	11-20	1/600
U12-type AT-AC intron	ooo|ATATCCTTT ---|●●●●●●$^{96}_{93}{}^{70}$	TTCCTTRACYCY $^{50}_{65}{}^{95}_{90}{}^{95}$●●●$^{86}_{95}{}^{75}_{85}$	ooooooooooooYAC|ooo ----------^{95}●^{96}|---	10-15	1/5000 to 1/10000
U2-type AT-AC intron	ARG|ATAAGTooo ●●●|●●$^{86}_{86}$●^{71}---	unknown unknown	ooooooooooooYAC|ooo ----------^{86}●●|---	n/a	unknown

[1] Splice junctions are indicated by vertical lines (|). Positions with no clear consensus sequences are indicated by open circles (o), whereas R indicates either purine (A or G), Y either pyrimidine (C or T) and K either A or C. Below each of the consensus sequences the degrees of conservation (%) are indicated. Black dots (•) indicate 100 % conservation in all the known sequences and horizontal lines (-) indicate the absence of strong conservation.

[2] Dist. = distance in nucleotide bases from the putative branch site to the 3'-splice junction. n/a = not available.

[3] The approximate frequency at which the intron type occurs (249).

Table 2. The taxonomic level and specific problems associated with the use of selected nuclear and mitochondrial protein-coding genes in taxonomic and phylogenetic studies.

Protein	GenBank hits[1]		Taxonomic levels studied	Specific problems experienced[2]
	Fungi	Total		
β-tubulin	802	5810	Species → Eukaryotic kingdoms (5, 11, 63, 75, 92, 142-145, 212, 213, 215, 216, 243, 245, 262, 278, 283, 290)	-Multi-copy (37, 101, 109, 112, 143, 167, 187, 222, 231, 259, 284) -Lateral transfer (260) -Distortion and lack of resolution of phylogenetic relationships due to lineage and site-specific accelerated evolutionary rates (144, 145)
Elongation factor 1α	283	15771	Individual → Eukaryotic and prokaryotic kingdoms (5, 10, 11, 42, 53, 135, 139, 149, 152, 157, 201, 207, 210, 215, 216, 237)	-Multi-copy (139, 206, 233, 237) -Distortion and lack of resolution of phylogenetic relationships due to lineage and site-specific accelerated evolutionary rates (10, 11, 123, 139, 233, 237)
Chitin synthases	246	272	Individuals → Species (140, 154, 191, 283)	-Multi-copy (29, 192, 193, 198, 199, 242, 301)
Actin	240	29930	Species → Eukaryotic kingdoms (4, 11, 19, 68, 87, 117, 157, 207, 221, 295)	-Multi-copy (69, 189, 296, 81) -Distortion of phylogenetic relationships because of lineage specific nucleotide substitution rates (1, 19, 68, 117)
Glyceraldehyde-3-phosphate dehydrogenase	218	5490	Species (2, 11, 18)	-Multi-copy (167, 300) -Distorted inter- and intrakingdom relationships because of lateral gene transfer (32, 45, 84, 118, 169, 258, 282, 195)
Calmodulin	133	22946	Individual → Eukaryotic kingdoms (11, 42, 92, 216, 245)	-Multi-copy (162, 303)
Mating type idiomorphs	129	123	Species[3] (256, 264, 279, 307)	-Sequence information from isolates with opposite mating types is not combinable (264)
Histone H3	104	4773	Species → Eukaryotic kingdoms (63, 179, 262, 263, 272, 273, 292)	-Multi-copy (79, 114, 179, 197, 219, 272, 273, 292)
Laccases	92	482	N/u	-Multi-copy (76, 95, 196, 305, 308)
70-kDa Heat-shock-protein (HSP70)	90	1140	Class → Eukaryotic and prokaryotic kingdoms (11, 27, 77, 94, 122, 157, 207, 245, 261)	-Multi-copy (261)

Table 2. Continued.

Protein	GenBank hits[1] Fungi	GenBank hits[1] Total	Taxonomic levels studied	Specific problems experienced[2]
Glucoamylase	69	128	Species[3] (92)	-Multi-copy (62)
Polygalacturonase	63	439	Species[3] (92)	-Multi-copy (40, 297)
O-methyltransferase	53	796	Species[3] (92)	
α-tubulin	51	6738	Species → Eukaryotic kingdoms (142, 144, 145)	-Multi-copy (143, 153, 167) -Distortion and lack of resolution of phylogenetic relationships due to lineage and site-specific accelerated evolutionary rates (144, 145)
ATPase subunit 6	50	754	Species → Class (11, 155, 207)	-Presence of hybrid genes (223) -Interspecific lateral transfer (223)
Phosphate permease	50	4	Species[3] (215)	
Glutamate dehydrogenase	49	1107	N/u	-Multi-copy (14)
ATPase subunit 9	46	93	Species → Eukaryotic and prokaryotic kingdoms (207, 232)	-Multi-copy (232)
Elongation factor 2	45	1970	Species → Eukaryotic and prokaryotic kingdoms (116, 123, 135, 139, 157, 201, 207)	-Multi-copy (139) -Distortion and lack of resolution of phylogenetic relationships due to lineage and site-specific accelerated evolutionary rates (123)
Histone H4	39	1287	Species → Eukaryotic kingdoms (63, 154, 179, 272, 283)	-Multi-copy (13, 79, 114, 186, 197, 219, 292)
γ-tubulin	39	428	Species → Eukaryotic kingdoms (142, 144, 145)	-Distortion and lack of resolution of phylogenetic relationships due to lineage and site-specific accelerated evolutionary rates (145)
RNA polymerase II second largest subunit	33	236	Species → Eukaryotic and prokaryotic kingdoms (56, 133, 157, 170, 207, 254, 276)	
Histone H2A	33	2701	Species → Eukaryotic kingdoms (12, 272, 280)	-Multi-copy (79, 114, 148, 186, 197, 219, 272)

Table 2. Continued.

Protein	GenBank hits[1] Fungi	Total	Taxonomic levels studied	Specific problems experienced[2]
Phosphoglycerate kinase	31	1360	Species → Eukaryotic kingdoms (11, 48)	-Multi-copy (167)
Nitrate reductase	31	338	Species → Eukaryotic kingdoms (92, 154, 283, 309, 310)	-Multi-copy (309)
Histone H2B	24	1028	Species → Eukaryotic kingdoms (12, 272)	-Multi-copy (79, 105, 114, 186, 197, 219)
Adenylate kinase	20	1062	N/u	-Multi-copy (89)
Eukaryotic initiation factors	20	2045	Species → Eukaryotic kingdoms (157)	
Malate dehydrogenase	20	2184	Species → Eukaryotic kingdoms (11, 134, 180)	
Glucose-6-phosphate dehydrogenase	15	594	Species[3] (92, 285)	
Serine proteinase	15	630	Species[3] (154)	
Orotidine 5'-monophosphate decarboxylase	15	92	Species[3] (154, 230)	
Serine-threonine kinase domain of protein kinases	14	257	Class → Eukaryotic kingdoms (156, 245)	-Multi-copy (156)
RNA polymerase II largest subunit (*RPB1*)	12	268	Species → Eukaryotic and prokaryotic kingdoms (15, 51, 123, 133, 157, 207, 227, 276)	
Cytochrome subunits	10	449	Individuals → Eukaryotic and prokaryotic kingdoms (3, 86, 102, 138, 146, 155, 175, 203, 233, 241, 288)	-Multi-copy (167) -Distortion of phylogenetic relationships due to lineage specific accelerated evolutionary rates (155, 233, 298)

Table 2. Continued.

Protein	GenBank hits[1] Fungi	Total	Taxonomic levels studied	Specific problems experienced[2]
Dihydroorotase	10	100	Eukaryotic and prokaryotic kingdoms (281)	-Multi-copy (281)
Trichothecene 3-*O*-acetyltransferase	9	9	Species[3] (215)	
UTP-ammonia ligase	8	2	Species[3] (215)	
Protease I	7	60	Species[3] (154)	
Triose phosphate isomerase	6	158	Class → Eukaryotic kingdoms (77, 207, 208)	-Multi-copy(167)
Valyl-tRNA synthetase	6	494	Class → Eukaryotic kingdoms (77)	
Aldehyde reductase	4	216	Class → Eukaryotic kingdoms (77, 147)	

[1] Total number of GenBank entries and only those associated with fungi (Updated November 25th 2000).
[2] Multi-copy refers to the presence of either non-orthologous or pseudo genes.
[3] Use of the protein-coding gene sequence has only been tested at the level indicated by these authors. N/u = gene not used for phylogeny.

8 FIGURES

Figure 1. Intron distribution matrix for selected β-tubulin genes in eukaryotes [modified from Dibb and Newman (58) and Liaud et al. (168)]. Intron positions and phases are indicated as defined by Dibb and Newman (58). The intron at position 5, for example, is phase 0, which means that codon number 5 is preceded by an intron. Phase 0 introns are indicated in blue by ⓪. The intron at position 10 is phase 1, which means that codon number 10 is split by an intron after the first base. Phase 1 introns are indicated in red by ①. The intron at position 21 is phase 2, which means that codon number 21 are split by an intron after the second base. Phase 2 introns are indicated in black by ②. A vertical line (-) indicates the absence of an intron. GenBank accession numbers are indicated in parentheses. For the correct intron positions in *Trichoderma viride* (GenBank accession number Z15054), a value of two should be added to all positions, since this strain has an insertion of two codon residues at the beginning of the sequence. All the intron positions after number 133 in *Histoplasma capsulatum* (GenBank accession number AH003038) should be increased by one, because of a single residue insertion. All the intron positions after number 350 in the human TUB4Q β-tubulin gene (GenBank accession number U83668) should be decreased by one because of a residue deletion.

Organism	Intron postion																																									
	5	7	9	10	12	13	16	17	20	21	22	35	54	56	57	62	94	106	133	176	183	192	207	214	223	230	247	259	277	282	294	314	318	326	350	389	410	429	430	431	439
FUNGI																																									
Gibberella zeae (UU34436, U34462, U34491)	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Gibberella fujikuroi (U27303)	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Neotyphodium coenophialum (X56847)	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Acremonium chrysogenum (X72789)	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Trichoderma viride (Z15055) tub2	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Epichloe typhina (X52616)	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Hypomyces odoratus (Y12256)	0	-	-	-	-	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mycosphaerella pini (AF044975)	0	-	-	-	-	0	-	-	-	-	-	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-
Colletotrichum gloeosporioides (U14138)	0	-	-	-	-	0	-	-	-	2	-	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Botryotinia fuckeliana (Z69263, U27198) tubA, benA	0	-	-	-	-	0	-	-	-	2	-	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Neurospora crassa (M13630)	0	-	-	-	-	0	-	-	-	2	-	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Penicillium digitatum (D78154)	0	-	-	-	-	0	-	-	-	2	-	2	0	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
Aspergillus nidulans (M17519) benA	0	-	-	-	-	0	-	-	-	2	-	2	0	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	2
Aspergillus parasiticus (L49386)	0	-	-	-	-	0	-	-	-	2	-	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	2
Histoplasma capsulatum (AH003038)	0	-	-	-	-	0	-	-	-	2	-	-	0	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Venturia inaequalis (M97951)	0	-	-	-	-	0	-	-	-	2	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Schizosaccharomyces pombe (AF042827)	0	-	-	-	-	0	-	-	-	2	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-
Blumeria graminis (X51326, M30927)	0	-	-	-	-	0	-	-	-	1	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
Trichoderma viride (Z15054) tub1	0	-	-	-	-	0	-	-	-	-	0	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Aspergillus flavus (M38265)	0	-	-	-	-	0	-	-	-	-	0	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	2
Aspergillus nidulans (M17520) tubC	0	-	-	-	-	0	-	-	-	2	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Galactomyces geotrichum (S69624)	0	-	-	-	-	0	-	-	-	2	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Rhynchosporium secalis (X81046)	0	-	-	-	-	-	-	2	-	2	-	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
Pneumocystis carinii (L05466)	0	-	-	1	-	-	-	0	-	0	-	2	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
Candida albicans (M19398)	0	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pleurotus sajor-caju (AF132911)	-	-	-	-	0	-	-	1	-	-	-	-	-	-	-	-	-	1	1	-	0	-	-	-	-	-	-	-	-	-	-	0	-	-	-	0	-	-	2	-
Schizophyllum commune (X63372, S45884)	-	-	-	-	0	-	-	1	-	-	-	-	-	-	-	-	-	1	2	-	0	-	-	-	-	-	-	1	-	-	-	0	-	-	-	0	-	-	-	-
Coprinus cinereus (AB000116)	-	-	-	-	0	-	-	1	-	-	-	-	-	-	-	-	-	1	1	-	0	-	-	-	-	-	-	1	-	-	-	0	-	-	-	-	-	-	2	-
METAZOA																																									
Caenorhabditis elegans (X15242)	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	1	-	-	-	-	-
Caenorhabditis briggsae (U55260)	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	1	-	-	-	-	-
Onchocerca gibsoni (X79930)	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	2	1	-	0	0	-	-	-	2	-	-	0	-	1	-	-	-	-	0	-	1	-	0	-	-	-
Trichuris trichiura (AF034219)	-	-	-	-	-	-	-	-	0	-	-	-	-	0	-	-	1	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-	1	-	-	-	-	-
Brugia pahangi (M36380)	-	-	-	-	-	-	-	-	0	-	-	-	1	-	-	-	-	-	0	0	-	-	-	-	-	1	-	-	-	-	0	-	-	0	-	-	-	-	-	-	-
Onchocerca volvulus (AF019886)	-	-	-	-	-	-	-	-	0	-	-	-	1	-	-	-	-	-	0	0	-	-	-	-	-	1	-	-	-	-	0	-	-	0	-	1	-	-	-	-	-
Cylicocyclus nassatus (AF181093)	-	-	-	-	-	-	-	-	0	-	-	-	1	-	-	-	-	-	0	-	-	-	-	-	-	1	-	-	-	-	0	-	-	0	-	1	-	-	0	-	-
Homo sapiens (X02344, U83668, X00734) HSTUBB2, TUB4Q, BETA	-	-	-	-	-	-	-	-	0	-	-	-	1	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VIRIDIPLANTAE																																									
Zea mays (X52878) tub1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Daucus carota (U63927) tub2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pisum sativum (X54844) gtub1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Lupinus albus (X70184, U47660) tub1, tub2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glycine max (M21296, M21297) s-beta-1, s-beta-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis thaliana (M84706) tub9	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arabidopsis thaliana (AF049870, M84700, M84701, M21415, M84702, M84703, M84703, M84705) tub1 to tub8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Polytomella agilis (M33372, M33371) tub1, tub3	-	0	-	-	-	-	0	-	-	-	-	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Polytomella agilis (M33373) tub2	-	-	-	-	-	-	0	-	-	-	-	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Volvox carteri (X12855) tub1	-	-	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Chlamydomonas incerta (AF001379)	-	-	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Chlamydomonas reinhardtii (M10064, M25918, K03281, M25919) tub1, tub2	-	-	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

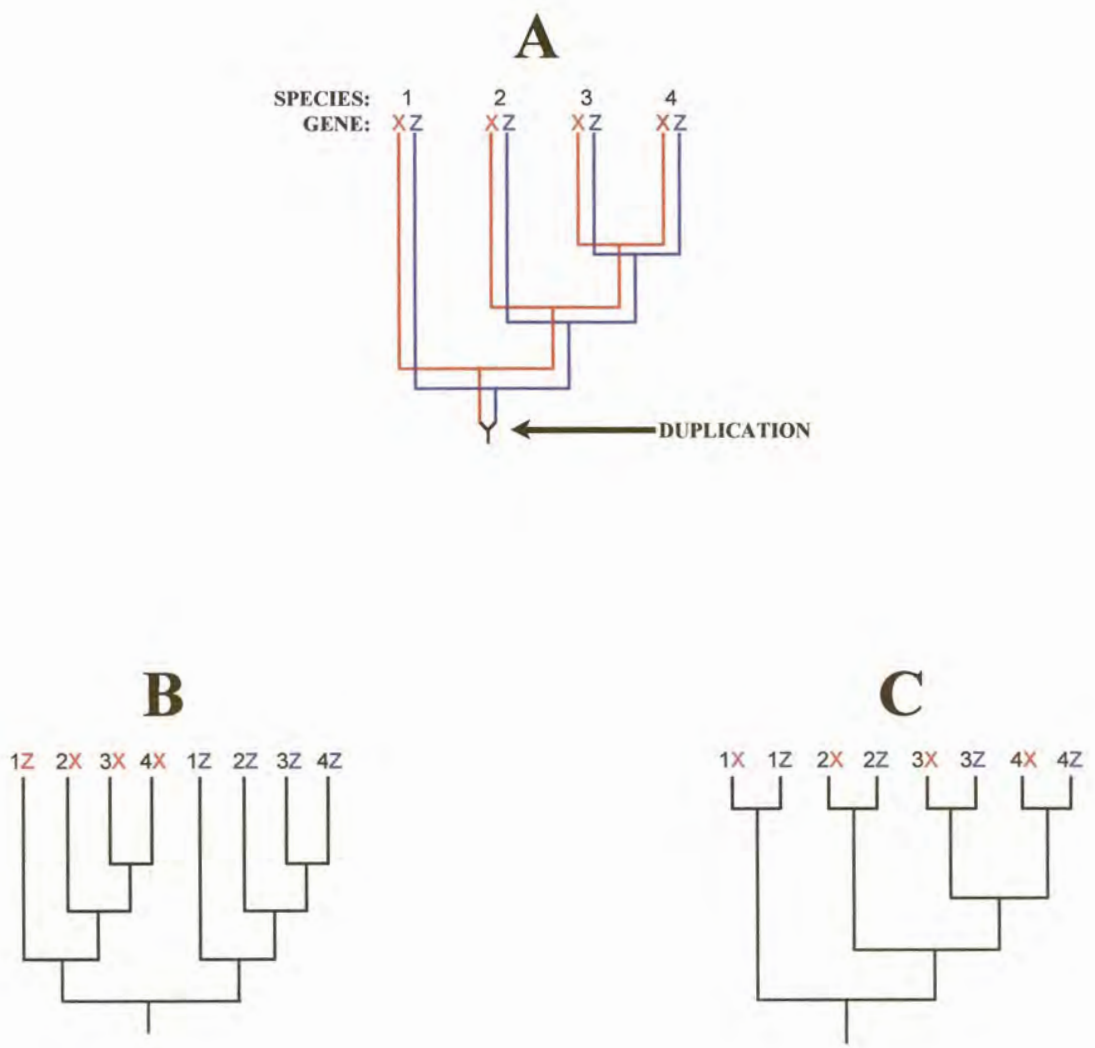

Figure 2. Evolutionary relationships among the hypothetical species 1, 2, 3, and 4 (240). A gene duplication event prior to the radiation of these species gave rise to a simple multigene family consisting of two genes (X and Z). **A:** The 'true phylogeny' reflecting the orthologous relationships among, for example, 1X, 2X, 3X and 4X, as well as the paralogous relationships between 1X and 1Z. **B:** The evolutionary relationships among the different genes and species prior to, or in the absence of concerted evolution. The tree is drawn in the standard output format of phylogenetic analyses software such as PAUP*4.0b (267). **C:** The evolutionary relationships among the different genes and species after, or in the presence of high levels of concerted evolution. The tree is drawn in the standard output format of phylogenetic analysis software.