Gene Genealogies, Cryptic Species, and Molecular Evolution in the Human Pathogen *Coccidioides immitis* and Relatives (Ascomycota, Onygenales)

Vassiliki Koufopanou,[1] *Austin Burt,*[1] *Timothy Szaro, and John W. Taylor*

Department of Plant and Microbial Biology, University of California at Berkeley

Previous genealogical analyses of population structure in *Coccidioides immitis* revealed the presence of two cryptic and sexual species in this pathogenic fungus but did not clarify their origin and relationships with respect to other taxa. By combining the *C. immitis* data with those of two of its closest relatives, the free-living saprophytes *Auxarthron zuffianum* and *Uncinocarpus reesii,* we show that the *C. immitis* species complex is monophyletic, indicating a single origin of pathogenicity. Cryptic species also were found in both *A. zuffianum* and *U. reesii,* indicating that they can be found in both pathogenic and free-living fungi. Our study, together with a few others, indicates that the current list of known fungal species might be augmented by a factor of at least two. However, at least in the *C. immitis, A. zuffianum,* and *U. reesii* complexes, cryptic species represent subdivisions at the tips of deep monophyletic clades and thus well within the existing framework of generic classification. An analysis of silent and expressed divergence and polymorphism values between and within the taxa identified by genealogical concordance did not reveal faster evolution in *C. immitis* as a consequence of adaptation to the pathogenic habit, nor did it show positive Darwinian evolution in a region of a dioxygenase gene (*tcrP* gene coding for 4-HPPD) known to cause antigenic responses in humans. Instead, the data suggested relative stasis, indicative of purifying selection against mostly deleterious mutations. Two introns in the same gene fragment were considerably more divergent than exons and were unalignable between species complexes but had very low polymorphism within taxa.

Introduction

Coccidioides immitis is a pathogenic fungus responsible for the recent epidemics of coccidioidomycosis (Valley fever) in California (Pappagianis et al. 1993; Pappagianis 1994). It has a dimorphic life cycle, living as a hyphal saprobe in the soil or as a unicellular systemic pathogen of mammals, and like many other fungi, no sexual state has ever been associated with it (Rippon 1988). Recently, however, analyses using molecular markers revealed a recombinant genetic structure in *C. immitis* isolated from humans in a single locality, implying that it does have sex occasionally and that a sexual stage ought to be found (Burt et al. 1996; Fisher et al. 2000*b*). Furthermore, genealogical concordance analysis of a global sample of isolates has indicated that it is not panmictic throughout its entire range, but instead consists of two reproductively isolated populations, one based in California and the other based outside of California, in the southwestern United States (Koufopanou, Burt, and Taylor 1997, 1998). Within the non-Californian population, there is significant geographic differentiation on the scale of 1,200 km (Burt et al. 1997; Fisher et al. 2000*a*). Similar analyses using molecular markers have also been successful in delineating mixis and reproductive isolation in other pathogenic fungi, including the human pathogens *Histoplasma capsulatum, Aspergillus flavus, Candida albicans,* and *Cryptococcus neoformans* (Sullivan et al. 1995; Carter et al. 1996; Geiser, Pitt, and Taylor 1998; Kasuga, Taylor, and White 1999; reviewed in Taylor et al. 1999; Burt, Koufopanou, and Taylor 2000) and the wheat pest *Fusarium graminearum* (O'Donnell et al. 2000).

Here, we extended the previous analyses of *C. immitis* by including isolates of two of the closest relatives, the nonpathogenic saprophytes *Auxarthron zuffianum* and *Uncinocarpus reesii* (Bowman and Taylor 1993; Pan, Sigler, and Cole 1994; Bowman, White, and Taylor 1996). Although these two species have also been collected from the lungs of rodents, they seem to be only transient and apparently harmless inhabitants of animals, and their life cycle does not include the production of spherules or endospores, stages which are presumed adaptations for the infective process (Sigler and Carmichael 1976). Unlike *C. immitis,* both *A. zuffianum* and *U. reesii* have well-identified sexual stages in their life cycles, and a recombinant population structure is therefore expected. We analyzed fragments of three nuclear genes: (1) *CHS1,* coding for chitin synthase, which is responsible for the synthesis of chitin, a major component of the fungal cell wall (Pan, Sigler, and Cole 1994); (2) *tcrP,* coding for a human T-cell reactive protein of *C. immitis* (this protein is homologous to 4-hydroxyphenyl-pyruvate dioxygenase (4-HPPD) proteins and mammalian F antigens; Wyckoff et al. 1995); and (3) *pyrG,* coding for orotidine 5′-monophosphate decarboxylase (OMPD), which catalyses a step in pyrimidine biosynthesis (Radford 1993).

By combining global samples of the three species, we aimed to root the relationships within *C. immitis* and determine whether the two taxa previously identified were both monophyletic or one was paraphyletic, still segregating ancestral alleles (Neigel and Avise 1986; Avise and Ball 1990). A consistent finding that one taxon was paraphyletic across a number of loci would imply a larger population size in that taxon (Hey and Kliman 1993). Data on the closest relatives would also help establish whether *C. immitis* as a whole was monophy-

[1] Present address: Department of Biology, Imperial College at Silwood Park, Ascot, Berks, United Kingdom.

Key words: gene genealogies, concordance, outgroups, cryptic species, molecular evolution, human pathogen, *Coccidioides immitis*, *Auxarthron zuffianum*, *Uncinocarpus reesii*, Ascomycota, Onygenales.

Address for correspondence and reprints: Vassiliki Koufopanou, Department of Biology, Imperial College at Silwood Park, Ascot, Berks SL5 7PY, United Kingdom. E-mail: v.koufopanou@ic.ac.uk.

Mol. Biol. Evol. 18(7):1246–1258. 2001
© 2001 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

letic, which would suggest a single origin of pathogenicity. Previous single-isolate analysis of 11 pathogenic and nonpathogenic relatives of *C. immitis* has indicated that the capacity to infect humans has arisen multiple times throughout the evolutionary history of the Onygenales (Bowman, White, and Taylor 1996). Second, we wanted to establish whether the finding of two cryptic species in *C. immitis* was a feature peculiar to it, perhaps due to its pathogenic nature leading to faster evolution and differentiation into several host niches, or whether cryptic species were a general feature of fungal species, including free-living ones. Finally, we wished to determine if adaptation to a new life style, that of being a pathogen, had led to positive Darwinian selection and increased rates of evolution, both for proteins in general and, more specifically, in gene regions known to produce antigenic responses in humans. For this purpose, we analyzed the numbers of silent and expressed changes within and between reproductively isolated taxa. The ratio $Q = (D/P)_E/(D/P)_S$ of between-taxa divergence (D) to within-taxon polymorphism (P) for expressed and silent substitutions (E and S, respectively) is a well accepted measure for detecting positive, adaptive evolution, with the neutral expectation being that $Q = 1$ (McDonald and Kreitman 1991; Charlesworth 1994; Hudson 1996). Positive selection, on the other hand, would lead to an excess of expressed between-taxa divergence ($Q > 1$) due to rapid fixation of favorable amino acid changes in different taxa, resulting from diversification and adaptation to new environments. A number of surface antigens of parasites or viruses have been shown to be positively selected, evolving rapidly by selective fixation of amino acids thought to be critical in the host-parasite interaction (Sphaer and Mullins 1993; Endo, Ikeo, and Gojobori 1996). We therefore expected the dioxygenase fragment containing the T-cell reactive site to be fast-evolving and under positive selection in *C. immitis.*

Materials and Methods

We analyzed five isolates from *A. zuffianum* and 10 from *U. reesii,* covering the known geographic range of the two species (table 1). All isolates were obtained from the Microfungus Collection of the University of Alberta. PCR primer sequences were based on sequences available in GenBank (see appendix): chitin synthase primers from *C. immitis, U. reesii,* and *A. zuffianum* sequences (Pan, Sigler, and Cole 1994); dioxygenase primers from a *C. immitis* sequence (Wyckoff et al. 1995); and orotidine decarboxylase primers from *Aspergillus niger, Aspergillus nidulans,* and *Penicillium chrysogenum* sequences (Radford 1993).

DNA was extracted from lyophilized material and PCR-amplified following protocols in Burt et al. (1994). PCR reactions were subjected to 40 cycles of 94°C/47–52°C/72°C for 1/1/1 min, and 7 min at 72°C. With primers that were imperfectly matched, we preceded this with three to six cycles of 94°C/35°C/72°C for 1/1/1 min and/or used nested primers in two successive amplifications, the second of which used a 10-fold dilution of the first PCR product as a template. Amplicons were sequenced directly using the AB automated DNA sequencer following the manufacturer's protocol. Accession numbers of sequences deposited in EMBL are AJ292100–AJ292114, AJ292088–AJ292098, and AJ291968–AJ291979.

Table 1
Origin and Identification of Strains

Strain	Identification No.	Origin
Auxarthron zuffianum		
IN	UAMH 2653	India, soil
UK	UAMH 3079	England, soil
AG	UAMH 4082	Argentina, soil
TX	UAMH 1875	Texas, prairie dog lungs (T)[a]
MX	UAMH 3135	Mexico, lizard dung
Uncinocarpus reesii		
TX	UAMH 160	Texas, *Sylvilagus auduboni,* lungs
UT	UAMH 3703	Utah, soil (MT+)
CA1	UAMH 1704	California, soil (MT+)
CA2	UAMH 1706	South California, soil (MT−)
AG	UAMH 1955	Argentina, soil
IT1	UAMH 2002	Italy, soil (MT−)
IT2	UAMH 2050	Italy, soil (MT−)
IT3	UAMH 3918	Itna, Sicily, soil
HG	UAMH 2847	Hungary, soil (MT+)
AUS	UAMH 2862	Australia, feathers (MT−)

NOTE.—MT+ and MT− refer to different mating types.
[a] "Type" species.

DNA sequences were aligned manually, and maximum-parsimony genealogies were constructed for each gene separately and for the three genes combined. Observed tree lengths were then compared with the minimum possible lengths in order to identify possible homoplasies within trees and incompatibilities between trees, following the analyses on the *C. immitis* data (Koufopanou, Burt, and Taylor 1997). Then, the three species were also analyzed together and gene genealogies were compared to examine the consistency of roots and of branching order of isolates. Finally, all data were analyzed in combination (phylogenetic analyses were performed using PAUP*, version 4.0; Swofford 1999).

Results and Discussion

Population Structure

Coccidioides immitis

Only information obtained from rooting the *C. immitis* isolates is discussed here, as the data have already been analyzed and discussed elsewhere (Koufopanou, Burt, and Taylor 1997, 1998). The two taxa previously identified as Californian (CA) and non-Californian (NCA) are both monophyletic in the chitin synthase and dioxygenase genealogies. For the orotidine decarboxylase gene, the Californian taxon has an ancestral allele and may be paraphyletic (fig. 1). Were the Californian taxon found to be paraphyletic at more loci than the non-Californian taxon, this would suggest that the Californian taxon had a larger population size than the non-Californian taxon (Hey and Kliman 1993). However, a much larger population size in the Californian taxon is

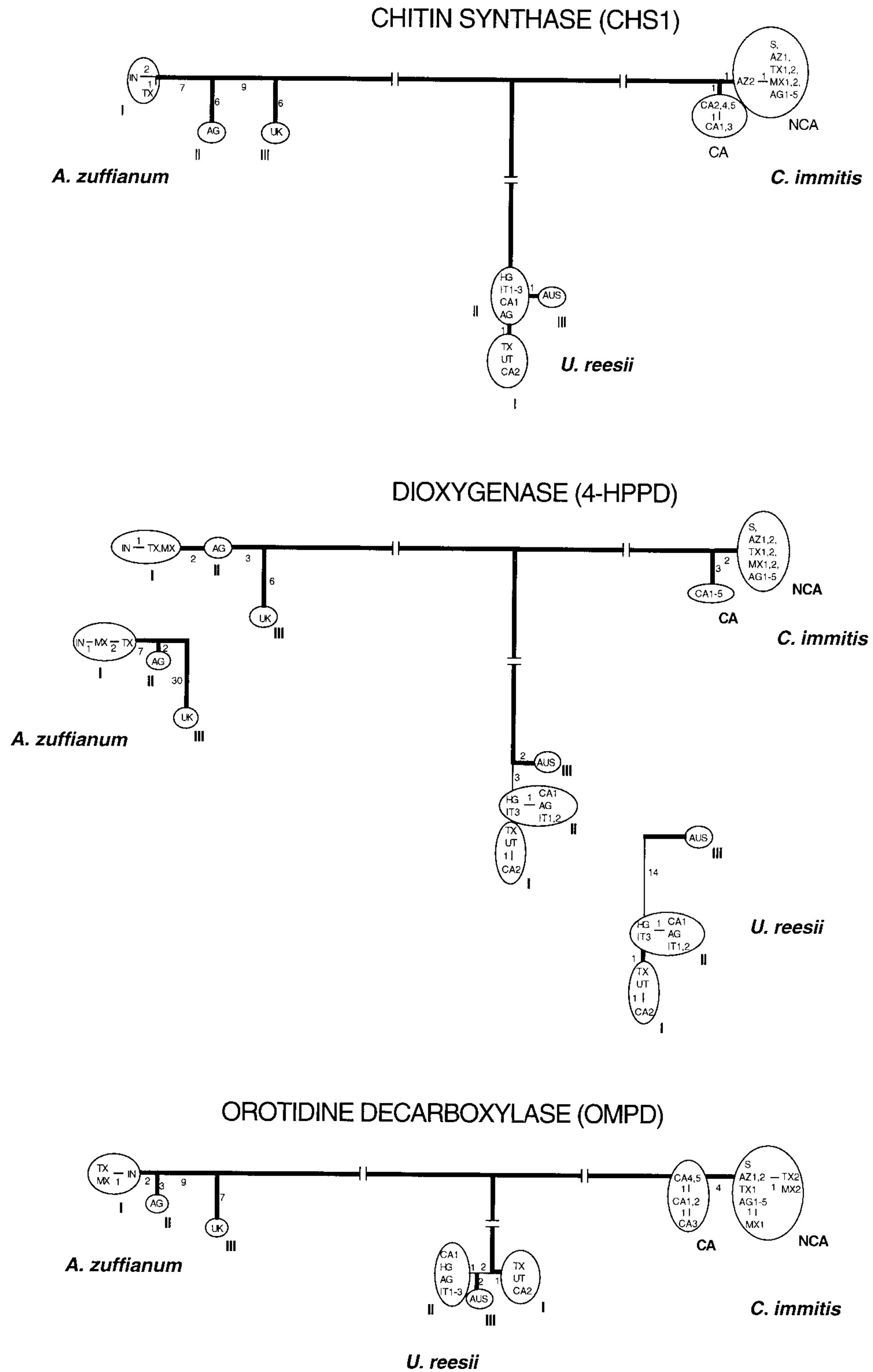

FIG. 1.—Gene genealogies for isolates of *Coccidioides immitis, Auxarthron zuffianum,* and *Uncinocarpus reesii* separately and combined. The figure shows the rooting for isolates of each species with respect to the outgroups, as well as branches shared by all three genealogies (thick lines). While there is homoplasy due to multiple hits along branches leading to each of the three species, there is no homoplasy at the tips; i.e., the sum of branch lengths shown within each species is equal to the number of variable sites in each locus. Note that *U. reesii* polymorphic sites 75 and 240, which separate isolates CA1, IT1, and IT2 from AG but introduce homoplasy within the orotidine decarboxylase locus, have been excluded from the tree (see text). For the dioxygenase locus, only the coding region can be aligned unambiguously between the three species, and separate genealogies including the intron regions are drawn for *A. zuffianum* and *U. reesii* isolates, respectively, with there being no polymorphic sites in the dioxygenase introns of *C. immitis.* All trees were constructed by maximum parsimony using PAUP*, version 4.0. Numbers indicate branch lengths within each species. Inferred reproductively isolated groups in the three species are circled. Branches between species are not drawn to scale.

unlikely, as it would have been accompanied by significantly higher values of polymorphism, not observed in our data.

Auxarthron zuffianum

From the 1,263 sites sequenced in *A. zuffianum,* 94 were polymorphic (7.4%; table 2 and fig. 2). This includes two introns in the dioxygenase locus, 48 and 59 bp long, with 27% of sites polymorphic. Note, however, that 87 out of 94 polymorphic sites are unique to two isolates (AG, UK), and if these are removed, polymorphism drops to 0.6%. No identical genotypes were found among the five isolates of *A. zuffianum.*

From the above polymorphisms, we constructed maximum-parsimony trees for each of the gene fragments (fig. 1). Chitin synthase and orotidine decarboxylase trees had minimal lengths, equal to the number of polymorphic sites in each locus (most-parsimonious tree lengths: CHS1—$L = 31$ steps; OMPD—$L = 22$ steps; table 2, *A. zuffianum* branches in fig. 1), but the dioxygenase tree is one step longer than minimum (4-HPPD, including intron sites: $L = 42$ steps). This homoplasy is removed when either of the UK, AG, or TX isolates is excluded or when site 18 is excluded. As the UK isolate is the most divergent, it is the one most likely to have had changes accumulated due to multiple hits. To test the compatibility of the three genealogies, we first excluded the UK isolate, pooled the data for the three loci, and looked for homoplasy in the combined data. The most-parsimonious tree from the combined data was one step longer than the minimum (35 vs. 34 steps) due to an incompatibility between the dioxygenase and orotidine decarboxylase loci in branches separating isolates IN, TX, and MX (group I). When the UK isolate was reincluded but site 18 was removed, the tree was still longer than minimum (94 vs. 93 steps), revealing additional incompatible sites between the two loci. Such incompatibilities are expected between isolates in a sexual species. However, there were also consistencies among the three genealogies in the branching order of isolates, with the UK isolate branching off first, followed by the AG isolate. The probability of this consistency arising if the five isolates were random samples from a panmictic sexual population is $P < 0.01$ (calculated by considering all possible rooted trees for the five isolates).

In short, three putative reproductively isolated groups in *A. zuffianum* were indicated by high genetic divergence and shared branches among three gene genealogies: group I, represented here by three isolates, and groups II and III, represented by only one isolate each (AG and UK, respectively). Group III was consistently the most divergent in the species. Mixis, on the other hand was implied by incompatible branches of isolates in group I, in accordance with the presence of a sexual state in the life cycle.

Uncinocarpus reesii

Of 1,273 nucleotides sequenced in the three genes in *U. reesii,* 27 were polymorphic, a rate of 2.1% (table 2 and fig. 2). The two introns in dioxygenase were 58 and 60 bp long, with 8.5% of sites being polymorphic. Here, too, 17 of the polymorphic sites were unique to a single isolate (AUS), and if this was removed, polymorphism dropped to 0.8%. Of the 10 isolates, only six had different genotypes. All identical genotypes were reconfirmed by independent DNA extractions, PCR, and DNA sequencing from material newly derived from the UAMH culture collection. Identical genotypes were merged for all of the analyses.

For chitin synthase and dioxygenase, trees have minimal lengths, equal to the number of polymorphic sites in each locus (most-parsimonious tree lengths: CHS1—$L = 2$ steps, 4-HPPD—$L = 17$ steps; table 2; *U. reesii* branches in fig. 1). For orotidine decarboxylase, there were two equally parsimonious trees, differing in the placement of the AG and AUS isolates, both being two steps longer than the minimum length; this homoplasy was removed when the AUS isolate was excluded, but not when any other *U. reesii* isolate was excluded, suggestive of multiple hits along branches to this isolate (for OMPD, $L = 10$ steps; excluding AUS, $L = 6$ steps $=$ minimum). To find the potentially homoplastic sites, we reincluded the AUS isolate and successively removed sites, first one and then two at a time. Minimal tree length was only obtained when both site 75 and site 240 (see table 2 and fig. 2) were excluded from the analysis. As before, the compatibility of the three genealogies was tested by excluding the AUS isolate and combining the data for the three loci. The most-parsimonious tree from the combined data was one step longer than minimum (11 vs. 10 steps), suggesting incompatible genealogies. By successively excluding each of the three loci, we located the incompatibility between the dioxygenase and orotidine decarboxylase loci in branches separating isolates HG, IT1, IT2, IT3, CA1, and AG (i.e., within group II; fig. 1). However, the homoplasy in the combined data also was removed when the AUS isolate was reincluded and the two homoplastic sites of the orotidine decarboxylase locus were excluded, implying no additional incompatibility among loci. The branching order in the dioxygenase tree and the extensive intron divergence of the AUS isolate suggest that this is the most widely divergent of the species. Although consistent with the chitin synthase genealogy, this is not consistent with the orotidine decarboxylase one, which shows the AUS isolate together with group II isolates. There is only one common branch in the three genealogies, separating the TX, UT, and CA2 (group I) isolates from the rest (fig. 1).

To summarize, in *U. reesii* the three gene genealogies share a common branch separating group I isolates from the rest, indicating complete sorting of alleles and reproductive isolation between them. Also, there is incompatibility in the placement of group II isolates, implying mixis among these isolates, again as predicted from the life cycle. Finally, the dioxygenase locus alone indicates that the AUS isolate is the most widely divergent of the species, and although this is consistent with the chitin synthase genealogy, it is inconsistent with the orotidine decarboxylase one, which shows the AUS iso-

Table 2
Distribution of Nucleotide Polymorhpisms Among Isolates of *Auxarthron zuffianum* and *Uncinocarpus reesii*

A. ZUFFIANUM

Locus	Chitin synthase (CHS1) (369bp)	Orotidine decarboxylase (OMPD) (396bp)
Site	2 3 3 4 6 7 8 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3	6 2 3 4 5 6 9 1 1 1 1 1 1 2 2 2 2 3 3 3 3 3
	1 0 3 5 6 5 8 0 1 1 5 5 6 6 9 9 0 1 1 2 2 4 5 6 7 7 8 1 2 6 6	3 7 8 4 1 9 0 2 3 4 7 8 1 1 9 9 0 3 7 7 8
	2 1 4 0 9 5 8 6 8 7 3 9 1 8 0 2 7 0 3 8 8 4 3 9	8 4 5 1 1 7 0 9 1 5 3 5 5 9 1
Nucleic acids		
Consensus	T A T T A G T G C A C C T C C C C C A A A G G G G C T A T T T	T G C G C A C A T T C C G C G A G T G G A C
TX	. . C	. . . . G . . . .
IN	. . . . C . G	. . . .
MX		. . . . G . . . .
AG	. . . G . . . . T G G T . G . C C A . . . G G G A	C . . . . G . . G . A . . . . . . C . . . .
UK	C G . . G T C A T G T T A T T G . . . . . C C . A T C . G G A	. C G A . . T G . C A G A T A G A C A A G G
Amino acids		
Consensus	A K I V K P L L Y Q D A I Y L L T V L K R P L E K N N G V T P	F R L K N I T K S I D L V F Q T D T R S I I
Substitutions	. . . . R	. T V . K V . . A . E . I . . . N . K . V .

Locus	Dioxygenase (4-HPPD) (498bp)
Site	1 1 2 2 2 3 3 3 4 4 4 4 4 5 5 5 7 7 1 1 2 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4
	3 8 2 4 7 4 5 9 0 1 2 3 6 4 5 7 6 9 2 5 8 3 3 6 6 7 7 1 4 4 4 5 5 5 6 7 8 8 8 8 8
	1 7 9 4 7 1 7 1 4 5 5 8 9 2 3 8 8 8 2 4 5 8 9
Nucleic acids	
Consensus	T C C C A G T G G C G T G G G T A C T G T T T C C A C C A G T C A A A G G G T - -
TX	. A . . . .
IN	G . . . . T . . . .
MX	G . . . .
AG	. A . T G . . A . . A . . A A . . . C . . C . . . . A . . . .
UK	. . A T G A G . A T A C A A A C G T C A C C C G T G . G G . C T C C G A A A C G T
Amino acids	
Consensus	* * * * * * * * * * * * * * * * E F I K V A G D L S L L * * * * * * * * * * * * *
Substitutions	. . . . E . G . .

Table 2
Continued

U. REESII

Locus	Chitin synthase (CHS1) (369bp)	Orotidine decarboxylase (OMPD) (396bp)	Dioxygenase (4-HPPD) (508bp)
Site	7 3	7 9 1 1 1 2 2 3	1 2 3 4 1 1 2 2 3 3 3 4 4 4 4 4 4
	2 1	5 0 2 2 2 0 4 3	8 0 9 2 5 7 6 9 0 1 8 3 4 5 5 8 8
	2	3 6 7 7 0 6	4 2 1 5 4 9 2 6 1 1 5 5 6
Nucleic acids			
Consensus	T G	G C C A T G C G	G T T T A C A C T T C C T A C C C
TX, UT	C .	. . . C C A . .	. . . A . . . .
CA2	C .	. . . C C A T .	. . . A . . . . . . T . . . . . .
HG, IT3	. .	A . T	. . . .
CA1, IT1,2	. .	A . T	. . . . . . G . . . .
AG	. .	. . T	. . . . . . G . . . .
AUS	. A	A T . . . . T A	A A G . G T . T C C . T C C T T T
Amino acids			
Consensus	N A	R I G A L V H R	* * * * G I K S Y L I * * * * * *
Substitutions	. .	. . . .	. . R . . . .

NOTE.—Site numbers indicate positions on sequences in figure 2. Dots indicate identity to the consensus, and asterisks indicate intron positions.

late clustering together with group II isolates. The possibility of uncovering conflicting gene genealogies in taxa that were separated a long time in the past but whose speciation has occurred close in time has already been discussed (Avise 1994). We conclude that there are three potentially noninterbreeding taxa in *U. reesii,* one of which is represented by a single isolate (AUS).

The inferred phylogenetic history of the three species is depicted in figure 3. Note that this is a ''species'' tree rather than a simple gene genealogy, in that it combines information on both divergence and compatibility of three genealogies, and the terminal taxa constitute putatively noninterbreeding units. The tree was rooted with Auxarthron as a monophyletic outgroup, as indicated in 18S rRNA and chitin synthase phylogenies (Pan, Sigler, and Cole 1994; Bowman, White, and Taylor 1996). Despite the extensive subdivision at the tips, each of the three named species represents a deep monophyletic clade, indicating early divergence and supporting the current generic classification. The maximum-likelihood analysis (see fig. 3 for parameters) gave no significant deviation from a ''molecular clock,'' implying that the parasitic *C. immitis* and the nonparasitic *U. reesii* evolve at similar rates, and providing no indication of accelerated evolution in *C. immitis* due to adaptation to a novel pathogenic habit. The amount of time the Californian and non-Californian taxa have been reproductively isolated was previously estimated to be 11 Myr based on five gene fragments, assuming a substitution rate at third-base positions of $u = 10^{-9}$/bp/yr (Koufopanou, Burt, and Taylor 1997, 1998). For the three fragments analyzed here, this estimate is 14 Myr, and the present tree implies that the UK isolate of Auxarthron may have been isolated from the others for as long as 28 Myr.

The finding of cryptic species in both *U. reesii* and *A. zuffianum* is consistent with the previous discovery of two species in *C. immitis* (Koufopanou, Burt, and Taylor 1997, 1998) and indicates that cryptic species are not confined to parasites but can also be found among the free-living saprophytes. Cryptic species have also been found in several other parasitic fungi (reviewed in Taylor et al. 1999; Burt, Koufopanou, and Taylor 2000), including the human pathogens *A. flavus* (2 cryptic spp.; Geiser, Pitt, and Taylor 1998), *H. capsulatum* (≥4 cryptic spp.; Kasuga, Taylor, and White 1999), *Cryptococcus*

CHITIN SYNTHASE (CHS1)

```
               1         2         3         4         5         6         7         8         9        10
C   AAGACCTGGGGCAAGGAAGCGTGGAAGAAGATTGTCGTCTGCGTCGTGAGTGACGGACGTGCAAAAATCAATCCACGGACGCGCGCCGTTCTGGCCGGTT
A   .....T....C.........T........A.........G..T..T.....C.....C.....T.....T.....G.....T..A..GT.G..A......C
U   ..C........A.....G..A...........C..G.....T.....C..C..T..C........G........CA.A..A..G..G............

C    K  T  W  G  K  E  A  W  K  K  I  V  V  C  V  V S2  D  G  R  A  K  I  N  P  R  T  R  A  V  L  A  G
A    .  .  .  A  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  L  .  .  .
U    N  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

               1         2         3         4         5         6         7         8         9        10
C   TGGGTGTATATCAAGACGGAATTGCGAAGCAACAGGTCAACGGAAAGGACGTCACTGCCCATATCTACGAATACACAACACAGATGGGAATGGAACTGAA
A   .......T..C................C.....G.................G..G.....C..T.........T..C..G.....C........G..C..
U   ....G..G........T........C.....G..A.................G.....T..C.....T.....T.....T..A.....C.....GT....

C   L  G  V  Y  Q  D  G  I  A  K  Q  Q  V  N  G  K  D  V  T  A  H  I  Y  E  Y  T  T  Q  M  G  M  E  L  K
A   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  I  .  .  .  .  .
U   .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

               1         2         3         4         5         6         7         8         9        10
C   AGGCAACCAGGTCCTTCTCAAGCCGCGCCGAGGTATGCCGGTGCAGATGCTGTTTTGCCTGAAGGAGAAGAATCAGAAGAAAATCAATTCCCACCGTTGG
A   G..A.C....A..G.AC..A..A..T..A..G........C.............C...T...........C.........G........G..T..A...
U   .........A..G..C............A..G........C..T..A............T.A.....A.....C..A...........C...........

C    G  N  Q  V  L  L  K  P  R  R  G  M  P  V  Q  M  L  F  C  L  K  E  K  N  Q  K  K  I  N S1  H  R  W
A    .  T  .  .  H  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
U    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

               1         2         3         4         5         6
C   TTTTTCCAGGCGTTCGGACGAGTCTTGGACCCGAATATATGCGTGCTCATCGACGCTGGTACCAAACCG
A   ..C........C..T.....G..TC.A..T........T..T..C..TC....T..A.....T.....T
U   ..C..T..A...........T...C....T........T.....C..TC.T.....A...........A

C    F  F  Q  A  F  G  R  V  L  D  P  N  I  C  V  L  I  D  A  G  T  K  P
A    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  L  .  .  .  .  .  .
U    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  L  .  .  .  .  .  .
```

DIOXYGENASE (4-HPPD)

```
               1         2         3         4         5         6         7         8         9        10
C   GACATCTGCACGTAAGTAGACTGATCCGTCCCTTCTTGCGGTGAATTAACCCAGAGCTGACACCGAAAATAGTGAATTCTCTGCACTGAAGAGTATCGTC
A   ..T..............T.TCTCCAAC..T.CGTAAAG.CGTCTGCTGA.A..G.TA.T-------------.........G..C..C............
U   ..T....A..........GAT........AGT.A.GCA.T...C.CAACGATGGCTAACCGTTGAA.T.GC---........C..T...........T...

C    D  I  C  T                    INTRON I                            E  F S1  A  L  K S2  I  V
A    .  .  .  .                                                        .  .  .  .  .  .  .  .  .
U    .  .  Y  .                                                        .  .  .  .  .  .  .  .  .

               1         2         3         4         5         6         7         8         9        10
C   ATGGCATCTCCAAATGATATCGTCAAGATGCCCATCAACGAGCCCGCCAAGGGAAAGAAACAATCCCAGATTGAAGAATATGTTGACTTCTACAATGGTG
A   .....C..G..GGGC..C..T.........................A.........G..G..G.................C...........C....
U   .......G.........G.......A.....A.............G..........G...........C..G........C...........GC....

                                      T-CELL REACTIVE SITE
C    M  A S1  P  N  D  I  V  K  M  P  I  N  E  P  A  K  G  K  K  Q S1  Q  I  E  E  Y  V  D  F  Y  N  G
A    .  .  .  .  G  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
U    .  .  .  .  .  .  V  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  . S2  .

               1         2         3         4         5         6         7         8         9        10
C   CTGGCGTTCAGCACATTGCTCTCCGAACCAACAACATCATCGATGCCATCACCAACCTCAAGGCGCGCGGCACCGAATTCATCAAGGTTCCAGAGACCTA
A   .C..T..............C..G..C..AG....T.................A..........C..T..TCTG.................G..A.....
U   ....T..G..A.........G........TG...........A.C..G...................T.....TGTG..G..............CAGC.....

C    A  G  V  Q  H  I  A  L  R  T  N  N  I  I  D  A  I  T  N  L  K  A  R  G  T  E  F  I  K  V  P  E  T  Y
A    .  .  .  .  .  .  .  .  .  .  D  .  .  .  .  .  .  .  .  .  .  .  .  .  L  .  .  .  .  .  .  .  .  .
U    .  .  .  .  .  .  .  .  .  .  D  .  .  .  N  .  .  .  .  .  .  .  .  .  V  .  .  .  .  .  . S2  .  .

               1         2         3         4         5         6         7         8         9        10
C   CTATGAAGACATGAAGATTCGCCTCAAGAGACAAGGCCTGGTCCTCGATGAGGACTTTGAGACCCTGAAGAGCCTGGACATCCTTATCGACTTTGACGAG
A   ...CC.G..T...CGCG.C...T.G....AGGCT..T...AC...GA.C.....T..C..C..T..CC.A.....A..T.....C...............
U   .......T.....C...C...........A...G..T.A......A.C..A..TC....A...T..C........A..T.....C........C..T...

C     Y  E  D  M  K  I  R  L  K  R  Q  G  L  V  L  D  E  D  F  E  T  L  K S2  L  D  I  L  I  D  F  D  E
A     .  Q  .  .  R  V  .  .  .  K  A  .  .  T  .  N  .  .  .  D  .  .  Q  .  .  .  .  .  .  .  .  .  .
U     .  D  .  I  .  L  .  .  .  K  .  .  .  .  .  N  .  .  L  .  .  .  Q  .  .  .  .  .  .  .  .  .  .
```

FIG. 2.—Consensus sequences of *Coccidioides immitis, Auxarthron zuffianum,* and *Uncinocarpus reesii.* DNA sequences are on top, amino acid sequences are at the bottom, introns are indicated in italics, and variable sites within species are underlined and in bold. The human T-cell reactive site of dioxygenase is also shown. Note that introns cannot be aligned reliably in the three species and all gaps were put to the right of intron sequences.

```
                1          2          3          4          5          6          7          8          9          10
C    AATGGGTATCTCCTGCAGCTTTTCACCAAGGTACGGAGATCTAAATTTACCCTAAAGGGGCAATTCCACATATTTCACTTCCAAACTACCTACTGACTC
A    .....C........C..A..C.....A......T.TCA..TCTT.C.GTAT.ATG.TAAAG..AAG..GT.GC.GACGA.GGCGTG..G----------
U    ..C..C......T.............A......T.CGCCCTGGTCG.GTGA..TCGATTCTG..ATA..TCC.GCT.ACGGT.TC.GCAG---------

C     N  G  Y  L  L  Q  L  F  T  K                                      INTRON II
A     .  .  .  .  .  .  .  .  .  .
U     .  .  .  .  .  .  .  .  .  .

                1          2          3
C    TGCATATAGCATCTCATGGATCGCCCAACC
A    ---------.............T..G..A
U    ---------..C........C..G......

C              H  L  M  D  R  P  T
A              .  .  .  .  .  .  .
U              .  .  .  .  .  .  .
```

OROTIDINE DECARBOXYLASE (OMPD)

```
                1          2          3          4          5          6          7          8          9          10
C    GACTTCTCCGCCGAAACCATAACCGGACTCCAATCCCTTTCCCAGAAGCACAACTTCCTGATCTTTGAGGACCGGAAATTCGTTGATATTGGCAACACAG
A    ..T..TAG..AAA.G..A.C..GA..CT.GA..GAA..GG..G.A..........T....T..C........C..G.....C........A..T..C.
U    .........T.........CG....C..........G.....A..A...........C.....C..A.....A..G.....C.....C..A.....G.

C     D  F S1  A  E  T  I  T  G  L  Q S1  L S1  Q  K  H  N  F  L  I  F  E  D  R  K  F  V  D  I  G  N  T
A     .  . S2  E  K  .  T  R  .  .  K  E  .  A  E  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
U     .  .  . S1  .  .  .  A  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

                1          2          3          4          5          6          7          8          9          10
C    TGCAAAAACAATACCATGGGGGCGCACTCCACATTTCGGAATGGGCTCATATCGTTAATGCAACCGTGCTTCCGGGCCCAGGTATCATAGATGCCCTAGC
A    .T..G............C...T.T...GG.....A..C.....A..CC.G..A...TGC..TA.A..C..T..AGAG..A...G.CC.G..GT.TT.
U    .......G..G..T.....C..T...T.G..G..C..AC.......G........C.....C..A.......C........G........C..G..G..

C    V  Q  K  Q  Y  H  G  G  A  L  H  I S1  E  W  A  H  I  V  N  A  T  V  L  P  G  P  G  I  I  D  A  L  A
A    .  .  .  .  .  .  .  . S1  .  R  .  .  D  .  .  .  L  .  .  C  .  I  .  .  .  E  .  .  V  Q  .  F S1
U    .  .  .  .  .  .  .  .  .  .  Q  .  .  Q  .  .  .  .  .  .  .  .  M  .  .  .  .  .  .  .  .  .  .  .

                1          2          3          4          5          6          7          8          9          10
C    GCAAGTTGCTTCTGCACCAGACTTCCCCCACGCTTCAGACAGAGGGCTCCTGATCTTGGCGACGATGACATCAAAGGGTAGTCTTGCTACTGGTCAGTAT
A    ...GACATTCAA...C.AG..T.....GT.T...GGT........A..GT.....C.C..TGAA........G..A..ATC.T.A..C..A..CG.T..C
U    A..G..G.....A..G..G.....T.....T..GG.GC...........T.....C.C..AGAA.....G..G.....C..C..G..C..C..CA.A..C

C      Q  V  A S1  A  P  D  F  P  H  A S1  D  R  G  L  L  I  L  A  T  M  T S1  K  G S2  L  A  T  G  Q  Y
A      .  T  F  N  .  Q  .  .  .  .  Y  .  G  .  .  .  .  .  .  .  .  E  .  .  .  .  . S1  .  .  .  .  D  .
U      .  .  .  .     .  .  .  .  .  .  .  A  H  .  .  .  .  .  .  .  E  .  .  .  .  .  .  .  .  .  .  K  .

                1          2          3          4          5          6          7          8          9
C    ACTGAGTTGTCGGTTGAGCTGGCGCGGAAGTACAAGGGCTTTGTCCTAGGCTTCGTGGCCAGCAGGTCCTTAGAAGGAGTTGAAACGGCTGGAAAG
A    ....CACGC.....A..CTG....A.A..AC...G...AACC..TG.T..G.....ATGT.C..A.G.G..GTCG.ACA.CTCCG.T.AG.T.CC.
U    ..G...C.T.....G...A..........A..T.G.........T..G..G..T.....T.CG..A..TC.G.CG.AT..G..C.......C....

C     T  E  L S1  V  E  L  A  R  K  Y  K  G  F  V  L  G  F  V  A S2  R S1  L  E  G  V  E  T  A  G  K
A     .  A  R  .  .  D  W  .  .  .  H  R  .  T  .  V  .  .  .  C  T  K  A  . S1  D  I S1  A  E  V  P
U     .  .  .  .  .  .  M  .  .  .  .  R  .  .  .  .  .  .  .  .  T  .  .  .  A  D  .  D  .  .  A  .
```

FIG. 2 *(Continued)*

neoformans (≥2 spp.; Franzot, Salkin, and Casadevall 1999), and *Candida albicans* (Sullivan et al. 1995), and in the wheat-scab-causing *F. graminearum* (7 spp.; O'Donnell et al. 2000). Together, these studies indicate that species designation based solely on phenotypic characteristics tends to underestimate the number of fungal species by a factor of at least two.

The coalescence time for *U. reesii* and *C. immitis* is more than 20 times as long as that of taxa within species complexes (fig. 3). Unless some exceptional circumstances have spurred the recent speciation, the long internal branches indicate that taxa at the tips are rather transient over evolutionary time, soon to go extinct and be replaced by others along the same main lineages. In other words, taxa are continuously being generated and going extinct along the named branches, which may represent some basic environmental niche and/or organismal structure or habit, and it is unlikely for more than one taxon from each clade to survive and diverge over long evolutionary periods, a phenomenon recognized by Williams (1992, p. 132) as "normalizing clade selection." Alternatively, the long internal branches may have resulted from poor sampling of the existing fungal diversity, with the implication that more taxa remain to be found. Although *U. reesii* is still the closest known relative of *C. immitis,* recently published 18S sequences suggest that *Onygena equina* and *Ascocalvatia alveolata* may be closer relatives of *C. immitis* than *A. zuffianum* (Landvik, Shailer, and Eriksson 1996; Sugiyama, Ohara, and Mikawa 1999). Nevertheless, our data suggest that the discovery of new cryptic subdivisions is likely to be within the existing framework of generic classification. No similar outgroup analysis is provided in the other studies that have discovered cryptic species, so it is dif-

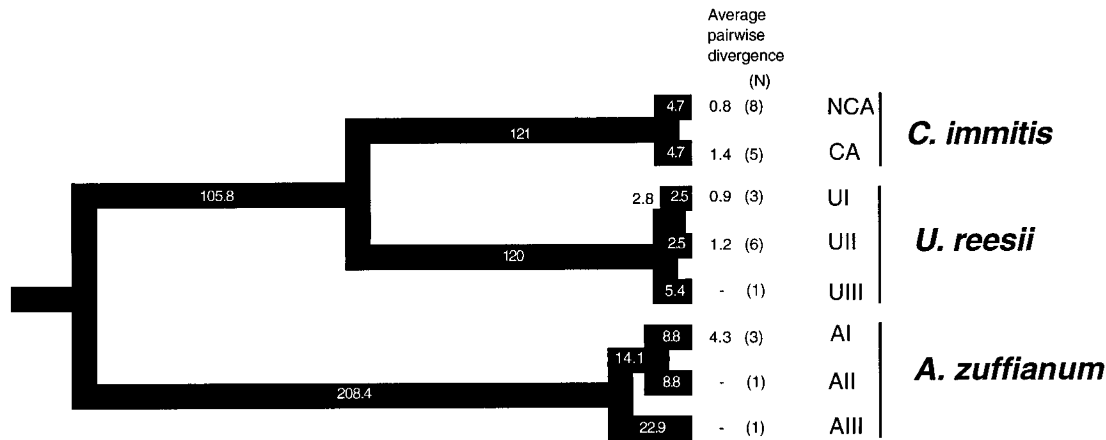

FIG. 3.—A "species" tree, containing only branches compatible among the three gene genealogies and showing potentially noninterbreeding taxa within *Coccidioides immitis, Auxarthron zuffianum,* and *Uncinocarpus reesii.* The tree is a maximum-parsimony tree on the data from the three genes combined, with polymorphisms within taxa coded as ambiguities for the respective nucleotides (branch-and-bound search using PAUP*, version 4.0). Maximum-likelihood branch lengths were calculated on this tree using the "general time-reversible" option for substitution rates and an "estimated proportion of invariable sites" of 0.4515, as these were the only added parameters that significantly increased the likelihood for the observed data; a "molecular clock," was enforced as it did not decrease this likelihood substantially. Uncorrected average pairwise divergences within taxa are also shown for comparison. Both branch lengths and divergences within taxa are per 1,000 sites; *N* is the number of isolates.

ficult to assess the validity of this hypothesis for fungi in general.

In contrast to *C. immitis,* both *A. zuffianum* and *U. reesii* are known to have sexual states, and mating tests could be performed to test the mating compatibilities of strains. No such data exist for *A. zuffianum,* but limited data on *U. reesii* (numbers of gymnothecia obtained from a cross; table IV in Sigler and Carmichael 1976) are consistent with our findings: the most divergent AUS isolate will not mate with either UT, CA1, or HG of groups I and II (but will mate with other strains not included in this study). Isolates IT1 and IT2 from group II will mate readily with CA1 and HG strains from the same group. The IT1 isolate (but not IT2), however, will also mate with the UT strain of group I. As no information is given on the fertility of ascospores in the gymnothecia, it is difficult to evaluate from these data the extent of the correlation between mating incompatibility and genetic divergence of strains. A good correlation has been found between mating incompatibility and sequence divergence at a single locus (nuclear ITS) for isolates from various geographic locations in the colonial green algae *Pandorina morum* and *Gonium pectorale* (Coleman, Suarez, and Goff 1994).

Molecular Evolution

The above analysis identified eight phylogenetically distinct taxa in the three species, three of which are represented by only one isolate. For clarity of presentation, we shall refer to the original species as "species complexes." To examine the evolution of the three loci, we calculated the numbers of silent and expressed nucleotide changes within and between the eight taxa (table 3). Changes among isolates within taxa were counted as polymorphisms (P), while changes among taxa, i.e., fixed within taxa, were counted as divergences (D). For the three taxa with only one representative, i.e., taxa II and III of *A. zuffianum* and taxon III of *U. reesii,* we assumed that all of the observed divergence was fixed within these taxa; in other words, sites that are contributing to the divergence of these isolates will not be found to be polymorphic once more isolates from these taxa have been analyzed. For expressed changes, the minimum number of nucleotide changes needed to convert one amino acid to another was calculated according to the Protpars matrix in PAUP* version 4.0 (Swofford 1999).

There was no significant difference in silent or expressed polymorphism across the three loci or across the five taxa with more than one representative, nor was there any difference in the ratio of expressed to silent polymorphism (table 3). Both silent and expressed polymorphisms are low, $<0.5\%$. Considering the divergence between taxa within species complexes, *A. zuffianum* is much more divergent than either *C. immitis* or *U. reesii.* Within *A. zuffianum,* there are also significant differences in the rates of evolution of the three loci: chitin synthase has more silent substitutions than either of the other two loci, but orotidine decarboxylase has 5 or 10 times as many expressed changes as the other two ($G = 14.59$, $df = 2$, $P < 0.001$). The ratios of $E/(E + S)$ were 0.04, 0.18, and 0.48 for CHS1, 4-HPPD, and OMPD, respectively, well below the 0.75 ratio expected under the neutral model if all codon changes have equal probabilities. This shows that the three loci are evolving under purifying selection, with most amino acid changes being deleterious and quickly eliminated from the population. Between the three complexes, silent and expressed divergences (i.e., divergent sites that are invariant within complexes) were 26% and 8.6%, respectively, much higher than the divergences within complexes (which were less than 4% and 1%, respectively). Such high divergences are likely to be saturated, and rates of evolution cannot be computed accurately from the be-

Table 3
Nucleotide Polymorphisms and Divergences (numbers of silent [S] and expressed [E] changes per locus) Within and Between Taxa of *Coccidioides immitis, Auxarthron zuffianum,* and *Uncinocarpus reesii* (groups of isolates as in fig. 2)

		CODING										INTRON			
												4-HPPD			
		CHS1 (369 bp)		4-HPPD (390 bp)		OMPD (396 bp)		Total (1,155 bp)				I (48–61 bp)	II (59–78 bp)		
	N	S	E	S	E	S	E	S	E	S + E	%[a]			Total	%[a]
Polymorphism															
C. immitis															
CA	5	1	0	0	0	0	2	1	2	3	0.26	0	0	0	0
NCA	8	1	0	0	0	2	0	3	0	3	0.26	0	0	0	0
A. zuffianum															
AI	3	3	0	1	0	0	1	4	1	5	0.43	2	0	2	1.87
U. reesii															
UI	3	0	0	1	0	1	0	2	0	2	0.17	0	0	0	0
UII	6	0	0	0	1	1	0	1	1	2	0.17	0	0	0	0
Total	25	5	0	2	1	4	3	11	4	15		2	0	2	
Divergence within species complexes															
C. immitis															
CA vs. NCA	2	2	0	5	0	4	0	11	0	11	0.95	0	0	0	0.00
A. zuffianum															
AI vs. AII vs. AIII	3	27	1	9	2	11	10	47	13	60	5.19	14	13	27	25.23
U. reesii															
UI vs. UII vs. UIII	3	2	0	5	0	6	0	13	0	13	1.13	4	6	10	8.47
Total	8	31	1	19	2	21	10	71	13	84		18	19	37	
D/P		6.20	∞	9.50	2.00	5.25	3.33	6.45	3.25			9.00	∞	18.50	
$Q = (D/P)_E/(D/P)_S$		∞		0.21		0.63		0.50							
P		1.00 NS		0.34 NS		0.67 NS		0.28 NS							
Divergence between species complexes															
C. immitis vs. *A. zuffianum* vs. *U. reesii*	3	91	7	84	26	127	66	302	99	401	34.70				

NOTE.—CA = Californian; NCA = non-Californian; D = divergence; P = polymorphism. Comparisons between taxa include only sites fixed within the taxa. *N* = number of isolates. *P* values are from Fisher's exact test.

[a] (S + E)/total length of sequence (×100).

tween-complexes divergence values (Maynard Smith 1994), especially since silent and expressed changes are likely to show different degrees of saturation. Between the species complexes, again, there is significant heterogeneity among loci in rates of silent and expressed substitution, with the E/(E + S) ratios being 0.07, 0.24, and 0.34 for CHS1, 4-HPPD, and OMPD, respectively, <0.75 and similar to those within the *A. zuffianum* complex.

To test whether the three loci were fixing amino acids in different taxa faster or slower than would be predicted by the amounts of standing polymorphism within the taxa, we calculated for each locus the ratio D/P of divergence over polymorphism for silent and expressed substitutions, respectively, from the pooled polymorphism and within-complex divergence (table 3). The ratio $Q = (D/P)_E/(D/P)_S$ is equal to one if loci are evolving according to the neutral hypothesis (McDonald and Kreitman 1991; Charlesworth 1994). In all three loci, Q values are <1, indicating negative, purifying selection rather than positive selection, but the deviations from unity are not significant (table 3). As there is no significant heterogeneity among loci in the divergence-to-polymorphism ratios ($G = 0.15$ and $G = 0.44$ for expressed and silent changes, respectively, NS), we were able to combine the data from the three loci, but the deviation was still nonsignificant. The Q ratios remained <1 and nonsignificant if the uncorrected between-complexes divergence values were included in the estimations. As silent sites are likely to be more saturated than replacement sites, correcting the between-complexes divergence values for saturation would tend to decrease the Q values and hence increase the deviation from unity.

Contrary to expectation, we found no positive selection or increased amino acid polymorphism in the 200-bp-long stretch of the T-cell reactive site of dioxygenase in *C. immitis* (fig. 2), implying no strong selection for diversification of the pathogen in this region. Indeed, an alignment of six dioxygenase proteins from a variety of organisms, including the human, mouse, rat, and the protist Tetrahymena, indicates that this stretch is among the most well conserved in the entire molecule (Wyckoff et al. 1995). Somewhat higher rates of expressed polymorphism were found in serine proteinase and chitinase gene fragments, which are also thought to be involved in the antigenic reaction of the host against this pathogen (Koufopanou, Burt, and Taylor 1997,

1998). These fragments were not included in this study, as we were not able to amplify them in *A. zuffianum* or *U. reesii* using primers designed from the *C. immitis* sequence. Also surprisingly, there was very low polymorphism in the dioxygenase introns: there were only two variable sites within group I of *A. zuffianum*, and in *U. reesii,* there was no intron polymorphism within either of the two taxa (table 3), although there was considerable intron divergence among taxa in these two species complexes. In *C. immitis,* only a single variable site was found previously among 317 intron sites from the dioxygenase, serine proteinase and chitinase loci, a rate of 0.3%, compared with 0.7% for third codon sites polymorphic within the two taxa (Koufopanou, Burt, and Taylor 1997, 1998). Although low in polymorphism, introns of different species complexes are so divergent, with a number of insertions/deletions, that they cannot be aligned unambiguously. Such disparity in the polymorphism and divergence of apparently neutral sites might suggest very quick fixation of mutations arising in these populations.

In short, we did not detect any significant deviation of amino acid divergence from the neutral expectation, as measured by the Q ratio of expressed and silent D/P values, providing no evidence for positive Darwinian selection by rapid fixation of favorable amino acids. Instead, the ratio was less than unity, suggestive of stasis by elimination of mostly deleterious alleles. Slow evolution and amino acid conservation were also indicated by the low values of expressed divergence alone, relative to total divergence. Although powerful in providing an intuitive null hypothesis for the detection of deviations, these tests are weak in that if positive and negative selection are acting on different sites or regions of the same gene, they would tend to cancel each other out, with the outcome possibly appearing as neutral evolution. Recent reviews have indicated that different regions, or even sites, are very likely to be under different selective pressures (Ngai et al. 1993; Golding and Dean 1998), thus making it very difficult to detect any particular type of selection when summing across sites. Furthermore, the Q ratio is very sensitive to small changes in the numbers of silent and expressed polymorphisms within populations, which tend to be small, especially when only potentially interbreeding populations are considered. Here, a stretch of 1,155 bp sequenced in a total of 25 isolates from 5 taxa and 3 loci was able to generate only a small number of polymorphic sites and an even smaller number of amino acid polymorphisms. A much larger sample size would thus be required to remove this sensitivity. No extreme heterogeneity was found among regions of the same fragment in our data, including the T-cell reactive site of dioxygenase.

Nevertheless, we did find significant heterogeneity in the rates of evolution of the three loci considered here. From comparisons across species, it is now obvious that proteins evolve at different rates, indicating differences in selective pressures among different molecules. Many proteins from distant organisms are greatly conserved, implying strong negative selection. On the other hand, surface antigens of parasites or viruses tend to be under strong positive selection more frequently than other genes (Endo, Ikeo, and Gojobori 1996). Fewer studies have scaled the divergence among species to the polymorphism within species (reviewed by Brookfield and Sharp 1994), and again the results have been mixed, with examples of positive (McDonald and Kreitman 1991; Eanes, Kirchner, and Yoon 1993; Kliman and Hey 1993), neutral (Kliman and Hey 1993), and negative evolution (Nachman, Boyer, and Aquadro 1994). The heterogeneity of findings probably reflects a true heterogeneity in the mode of evolution of different genes and populations.

Conclusions

Our study illustrates the utility of outgroups in population genealogical analyses, as they allow rooting of relationships within taxa. When close relatives are included in the analyses, the extent of diversification within taxa can also be evaluated. By combining data from the three species, we established the monophyly of *C. immitis* as a whole, and hence a single origin of pathogenicity in this fungus. As with *C. immitis,* cryptic species were also revealed in both *A. zuffianum* and *U. reesii,* the most divergent of which may have been reproductively isolated for as long as 28 Myr. Our studies, together with a few others, indicate that species designation based solely on phenotypic characteristics tends to underestimate the number of fungal species by a factor of at least two. Our analysis also indicates, however, that cryptic species are likely to be well within the existing framework of generic classification, as they all represent subdivisions at the tips of deep monophyletic clades. It is unclear at present whether the same will be true for fungi other than those considered here.

Adaptation to the pathogenic lifestyle in *C. immitis* does not seem to have been accompanied by accelerated evolution or positive Darwinian selection in the three protein fragments analyzed, one of which included a region known to produce antigenic responses in humans. Instead, our data suggest relative stasis by the action of negative, purifying selection against mostly deleterious amino acid changes.

Acknowledgment

This work was supported by an NIH-NIAID grant to J.W.T.

LITERATURE CITED

AVISE, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.

AVISE, J. C., and R. M. J. BALL. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. Oxf. Surv. Evol. Biol. **7**:45–67.

BOWMAN, B. H., and J. W. TAYLOR. 1993. Molecular phylogeny of pathogenic and non-pathogenic Onygenales. Pp. 169–178 *in* D. R. REYNOLDS and J. W. TAYLOR, eds. The fungal holomorph: mitotic, meiotic and pleomorphic speciation in fungal systematics. CAB International, Wallingford, England.

BOWMAN, B. H., T. J. WHITE, and J. W. TAYLOR. 1996. Evolutionary relationships of human pathogenic fungi: multiple origins of pathogenicity in the fungal order Onygenales. Mol. Phylogenet. Evol. **6**:89–96.

BROOKFIELD, J. F. Y., and P. SHARP. 1994. Neutralism and selectionism face up to DNA data. Trends Genet. **10**:109–111.

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.

BURT, A., D. A. CARTER, T. J. WHITE, and J. W. TAYLOR. 1994. DNA sequencing with arbitrary primer pairs. Mol. Ecol. **3**: 523–525.

BURT, A., B. M. DECHAIRO, G. L. KOENIG, D. A. CARTER, T. J. WHITE, and J. W. TAYLOR. 1997. Molecular markers reveal differentiation among isolates of *Coccidioides immitis* from California, Arizona and Texas. Mol. Ecol. **6**:781–786.

BURT, A., V. KOUFOPANOU, and J. W. TAYLOR. 2000. Population genetics of human-pathogenic fungi. Pp. 229–244 *in* R. C. A. THOMPSON, ed. Molecular epidemiology of infectious diseases. Arnold, London.

CARTER, D. A., A. BURT, J. W. TAYLOR, G. L. KOENIG, and T. J. WHITE. 1996. Clinical isolates of *Histoplasma capsulatum* from Indianapolis have a recombining population structure. J. Clin. Microbiol. **34**:2577–2584.

CHARLESWORTH, B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. **63**:213–227.

COLEMAN, A. W., A. SUAREZ, and L. J. GOFF. 1994. Molecular delineation of species and syngens in volvocacean green algae (Chlorophyta). J. Phycol. **30**:80–90.

EANES, W. F., M. KIRCHNER, and J. YOON. 1993. Evidence for adaptive evolution of the *G6pd* gene in the *Drosophila melanogaster* and *Drosophila simulans* lineages. Proc. Natl. Acad. Sci. USA **90**:7475–7479.

ENDO, T., K. IKEO, and T. GOJOBORI. 1996. Large-scale search for genes on which positive selection may operate. Mol. Biol. Evol. **13**:685–690.

FISHER, M. C., G. L. KOENIG, T. J. WHITE, and J. W. TAYLOR. 2000*a.* A test for concordance between the multilocus genealogies of genes and microsatellites in the pathogenic fungus *Coccidioides immitis.* Mol. Evol. Biol. **17**:1164–1174.

———. 2000*b.* Pathogenic clones versus environmentally driven population increase: analysis of an epidemic of the human fungal pathogen *Coccidioides immitis.* J. Clin. Microbiol. **38**:807–813.

FRANZOT, S. P., I. F. SALKIN, and A. CASADEVALL. 1999. *Cryptococcus neoformans* var. *grubii*: separate varietal status for *Cryptococcus neoformans* serotype A isolates. J. Clin. Microbiol. **37**:838–840.

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.

GOLDING, G. B., and A. M. DEAN. 1998. The structural basis of molecular adaptation. Mol. Biol. Evol. **15**:355–369.

HEY, J., and R. M. KLIMAN. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol. Biol. Evol. **10**:804–822.

HUDSON, R. R. 1996. Molecular population genetics of adaptation. Pp. 291–309 *in* M. R. ROSE and G. V. LAUDER, eds. Adaptation. Academic Press, New York.

KASUGA, T., J. W. TAYLOR, and T. J. WHITE. 1999. Phylogenetic relationships of varieties and geographical groups of the human pathogenic fungus *Histoplasma capsulatum* Darling. J. Clin. Microbiol. **37**:653–663.

KLIMAN, R. M., and J. HEY. 1993. DNA sequence variation at the period locus within and among species of the *Drosophila melanogaster* complex. Genetics **133**:375–387.

KOUFOPANOU, V., A. BURT, and J. TAYLOR. 1997. Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis.* Proc. Natl. Acad. Sci. USA **94**:5478–5482.

———. 1998. Correction: concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis.* Proc. Natl. Acad. Sci. USA **95**:8414.

LANDVIK, S., N. F. J. SHAILER, and O. E. ERIKSSON. 1996. SSU rDNA sequences support for a close relationship between the Elaphomycetales and the Eurotiales and Onygenales. Mycoscience **37**:237–241.

MCDONALD, J. H., and M. KREITMAN. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila.* Nature **351**:652–654.

MAYNARD SMITH, J. 1994. Estimating selection by comparing synonymous and substitutional changes. J. Mol. Evol. **39**: 123–128.

NACHMAN, M. W., S. N. BOYER, and C. F. AQUADRO. 1994. Nonneutral evolution at the mitochondrial NADH dehydrogenase subunit 3 gene in mice. Proc. Natl. Acad. Sci. USA **91**:6364–6368.

NEIGEL, J. E., and J. C. AVISE. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515–534 *in* E. NEVO and S. KARLIN, eds. Evolutionary processes and theory. Academic Press, New York.

NGAI, J., M. M. DOWLING, L. BUCK, R. AXEL, and A. CHESS. 1993. The family of genes encoding odorant receptors in the channel catfish. Cell **72**:657–666.

O'DONNELL, K. O., H. C. KISTLER, B. K. TACKE, and H. H. CASPER. 2000. Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum,* the fungus causing wheat scab. Proc. Natl. Acad. Sci. USA **97**:7905–7910.

PAN, S., L. SIGLER, and G. T. COLE. 1994. Evidence for a phylogenetic connection between *Coccidioides immitis* and *Uncinocarpus reesii* (Onygenaceae). Microbiology **140**: 1481–1494.

PAPPAGIANIS, D. 1994. Marked increase in cases of Coccidioidomycosis in California: 1991, 1992, and 1993. Clin. Infect. Dis. **19**(Suppl. 1):S14–S18.

PAPPAGIANIS, D., R. K. SUN, S. B. WERNER, G. W. I. RUTHERFORD, R. W. ELSEA, G. B. J. MILLER, N. BOOTWALA, and R. S. HOPKINS. 1993. Coccidioidomycosis—United States, 1991–1992. Morbid. Mortal. Wkly. Rep. **42**:21–24.

RADFORD, A. 1993. A fungal phylogeny based upon orotidine 5′-monophosphate decarboxylase. J. Mol. Evol. **36**:389–395.

RIPPON, J. W. 1988. Medical mycology. 3rd edition. W. B. Saunders, Philadelphia.

SIGLER, L., and J. W. CARMICHAEL. 1976. Taxonomy of *Malbranchea* and some other hyphomycetes with arthroconidia. Mycotaxon **4**:349–488.

SPHAER, E. G., and J. I. MULLINS. 1993. Rates of amino acid change in the envelope protein correlate with pathogenicity of primate lentiviruses. J. Mol. Evol. **37**:57–65.

SUGIYAMA, J., A. OHARA, and T. MIKAWA. 1999. Molecular phylogeny of onygenalean fungi based on small subunit ribosomal DNA (SSU rDNA) sequences. Mycoscience **40**: 251–258.

SULLIVAN, D. J., T. J. WESTERNENG, K. A. HAYNES, D. E. BENNETT, and D. C. COLEMAN. 1995. *Candida dublinensis* sp. nov.: phenotypic and molecular characterization of a novel species associated with oral candidiosis in HIV-infected individuals. Microbiology **141**:1507–1521.

SWOFFORD, D. L. 1999. PAUP*4.0: phylogenetic analysis using parsimony. Sinauer, Sunderland, Mass.

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.

WILLIAMS, G. C. 1992. Natural selection: domains, levels and challenges. Oxford University Press, Oxford, England.

WYCKOFF, E. E., E. J. PISHKO, T. N. KIRKLAND, and G. COLE. 1995. Cloning and expression of a gene encoding a T-cell reactive protein from *Coccidioides immitis*: homology to 4-hydroxyphenylpyruvate dioxygenase and the mammalian F antigen. Gene **161**:107–111.

RICHARD THOMAS, reviewing editor

Accepted March 14, 2001

APPENDIX

List of PCR Primers

Primer	Position[a]	Primer (5′–3′)
Chitin synthase (CHS1)		
CS1.	28	GGT GTC TTC AAG AAC ATC GA
CS4A*	525	CAT AAC CTT GAT CTC TCC AC
CS11.	28	GGC CTC TTC AAG AAT ATC GA
CS4*.	488	GG TTC CAG GTC GAA GGC G
CS1A	52	ATG TGC TCT CGG ACA AAC AG
CS14A*	521	AC TTT GAT CTC CCC GCA AGC
Dioxygenase (4-HPPD)		
DO7	825	GAG AAG ATC CTC GGA TTC CA
DO10*	1463	GC CCT GAA GTT GCC CGC
Orotidine decarboxylase (OMPD)		
OR2	138	G CTC CTG GAT CTT GCT GAC
OR7*	763	GG CGT CTG GTA CTG CTG ACC
OR13	228	CTC GGT CCC TAC ATG GTC G
OR6*	707	CGT GAA GAC GAC AAA ATC CTC

NOTE.—Consecutive primers were used as pairs. Asterisks indicate 3′ primers, presented as reverse complements to the original sequences.

[a] With respect to sequence of reference in GenBank (see text for references).