Molecular Ecology (2003) **12**, 3383–3401 doi: 10.1046/j.1365-294X.2003.01995.x

Phylogeography of the fungal pathogen *Histoplasma capsulatum*

TAKAO KASUGA,[1] THOMAS J. WHITE,[2] GINA KOENIG,[3] JUAN MCEWEN,[4] ANGELA RESTREPO,[5] ELIZABETHA CASTAÑEDA,[5] CARLOS DA SILVA LACAZ,[6] ELISABETH M. HEINS-VACCARI,[6] ROSELI S. DE FREITAS,[6] ROSELY M. ZANCOPÉ-OLIVEIRA,[7] ZHENYU QIN,[8] RICARDO NEGRONI,[9] DEIDRE A. CARTER,[10] YUZURU MIKAMI,[11] MIKI TAMURA,[12] MARÍA LUCÍA TAYLOR,[13] GEORGINA F. MILLER,[14] NATTEEWAN POONWAN[15] and JOHN W. TAYLOR[1]

[1]*Department of Plant and Microbial Biology, 321 Koshland Hall, University of California, Berkeley, CA 94720, USA,* [2]*Celera Diagnostics, Alameda, CA, USA,* [3]*Roche Molecular Systems, Alameda, CA, USA,* [4]*Corporación para Investigaciones Biológicas, Medellin, Colombia,* [5]*Instituto Nacional de Salud Santafé de Bogotá, Colombia,* [6]*Instituto de Medicina Tropical de São Paulo, Brazil,* [7]*Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil,* [8]*Center of Medical Mycology, Institute of Dermatology, Chinese Academy of Medical Sciences, Nanjing, China,* [9]*Hospital de Enfermedades Infecciosas, Buenos Aires, Argentina,* [10]*Microbiology Department, University of Sydney, Australia,* [11]*Research Center for Pathogenic Fungi and Microbial Toxicoses, Chiba University, Chiba, Japan,* [12]*Institute of Molecular and Cellular Biosciences, University of Tokyo, Japan,* [13]*Departamento de Microbiología y Parasitología, Facultad de Medicina, Universidad Nacional Autónoma de México, México,* [14]*Veterinary Resources Program, National Institutes of Health, Bethesda, MD, USA,* [15]*Department of Medical Sciences, National Institute of Health, Nonthaburi, Thailand*

Abstract

Until recently, *Histoplasma capsulatum* was believed to harbour three varieties, var. *capsulatum* (chiefly a New World human pathogen), var. *duboisii* (an African human pathogen) and var. *farciminosum* (an Old World horse pathogen), which varied in clinical manifestations and geographical distribution. We analysed the phylogenetic relationships of 137 individuals representing the three varieties from six continents using DNA sequence variation in four independent protein-coding genes. At least eight clades were identified: (i) North American class 1 clade; (ii) North American class 2 clade; (iii) Latin American group A clade; (iv) Latin American group B clade; (v) Australian clade; (vi) Netherlands (Indonesian?) clade; (vii) Eurasian clade and (viii) African clade. Seven of eight clades represented genetically isolated groups that may be recognized as phylogenetic species. The sole exception was the Eurasian clade which originated from within the Latin American group A clade. The phylogenetic relationships among the clades made a star phylogeny. *Histoplasma capsulatum* var. *capsulatum* individuals were found in all eight clades. The African clade included all of the *H. capsulatum* var. *duboisii* individuals as well as individuals of the other two varieties. The 13 individuals of var. *farciminosum* were distributed among three phylogenetic species. These findings suggest that the three varieties of *Histoplasma* are phylogenetically meaningless. Instead we have to recognize the existence of genetically distinct geographical populations or phylogenetic species. Combining DNA substitution rates of protein-coding genes with the phylogeny suggests that the radiation of *Histoplasma* started between 3 and 13 million years ago in Latin America.

Keywords: allopatric speciation, glacial refugia, last glacial maxima, phylogenetic species, population diversity, star phylogeny

Received 11 April 2003; revision received 29 July 2003; accepted 28 August 2003

Correspondence: Takao Kasuga. Fax: 510 642 4995; E-mail: kasugat@uclink.berkeley.edu; John Taylor. E-mail: jtaylor@socraks.berkeley.edu

Table 1 Abbreviations of varieties and geographical groups of *Histoplasma capsulatum*

Abbr.	Variety and population
Hc var. *capsulatum*	*H. capsulatum* var. *capsulatum*
Hc var. *farciminosum*	*H. capsulatum* var. *farciminosum*
Hc var. *duboisii*	*H. capsulatum* var. *duboisii*
NAm 1	North American class 1
NAm 2	North American class 2
LAm A*	Latin American group A
LAm B*	Latin American group B

*Latin American group A and Latin American group B are synonyms of South American group A and South American group B in Kasuga *et al.* (1999), respectively.

Introduction

We investigated the population structure and phylogeny of the pathogenic ascomycete fungus *Histoplasma capsulatum* Darling (mieosporic state or teleomorph: *Ajellomyces capsulatus* (Kwon-Chung) McGinnis *et* Katz). This dimorphic fungus causes deep mycosis in various mammalian species including humans (Rippon 1988; Kwon-Chung & Bennett 1992). It exists in the mycelial phase in soil enriched with bird and bat guano. In the lung, inhaled airborne microconidia or hyphal fragments transform to the pathogenic yeast form and start the mycosis. The disease is noncontagious between humans. *Histoplasma capsulatum* occurs in temperate and tropical regions worldwide and distinct genotypes are known which show different clinical manifestations and geographical distributions. On the basis of morphology and pathogenicity, the genus *Histoplasma* has been thought to consist of three distinct varieties: *H. capsulatum* (*Hc*) var. *capsulatum*, *Hc* var. *duboisii* and *Hc* var. *farciminosum* or three independent species: *H. capsulatum*, *H. duboisii* and *H. farciminosum* (Kwon-Chung & Bennett 1992; Rippon 1988).

Recently, 46 isolates of *H. capsulatum* representing the three varieties from various geographical locations were subjected to phylogenetic analysis using DNA sequence variation in four independent protein-coding genes (Kasuga *et al*. 1999). This study showed that *H. capsulatum* consisted of at least six clades: (i) North American class 1 *Hc* var. *capsulatum* (NAm 1; see Table 1 for abbreviations); (ii) North American class 2 *Hc* var. *capsulatum* (NAm 2); (iii) Panamanian *Hc* var. *capsulatum*; (iv) South American group A *Hc* var. *capsulatum*; (v) South American group B *Hc* var. *capsulatum* and (vi) *Hc* var. *duboisii*. *Histoplasma capsulatum* var. *farciminosum* was found within the South American group A clade. Under a genealogical concordance–phylogenetic species concept (GC–PSC) (Mayden 1997), based on possession of multiple shared derived characters as well as concordance of four gene genealogies, *H. capsulatum* was claimed to harbour six species instead of three varieties or three species. However, this study did not include individuals from many regions of the globe.

To challenge the hypothesis that *H. capsulatum* comprises six phylogenetic species, we analysed phylogenetic relationships of 92 additional *H. capsulatum* isolates which, together with 45 of the 46 analysed before, represent 25 countries in six continents. We challenged the validity of the six-species hypothesis by constructing a more detailed phylogeographical map and searching for possible hybrids between geographical populations. In this research, we applied the GC–PSC to define genetically isolated populations in *H. capsulatum*. From neutral mutation rates in protein-coding genes (Kasuga *et al*. 2002) and genetic distances between geographical populations, we estimated the ages of populations and used this information to discuss the history of each population and the origin of the *H. capsulatum* complex.

Under a neutral model of evolution, genetic drift will inevitably lead to fixation of formerly polymorphic loci following genetic isolation and, after sufficient time, to genealogical concordance of multiple gene trees. New polymorphisms will continue to arise and accumulate in these loci in each interbreeding population. Thus, these genetically isolated populations or species will be recognized as reciprocally monophyletic groups. Recombination among individuals within a species will result in discordance among the gene genealogies. Therefore, in GC–PSC, the transition from deep genealogical concordance to shallow genealogical discordance is used to delimit species boundaries (Avise & Ball. 1990; Baum & Shaw 1995). The GC–PSC is especially compatible with DNA analyses and has been demonstrated to recognize genetically distinct populations or species without actually observing matings or gene flow in nature (reviewed in Taylor *et al*. 2000).

We found that *H. capsulatum* comprises seven phylogenetic species plus a Eurasian clade that emerges from the largest clade, South American group A.

In this research, numbers of Mexican and Central American isolates were found in the South American group A clade; therefore, we replace the clade name 'South America' with 'Latin America'. Individuals identified as *Hc* var. *duboisii* were limited to Africa but the African clade included individuals morphologically identified as *Hc* var. *capsulatum* and *Hc* var. *farciminosum*. Individuals identified as *Hc* var. *farciminosum* were accommodated in three different phylogenetic species, supporting the claim that *Hc* var. *farciminosum* is a collection of individuals from different clades that share the ability to cause disease in horses and not a phylogenetic species. There was no resolution of the branching order of the clades, supporting the conclusion that *H. capsulatum* radiated rapidly over a short period, which we estimate occurred 3–13 million years ago (Ma).

Materials and methods

Organisms

Table 2 lists the isolates used in the study. The 45 isolates labelled H2–H140 were used in the previous study (Kasuga *et al*. 1999). H10 was excluded due to doubts about its source. The additional 92 isolates were from newly investigated populations such as Australia, Mexico, Brazil, Argentina, China, Thailand and several European and African countries. The analysed fungal samples included soil isolates as well as veterinary and clinical isolates. Culture conditions, DNA isolation methods, polymerase chain reaction (PCR) and sequencing conditions were published by Kasuga *et al*. (1999). Placement of varieties, i.e. var. *capsulatum*, var. *farciminosum* and var. *duboisii*, was done by medical mycologists who originally isolated fungal strains, based mostly on pathogenicity and morphology. In short, *Hc* var. *capsulatum* mainly caused pulmonary infections. *Histoplasma capsulatum* var. *duboisii* was mainly found in Africa and caused lesions of cutaneous, subcutaneous and osseous tissues. The diameter of a yeast cell of *Hc* var. *duboisii* was 12–15 μm, which was larger than that of *Hc* var. *capsulatum*, which was 2–4 μm in diameter. *Histoplasma capsulatum* var. *farciminosum* caused infections in horses, donkeys and mules. Yeast cells of *Hc* var. *farciminosum* in tissue section were indistinguishable from those of *Hc* var. *capsulatum* (Rippon 1988).

DNA analyses

DNA sequences of partial protein-coding genes used in this study are arf, ADP-ribosylation factor (Lodge *et al*. 1994); H-anti, H antigen precursor (Deepe & Durose 1995); ole, delta-9 fatty acid desaturase (Gargano *et al*. 1995) and tub1, alpha-tubulin (Harris *et al*. 1989).

Phylogenetic analyses [both maximum parsimony and neighbour-joining (NJ)] were performed by using PAUP 4.0b 3a (Sinauer Associates). Most parsimonious (MP) trees were generated by the heuristic search procedure using 500 replications of the random addition sequence option. Nucleotide sites were weighted equally, with character state transformations treated as unordered and of equal cost. Insertions and deletions (indels) that were consistently and unambiguously alignable across all taxa were treated as single evolutionary events by recording a single site within the indel as a multistate character. For MP analysis of the combined data set, characters from the arf, H-anti, ole and tub1 loci were weighted as 1.00, 0.73, 0.82 and 0.56, respectively. These values were inversely proportional to the total number of phylogenetically informative sites per locus. Indices of support (bootstrap values) for internal branches were generated by 500 replications of the bootstrap procedure (Felsenstein 1985). Neighbour-joining trees were generated using the Kimura (1980) two-parameter correction for multiple hits.

Population histories of Latin American A (LAm A) and NAm 2 were inferred by use of the nested clade analysis (NCA). Gene genealogies of each of the four loci, arf, H-anti, ole and tub1, were reconstructed by the statistical parsimony method using software TCS 1.13 (Clement *et al*. 2000). Nested cladograms were constructed according to published methods (Templeton *et al*. 1992; Crandall 1996). The NCA was then performed using the nested cladograms and software GEODIS 2.0 (Posada *et al*. 2000). Isolates belonging to NAm 2 clades were divided into three populations: Midwest (isolates from Indiana and Missouri), South (Arkansas, Texas and Louisiana) and Southeast (Alabama, Georgia, South Carolina and Virginia) and coordinates of St Louis, New Orleans and Atlanta were used, respectively. The LAm A isolates were divided into four populations: Mexico (Mexico and Guatemala), Colombia (Colombia and Panama), Rio de Janeiro (Brazil) and São Paulo (Brazil) and coordinates of Mexico City, Bogotá, Rio de Janeiro and São Paulo were used, respectively. The single Surinamese isolate H145 was not included in the data set due to the lack of population sample from the close vicinity. Inference of population history was made according to the inference key for the nested haplotype tree analysis of geographical distances (Templeton 1998).

Results

Polymorphism summary

Multiple loci used for the recognition of phylogenetic species are preferably functionally and genetically unlinked. The four chosen loci for this study, arf, H-anti, ole and tub1, are likely to be functionally independent but their locations on chromosomes are still unknown. Significant linkage disequilibrium between loci was not detected in the randomly recombining North American population ($P < 0.05$). None of these loci were found on the same bacterial artificial chromosome (BAC) clone in the genomic library used for the ongoing *Histoplasma* genome project (http://www.genome.wustl.edu/projects/hcapsulatum/index.php). Thus, these four loci are likely to have been evolving independently in the *Histoplasma* genome.

Combined DNA sequence data for the four loci gave us 1585 aligned sites, of which 399 were variable and 193 were phylogenetically informative. The 193 phylogenetically informative sites, 399 variable sites and 1585 aligned sites were distributed as follows: for arf, 36 informative sites/78 variable sites/470 aligned sites; for H-anti, 49/85/412; for ole, 44/109/425 and for tub, 64/127/278. Among the 399 variable bases, 147 had indels and 296 had substitutions; of these, 44 sites had both. Introns were clearly more variable than exons in the arf and tub1 loci but not so in the H-anti

Table 2 List of fungal isolates used in this study

Isolate	Variety	Collection number[a] Population	ATCC	RMSCC	Others	Location	Source	Yr isolated	Sender of isolate[b]
H2	*capsulatum*	NAm 2			D14	Georgia/USA	Human/HIV+	1990	E. Keath from P. Connolly
H5	*capsulatum*	NAm 2			D20	Indiana/USA	Human/HIV+	1989	E. Keath from P. Connolly
H6	*capsulatum*	NAm 2			E14	Indiana/USA	Human	1980	E. Keath from P. Connolly
H8	*capsulatum*	NAm 2	26032	1000	M.D. Berliner G217B	Louisiana/USA	Human	1973 or before	
H9	*capsulatum*	NAm 1	38904		Downs	Missouri/USA	Human	1968	E. Keath & G. Kobayashi
H11	*capsulatum*	NAm 2		1003	848	Missouri/USA	Human	1993 or before	G. Kobayashi
H18	*capsulatum*	NAm 2	4745	1019	5-1MD	Missouri /USA	Human	1993 or before	G. Kobayashi
H59	*capsulatum*	LAm B		2349	H-0057-I-10	Bogota/Colombia	Human	1990	A. Restrepo, E. Castaneda & J. McEwen
H60	*capsulatum*	LAm A		2350	H-0057-I-11	Bogota/Colombia	Human/HIV+	1990	A. Restrepo, E. Castaneda & J. McEwen
H61	*capsulatum*	LAm A		2351	H-0057-I-14	Bogota/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H62	*capsulatum*	LAm A		2352	H-0057-I-15	Bogota/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H63	*capsulatum*	LAm A		2353	H-0057-I-18	Bogota/Colombia	Human	1989	A. Restrepo, E. Castaneda & J. McEwen
H64	*capsulatum*	LAm A		2354	H-0057-I-22	Bogota/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H66	*capsulatum*	H66 lineage		2357	13594, GH	Medellin/Colombia	Human	1986	A. Restrepo, E. Castaneda & J. McEwen
H67	*capsulatum*	LAm A		2358	30177, JE	Medellin/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H68	*capsulatum*	LAm B		2359	30318, CH	Medellin/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H69	*capsulatum*	H69 lineage		2360	21402, JVM	Medellin/Colombia	Human	1991	A. Restrepo, E. Castaneda & J. McEwen
H70	*capsulatum*	LAm B		2363	30956, WS	Medellin/Colombia	Human/HIV+	1994	A. Restrepo, E. Castaneda & J. McEwen
H71	*capsulatum*	LAm A		2364	21337, JJM	Medellin/Colombia	Human	1989	A. Restrepo, E. Castaneda & J. McEwen
H73	*capsulatum*	LAm A		2355	H-0057-I-24	Bogota/Colombia	Human	1994	A. Restrepo, E. Castaneda & J. McEwen
H74	*capsulatum*	LAm A		2362	26760, GM	Medellin/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H75	*capsulatum*	LAm B		2365	14056, HC	Medellin/Colombia	Human	1986	A. Restrepo, E. Castaneda & J. McEwen
H76	*capsulatum*	LAm A		2367	T29302, GC	Medellin/Colombia	Human	1993	A. Restrepo, E. Castaneda & J. McEwen
H77	*capsulatum*	NAm 2	10886	2404	C.W. Emmons 6613	Virginia/USA	Brown rat	1940s	
H79	*capsulatum*	NAm 1	11408	2428	C.W. Emmons 6617	Georgia/USA	Skunk	1940s	
H81	*capsulatum*	H81 lineage	26028	2431	M.D. Berliner G184B	Panama	Human	1967 or before	
H82	*capsulatum*	H81 lineage	26029	2432	M.D. Berliner G186A	Panama	Human	1967 or before	
H83	*capsulatum*	H81 lineage	26030	2433	M.D. Berliner G186B	Panama	Human	1967 or before	
H84	*capsulatum*	NAm 2	26320	2434	C.W. Emmons 6623	Georgia/USA	Opossum	1940s	
H85	*capsulatum*	LAm B	28308	2435	CDC B923	Argentina	Soil	1966 or before	
H86	*capsulatum*	NAm 2	32682	2436	A.F. DiSalvo SC74	S. Carolina/USA	Soil	1974?	
H87	*duboisii*	Africa	24294	2437	D. Grigoriu 8107A	Guinea-Liberian Border	Human	1970	
H88	*duboisii*	Africa	32281	2438	RV26821	Belgium	Human	1975 or before	
H90	*farciminosum*	Eurasia	58332	2441	CDC B-3786	Egypt	Horse	1983 or before	
H91	*duboisii*	Africa	24295	2444	D. Grigoriu 8123	Guinea-Liberian Border	Human	1970	
H95	*farciminosum*	Eurasia	58333	2442	CDC B-3787	Egypt	Horse	1983 or before	
H96	*farciminosum*	Eurasia	60358	2443	A.F. DiSalvo 85-1610	India	Horse	1985?	
H97	*capsulatum*	NAm 2		2472	0001	Alabama/USA	Human	1995 or before	W. Dismukes, S. Moser & B. Hines
H126	*capsulatum*	NAm 1				Missouri/USA	Human/HIV+	1987	G. Kobayashi
H127	*capsulatum*	NAm 1				Missouri/USA	Human/HIV+	1987	G. Kobayashi
H130	*capsulatum*	NAm 2		2767	15	Alabama/USA	Human	1995 or before	W. Dismukes, S. Moser & B. Hines
H137	*duboisii*	Africa	28536		Kwon-Chung Hd27	Zaire	Human	1962	
H138	*capsulatum*	NAm 2	22635		Kwon-Chung T-3-1	Arkansas/USA	soil	1975	
H139	*capsulatum*	NAm 2	22636		Kwon-Chung T-4-2	Arkansas/USA	soil	1975	
H140[c]	*not known*	H140 lineage			MK9500885	Maryland/USA/Peru	Owl monkey	1997	G. Miller
H141	*capsulatum*	LAm A		4710	CBS 207.55	Indonesia	Human	1955	
H142	*capsulatum*	Eurasia		4718	CBS 214.53	England	Human	1940	
H143	*capsulatum*	Africa		4719	CBS 287.54	South Africa	Human	1954	

Table 2 *Continued*

Isolate	Variety	Collection number[a] Population	ATCC	RMSCC	Others	Location	Source		Yr isolated	Sender of isolate[b]
H144	*capsulatum*	Netherlands		4720	CBS 381.65	Netherlands	Human		1965	
H145	*capsulatum*	LAm A		4721	CBS 682.89	Surinam	Human/HIV+		1989	
H146	*capsulatum*	LAm A		4722	CBS 719.79	Brazil	Human		1979	
H147	*duboisii*	Africa		4723	CBS 175.57	Senegal	Human		1957	
H148	*farciminosum*	Eurasia		4724	CBS 205.35		Horse?		1935	(G. Puntoni)
H149	*capsulatum*	LAm A			N.A.G.	Sao Paulo/Brazil	Human/HIV+		1996	C. da S. Lacaz
H150	*capsulatum*	LAm A			C.S.	Sao Paulo/Brazil	Human		1996	C. da S. Lacaz
H151	*capsulatum*	LAm A			M.A.C.S.	Sao Paulo/Brazil	Human/HIV+		1997	C. da S. Lacaz
H152	*capsulatum*	LAm A			W.S.A.	Sao Paulo/Brazil	Human/HIV+		1997	C. da S. Lacaz
H153	*capsulatum*	H153 lineage			G.M.L.	Sao Paulo/Brazil	Human		1997	C. da S. Lacaz
H154	*capsulatum*	LAm A			P.L.F.	Sao Paulo/Brazil	Human		1997	C. da S. Lacaz
H155	*capsulatum*	LAm A			I.A.F.	Sao Paulo/Brazil	Human/HIV+		1998	C. da S. Lacaz
H156	*capsulatum*	LAm A			S.P.G.	Sao Paulo/Brazil	Human		1997	C. da S. Lacaz
H157	*capsulatum*	Australia			12-6	Adelaide/Australia	Human	caver	1970s	D. Muir
H158	*capsulatum*	Australia			12-12	Wee Jasper Cave/AU	soil + bat guano	Mouse passage	1984	D. Muir
H159	*capsulatum*	Australia			12-13	Westmead/Australia	Human	Jasper visitor	1984	D. Muir
H160	*capsulatum*	Australia			12-17	Brisbane/Australia	Human/HIV+	dissemi/bowel	1988	D. Muir
H161	*capsulatum*	Australia			12-20	Penrith/Australia	Human/HIV+	dissemi/blood	1990	D. Muir
H162	*capsulatum*	LAm B		4725	Sali	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H163	*capsulatum*	LAm B		4726	Carrillo	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H164	*capsulatum*	LAm B		4727	130504	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H165	*capsulatum*	LAm B		4728	125343	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H166	*capsulatum*	LAm B		4729	Fernandez	Argentina	Human/HIV+	dissemi/skin	1998–1999	R. Negroni
H167	*capsulatum*	H167 lineage		4730	135483	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H168	*capsulatum*	LAm B		4731	M. Almeida	Argentina	Human/HIV+	dissemi/skin	1998–1999	R. Negroni
H169	*capsulatum*	LAm B		4732	Villarroel	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H170	*capsulatum*	LAm B		4733	Fontana	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H171	*capsulatum*	LAm B		4734	140147	Argentina	Human/HIV+	dissemi/blood	1998–1999	R. Negroni
H172	*capsulatum*	LAm B		4735	A. Gomez	Argentina	Human/HIV+	dissemi/skin	1998–1999	R. Negroni
H173	*farciminosum*	NAm 2	32136	4736	CBS 176.57		Horse		1957	
H174	*farciminosum*	Eurasia	32138	4739	CBS 477.64	Poland	Horse		1959	
H175	*farciminosum*	Eurasia	32139	4740	CBS 478.64	Poland	Horse		1962	
H176	*capsulatum*	Netherlands		4741	CBS 243.69	Netherlands	Human		1969	
H177	*capsulatum*	Eurasia		4742	D16a	Beijing/China	Human			Z. Qin
H178	*capsulatum*	Eurasia		4743	D16b	Beijing/China	Human			Z. Qin
H179	*capsulatum*	NAm 2	10230	4744	C.W. Emmons 6510	USA	Human			Z. Qin
H181	*capsulatum*	NAm 2		4746	D16e	Beijing/China	Human			Z. Qin
H185[c]	*not known*	H140 lineage			2612	Maryland/USA/Peru	Owl monkey	spleen	1999	G. Miller
H187	*duboisii*	Africa	76543	4748	M236	Nigeria	bat cave		1991	
H188	*capsulatum*	LAm A	11656	4749	CDC A-721	Panama	Soil			
H189	*farciminosum*	Africa	28798	4750						
H190	*farciminosum*	Eurasia	32140	4751	Mariat 848					
H191	*farciminosum*	Eurasia	32141	4752	Mariat 92					
H192	*capsulatum*	Eurasia	64799	4753	Thammayya ST2483	India	Human			
H193	*farciminosum*	Eurasia	58334	4754	Sleim 2803	Egypt	Horse		1984 or before	
H194	*farciminosum*	Eurasia	58335	4755	Sleim 2801	Egypt	Horse		1984 or before	
H196	*capsulatum*	LAm A		4757	78642	Rio de Janeiro/Brazil	Human/HIV–	BL	1997	R. Oliveira-Zancope
H197	*capsulatum*	LAm A		4758	84392	Rio de Janeiro/Brazil	Human/HIV+	dissemi/blood	1999	R. Oliveira-Zancope

Table 2 *Continued*

Isolate	Variety	Collection number[a] Population	ATCC	RMSCC	Others	Location	Source		Yr isolated	Sender of isolate[b]
H198	*capsulatum*	LAm A		4759	84422	Rio de Janeiro/Brazil	Human/HIV+	dissemi/blood	1999	R. Oliveira-Zancope
H199	*capsulatum*	LAm A		4760	3237	Rio de Janeiro/Brazil	Human/HIV–	chronic/sputum	1988	R. Oliveira-Zancope
H200	*capsulatum*	LAm A		4761	RPS09	Rio de Janeiro/Brazil	Soil at chicken house A		1983	R. Oliveira-Zancope
H201	*capsulatum*	LAm A		4762	RPS45	Rio de Janeiro/Brazil	Soil at chicken house B		1983	R. Oliveira-Zancope
H202	*capsulatum*	LAm A		4763	RS36	Rio de Janeiro/Brazil	Wild rodent		1983	R. Oliveira-Zancope
H203	*capsulatum*	LAm A		4764	RS93	Rio de Janeiro/Brazil	matachirus opposum		1983	R. Oliveira-Zancope
H204	*capsulatum*	Eurasia	66368		Randhawa VPCI/192	India	Human	Semen		
H205	*capsulatum*	Eurasia			HP4, NIH 37-384-23	Thailand	Human	blood	1994	N. Poonwan & Y. Mikami
H206	*capsulatum*	Eurasia			HP9, NIH 36-332-61	Thailand	Human	liver	1994	N. Poonwan & Y. Mikami
H207	*capsulatum*	Eurasia			HP12, NIH 36-395-15	Thailand	Human	blood	1993	N. Poonwan & Y. Mikami
H208	*capsulatum*	Eurasia			HP16, NIH 37-1676-85	Thailand	Human	BAL	1994	N. Poonwan & Y. Mikami
H209	*capsulatum*	Eurasia			HP18, NIH 37-466-131	Thailand	Human	lymph node	1994	N. Poonwan & Y. Mikami
H210	*capsulatum*	Eurasia			HP23, NIH 39-205-205	Thailand	Human	skin	1996	N. Poonwan & Y. Mikami
H211	*capsulatum*	LAm A			HP13, NIH 36-502-23	Thailand	Human/HIV+	BAL	1996	N. Poonwan & Y. Mikami
H212	*farciminosum*	Eurasia			IFM 5418, 848.63 IP	Algeria	Horse		1940s	Institut Pasteur
EH46	*capsulatum*	LAm A		4707		Guerrero/Mexico	Human		1979	M.L. Taylor
EH53	*capsulatum*	LAm A		4708		Hidalgo/Mexico	Human		1977	M.L. Taylor
EH303	*capsulatum*	LAm A		4667		Guatemala	Human			M.L. Taylor
EH304	*capsulatum*	LAm A		4668		Guatemala	Human		1991	M.L. Taylor
EH315	*capsulatum*	EH315 lineage		4669		Guerrero/Mexico	bat		1994	M.L. Taylor
EH316	*capsulatum*	LAm A		4670		Guerrero/Mexico	Human/HIV+		1993	M.L. Taylor
EH317	*capsulatum*	LAm A		4671		Morelos/Mexico	Human/HIV+		1992	M.L. Taylor
EH319	*capsulatum*	LAm A		4688		Mexico City/Mexico	Human/HIV+			M.L. Taylor
EH325	*capsulatum*	LAm A		4689		Chiapas/Mexico	Human/HIV+			M.L. Taylor
EH332	*capsulatum*	LAm A		4690		Guatemala	Human		1994	M.L. Taylor
EH333	*capsulatum*	LAm A		4691		Guatemala	bird guano		1991	M.L. Taylor
EH359	*capsulatum*	LAm A		4692		Oaxaca/Mexico	Human		1995	M.L. Taylor
EH362	*capsulatum*	LAm A		4693		Guatemala	zanate excreta		1996	M.L. Taylor
EH363	*capsulatum*	LAm A		4694		Guatemala	Human		1996	M.L. Taylor
EH364	*capsulatum*	LAm A		4695		Guatemala	Human		1996	M.L. Taylor
EH372	*capsulatum*	LAm A		4696		Morelos/Mexico	bat intestine		1997	M.L. Taylor
EH373	*capsulatum*	LAm A		4697		Morelos/Mexico	bat lung		1997	M.L. Taylor
EH374	*capsulatum*	LAm A		4698		Morelos/Mexico	bat spleen		1997	M.L. Taylor
EH376	*capsulatum*	LAm A		4699		Morelos/Mexico	bat lung		1997	M.L. Taylor
EH377	*capsulatum*	LAm A		4700		Morelos/Mexico	bat lung		1997	M.L. Taylor
EH378	*capsulatum*	LAm A		4701		Morelos/Mexico	bat lung		1997	M.L. Taylor
EH379	*capsulatum*	LAm A		4702		Estado de Mexico	Human		1996	M.L. Taylor
EH383	*capsulatum*	LAm A		4703		Morelos/Mexico	bat lung		1997	M.L. Taylor
EH391	*capsulatum*	LAm A		4704		Morelos/Mexico	bat liver		1997	M.L. Taylor
EH394	*capsulatum*	LAm A		4705		Oaxaca/Mexico	bat lung		1997	M.L. Taylor
EH408	*capsulatum*	LAm A		4706		Puebla/Mexico	bat lung etc.		1998	M.L. Taylor
Blastomyces dermatitidis			60915		D. Stevens A	South Carolina/USA	Human		1970	

[a]ATCC, American Type Culture Collection, Rockville, MD, USA; RMSCC, Roche Molecular Systems Culture Collection, Alameda, CA, USA; CDC, Centers for Disease Control, Atlanta, GA, USA. CBS, Centraalbureau voor Schimmelcultures, Baarn, The Netherlands.

[b]E. Keath: Saint Louis University, St. Louis, Missouri, USA. P. Connolly: Indiana University Medical Center, Indianapolis, Indiana, USA. G. Kobayashi: Washington University School of Medicine, St. Louis, Missouri, USA. W. Dismukes, S. Moser & B. Hines: Univ. of Alabama, Birmingham, Alabama, USA. D. Muir: The Royal North Shore Hospital, S. Leonards, Australia. Restrepo, Castaneda, McEwen, Miller, Lacaz, Negroni, Qin, Oliveira-Zancope, Poonwan, Mikami, ML Taylor are co-authors of this paper.

[c]DNA was directly isolated from a yeast-infested monkey liver. The monkey was wild-caught in Peru, then kept in Maryland USA.

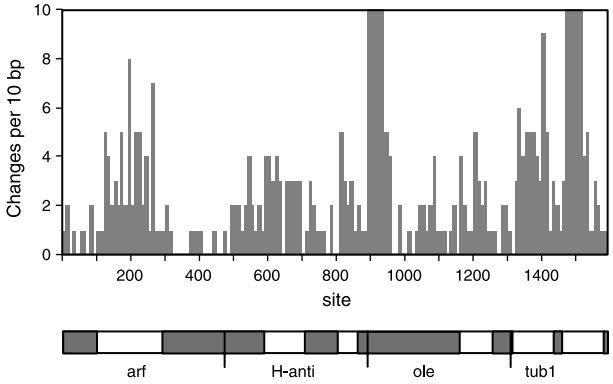

Fig. 1 Distribution of polymorphic sites per 10 bp in the four unlinked loci. Exons in the gene map are shaded in grey and introns are not shaded. Locations of loci are: arf, 1–470; H-anti, 471–882; ole, 883–1307 and tub1, 1308–1585.

and ole loci (Fig. 1). Most indels in the introns were just 1–3 bp but isolates H163 and H172 had a 50-bp deletion in the third intron of tub1 gene. In exons, there were 107 base substitutions, of which 28 were nonsynonymous substitutions. Only one large deletion of 54 bp, which corresponded to 18 amino acids, was found in the first exon of the ole gene in isolate H167. Of 137 isolates, 80 multilocus genotypes were identified, an increase of 47 multilocus genotypes over the previous study.

Phylogeny of multilocus genotypes and recognition of phylogenetic species

Maximum parsimony and NJ methods were used to analyse phylogenetic relationships among the 80 unique multilocus genotypes. In each of the four gene trees, isolates tended to cluster together according to their geographical origin. Multilocus genotypes from each of the Australian, Dutch and Eurasian populations as well as five previously identified clades, NAm 1, NAm 2, LAm A, LAm B and Africa, formed a monophyletic group in at least one of the four loci in MP bootstrap consensus trees and NJ trees (Figs 2 and 3).

Dettman *et al*. (2003a,b) have recently developed an approach to identify phylogenetic species from multiple-gene genealogies: a clade was recognized as an independent evolutionary lineage if it was well supported in at least one single-locus genealogy, as judged by both MP bootstrap values of at least 70% (Hillis & Bull 1993) and Bayesian posterior probabilities of at least 0.95 (Rannala & Yang 1996), and was also not contradicted in any other single-locus genealogy at the same level of support (Dettman *et al*. 2003a). Boundaries of fungal species recognized by this approach were shown to be in good agreement with those identified by a mating test (Dettman *et al*. 2003b).

The criteria proposed by Dettman *et al*. (2003a,b) using bootstrap support were used to define phylogenetic species of *Histoplasma*. All of the uncontradicted branches in the semistrict consensus tree produced from the four gene trees (Fig. 2a–d) were indicated as bold branches in the combined MP tree (Fig. 4a). Thirteen groups were uncontradicted in the semistrict consensus tree and were supported by bootstrap values of at least 70%. Four of the 13 groups correspond to the previously identified clades, LAm B, NAm 1, NAm 2 and Africa (*Hc* var. *duboisii*), and two groups are newly investigated Australian and Dutch populations, all of which were supported by bootstrap values of at least 99%. In the data set of Dettman *et al*. (2003a,b) all of the branches supported by bootstrap values of 95% or larger were also supported by a Bayesian posterior probability of 1.0. Thus, these six groups of *Histoplasma* should be recognized as phylogenetic species by the criteria proposed by Dettman *et al*. (2003a,b). H167 from Argentina and NAm 1 formed a well-supported clade; however, we judged that H167 and NAm 1 do not form a single phylogenetic species due to the large genetic and geographical distance observed between them. There are two well-supported branches in the African clade. It might be that these two branches correspond to two independent phylogenetic species. We, however, decided not to divide the African clade due to the insufficient sample size. Likewise, two isolates, EH317 and EH325 from Morelos and Chiapas Mexico, respectively, formed a well-supported clade. In H-anti and tub1 loci, these two isolates are distinct from other LAm A isolates (Figs 2b and d and 3b and d). As EH317 and EH325 share alleles with other LAm A isolates at arf and ole loci, we judged that these two isolates do not form an independent phylogenetic species. The LAm A clade identified previously (Kasuga *et al*. 1999) formed a monophyletic group in the combined MP tree and NJ tree with bootstrap support of 87 and 96% (Fig. 4) as well as in gene genealogy of H-anti with bootstrap support of 92% (Fig. 2b). However, the LAm A clade is not recognized in the semistrict consensus tree produced from the four gene trees due to the discordance in tree topologies in the arf and tub1 loci. LAm A cannot be recognized as a phylogenetic species according to the criteria proposed by Dettman *et al*. (2003a). Nevertheless, we maintain the LAm A clade as the most diverse phylogenetic species including isolates from Brazil to Mexico. The Eurasian clade is likely to correspond to a genetically isolated population arising from LAm A individuals. In order to maintain the LAm A clade as an monophyletic phylogenetic species, we did not give phylogenetic species status to the Eurasian clade.

Of the 80 multilocus genotypes, 73 could be included in one of eight clades or seven phylogenetic species. The remaining seven multilocus genotypes (i.e. H140, H81, H66, H69, H153, H167 and EH315) do not belong to any of the eight clades and are also distant from each other (hereafter called lone lineages, indicated with filled circles in Figs 2, 3 and 4).

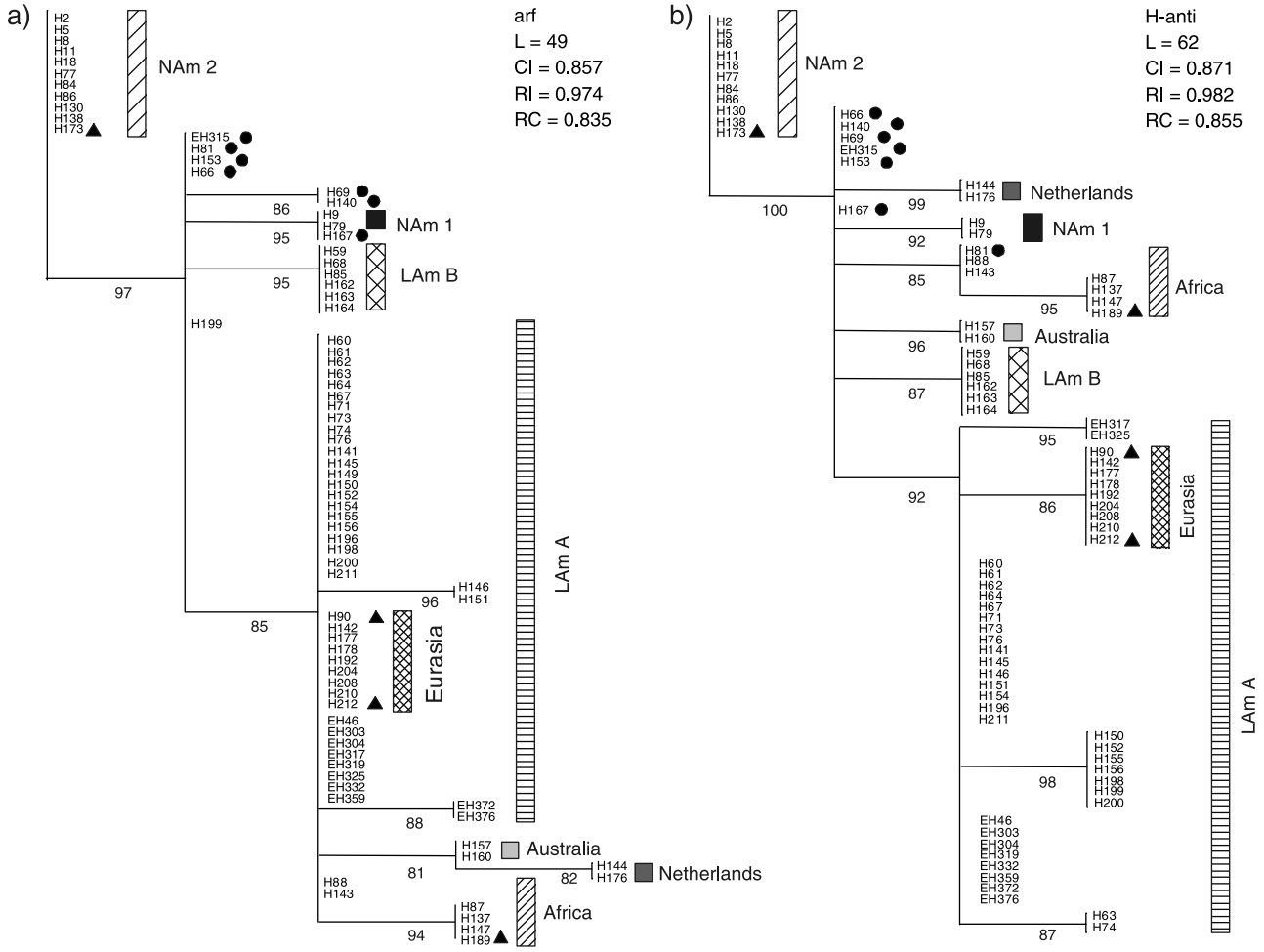

Fig. 2 A bootstrap consensus tree resulting from maximum parsimony analysis of DNA sequences of the 80 multilocus genotypes from each of the four gene regions sequenced. When more than one isolate had the same genotype, the isolate name with the smallest number was used for the genotype and shown here. Figure 4 shows all of the isolates. L, tree length; CI, consistency index; RI, retention index; RC, rescaled consistency index. Numbers below branches represent indices of support based on 500 bootstrap replications of the parsimony procedure. Branches with bootstrap support smaller than 70% were reduced to polytomies. ●, Lone lineages; ▲, isolates of *Histoplasma capsulatum* var. *farciminosum*. Abbreviations of groups are listed in Table 1.

In the four gene genealogies, as well as the combined genealogy, relationships between the clades were unresolved, resulting in a star phylogeny (Slatkin & Hudson 1991). The unresolved branching order may be the result of hybridization and intralocus recombination or an ancient radiation. To test for recombination, we performed split decomposition analysis (Dress *et al*. 1996). Recombination breaks, which separate linked clusters, were not detected in any loci (data not shown). This result indicates that our data set does not include mosaic genotypes generated by recent hybridization followed by intralocus recombination among alleles associated with the diverged clades. Internal branches in the combined NJ tree are, on average, only one-tenth of the length of branches leading to the six phylogenetic species and six lone lineages (the Netherlands clade and H167 lineage were excluded due to their obvious close association to one of the clades). This finding implies a rapid radiation of *Histoplasma* species over a short period of time.

In most cases, DNA polymorphism and clades were strongly associated with geographical locations. For each clade, observed population structure and known traits are detailed in the following paragraphs.

Mexican, Central and South American (Latin American) populations

Histoplasmosis is endemic from Mexico to Argentina. In our previous study, 21 clinical isolates from Colombia and Panama and one soil isolate from Argentina were examined and three distinct clades were identified, LAm A, LAm B and the H81 lineage from Panama. In this

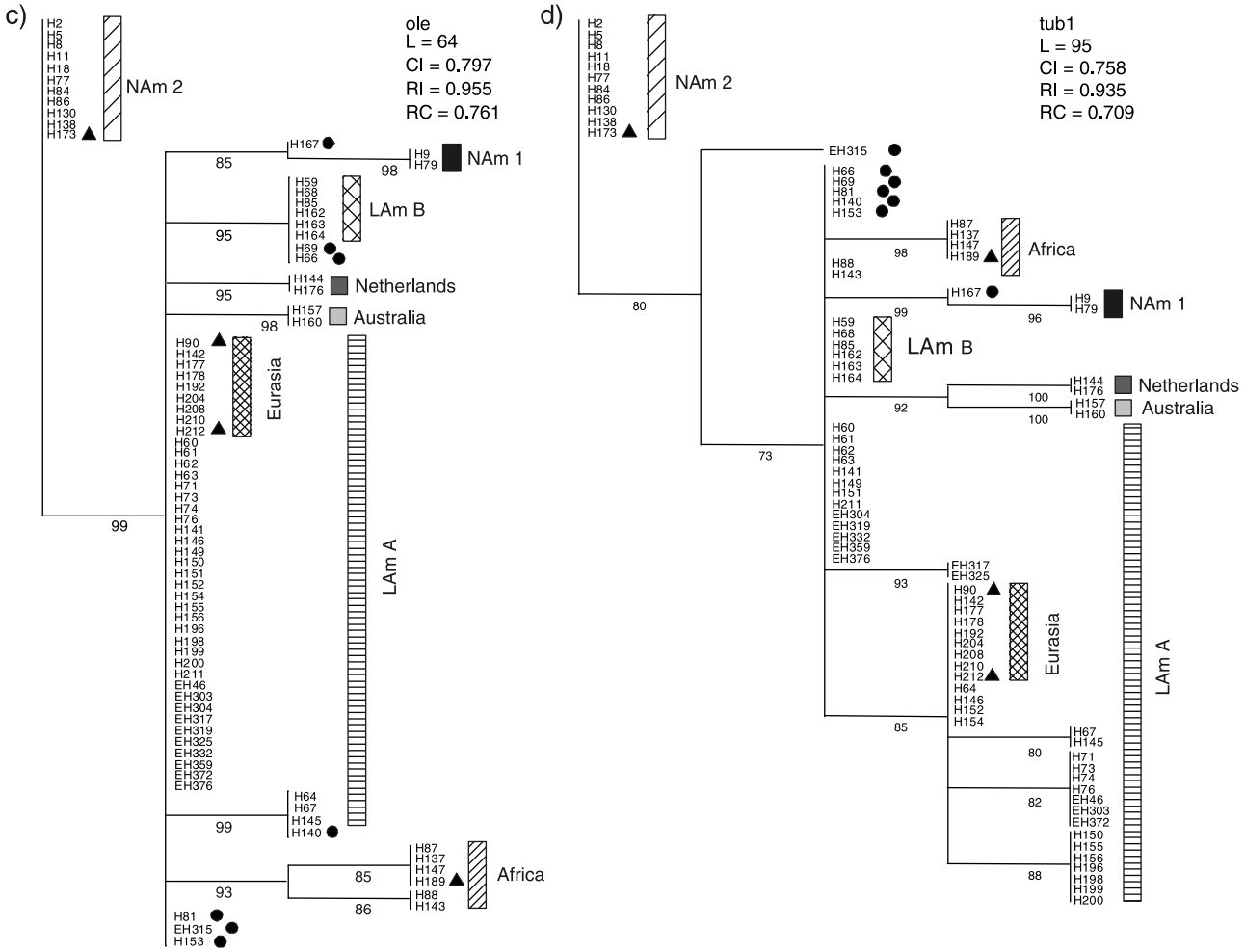

Fig. 2 *(Continued)*

research, 38 clinical isolates and 19 environmental isolates from Mexico, Guatemala (Reyes-Montes *et al*. 1999; M.L. Taylor *et al*. 1999), Panama, Surinam, Brazil (de Medeiros Muniz *et al*. 2001) and Argentina were added and the population structure of *Histoplasma* in Latin American countries was re-examined. *Histoplasma capsulatum* isolates in Latin America are the most phylogenetically diverse. Most (42 of 44) isolates from Mexico, Guatemala, Surinam and Brazil were found in the LAm A whereas most (11 of 12) isolates from Argentina were found in the LAm B. One soil isolate from Panama (H188) was distant from the H81 lineage from Panama but belonged to the LAm A. In Colombia, clinical isolates belonging to both LAm A (10 isolates) and LAm B (four isolates) were identified.

In addition to the two major Latin American clades A and B, seven lone lineages, each containing one multilocus genotype, were found in the Latin American countries. The first lone lineage comprises two DNA samples (H140 and H185) from unculturable yeasts infesting the internal organs of Peruvian owl monkeys (Miller & Owens 1999) (the H140 lineage). Five other lone lineages are represented by isolates H66 and H69, both from Colombia, H153 from Brazil, H167 from Argentina and EH315 from Mexico. The H81 lone lineage from Panama (Berliner 1968) consists of three isolates G184A, G186A and G186B (corresponding to H81, H82 and H83, respectively) but just one multilocus genotype. So far, no other isolate closely related to the H81 lineage has been identified. The widely studied type cultures (e.g. Carr & Shearer 1998; Sebghati *et al*. 2000), from which the genome sequence of H82 (G186A) is being determined, appear to be distantly related to the genotypes responsible for histoplasmosis in North America or Latin America.

Eurasian population

Histoplasmosis is not endemic to Europe although it has been reported from wild badgers (Bauder *et al*. 2000). As a consequence of HIV infection, the incidence of histoplasmosis in Europe is increasing but most cases are attributable

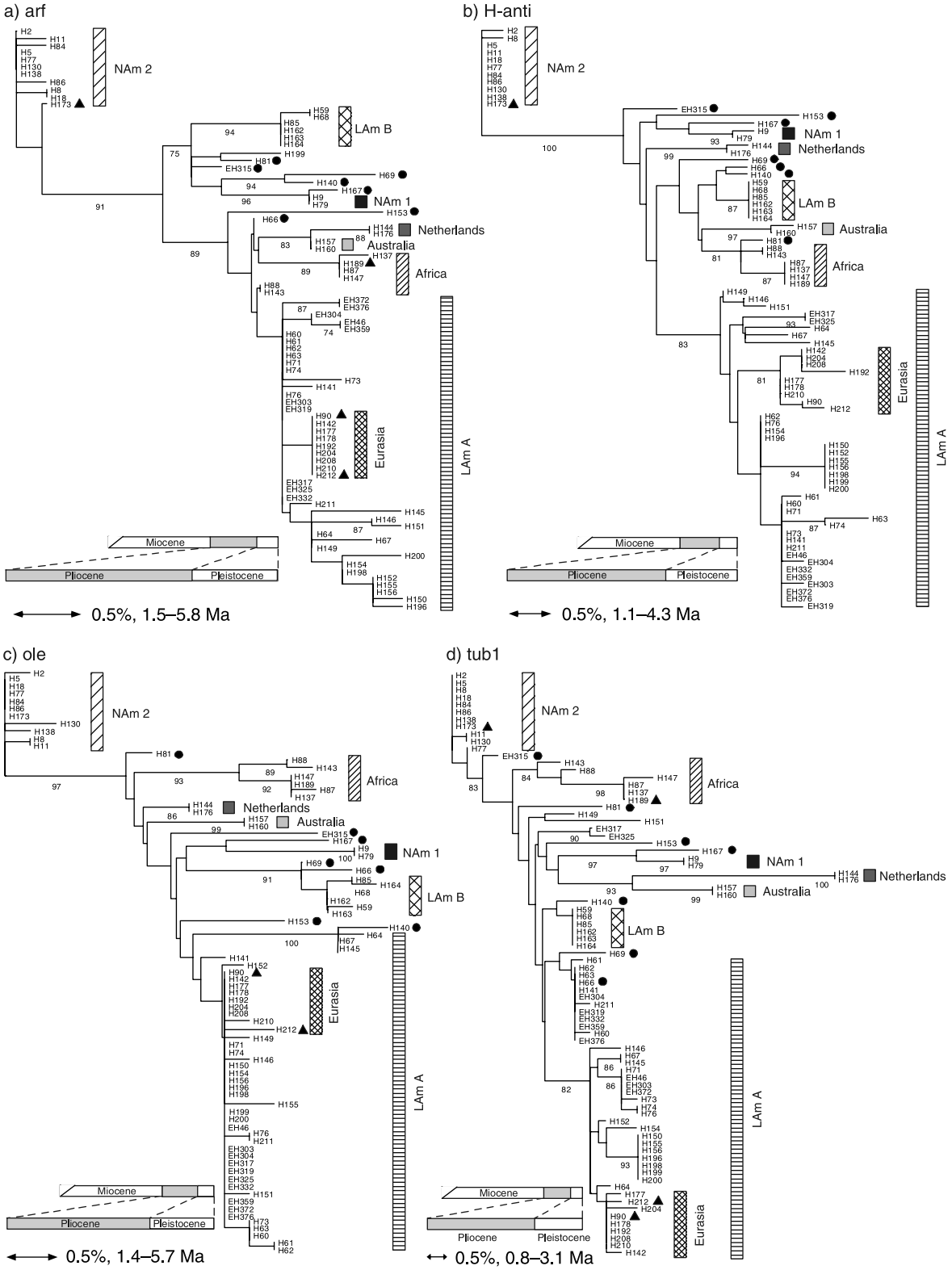

Fig. 3 A neighbour-joining representation of the data set used for Fig. 2 (see Fig. 2 for details). Branch lengths are proportional to Kimura's two-parameter distance. Bootstrap values ≥ 70% are shown below branches. Three scales are shown at the lower left for each of the gene trees. The top and middle scales represent durations of the Pliocene [5.3 million years ago (Ma)] and Pleistocene (1.8 Ma) epochs based on the *Histoplasma–Blastomyces* divergence of 127.8 and 31.8 Ma, respectively. The bottom scale represented by an arrow indicates 0.5% DNA substitutions and corresponding durations in million years.

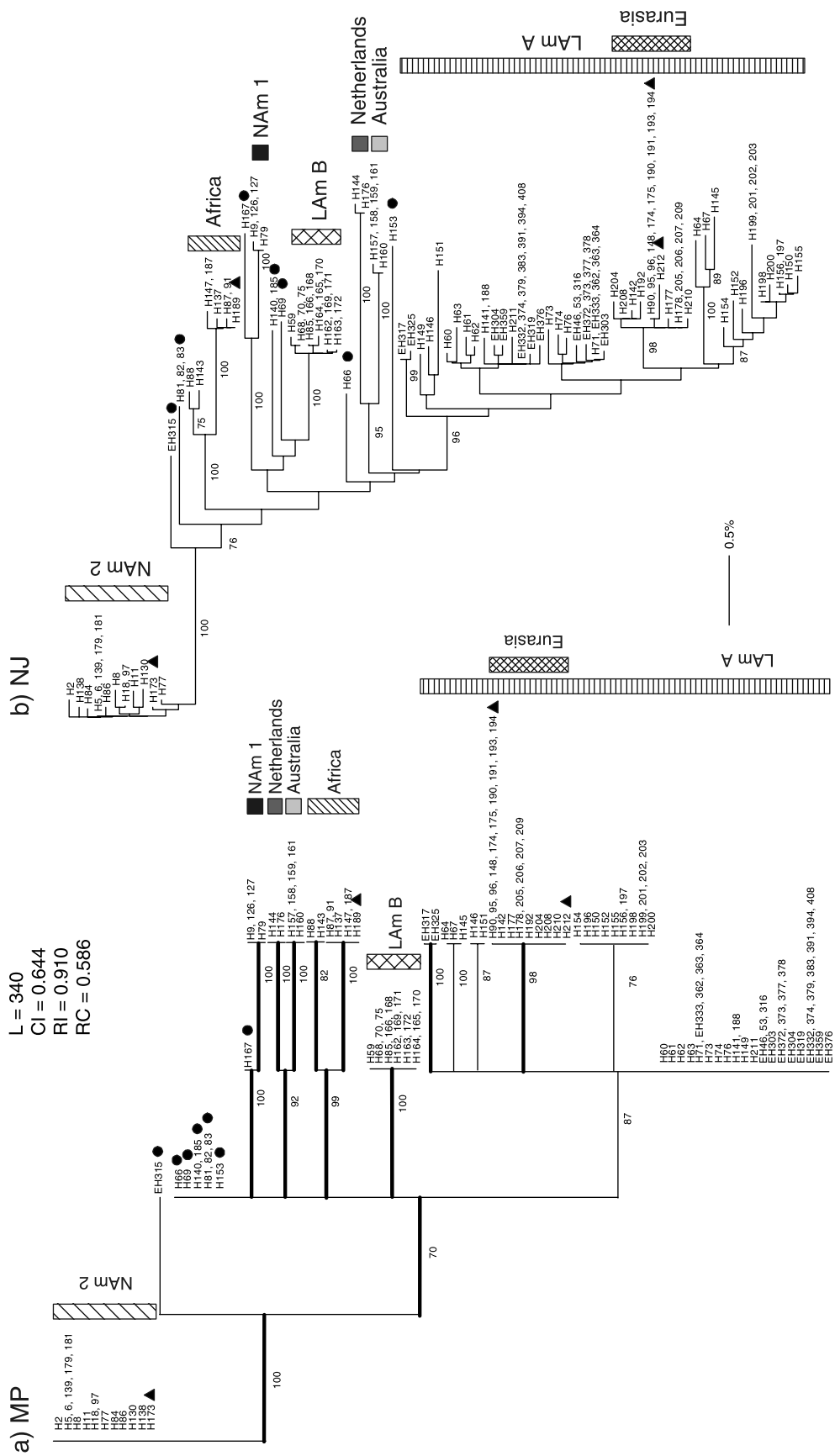

Fig. 4 (a) A bootstrap consensus tree derived from maximum parsimony analysis of the combined data of the four loci. All isolates sharing each of the 80 multilocus genotypes are shown. Numbers below branches represent indices of support based on 500 bootstrap replications of the parsimony procedure; only values ≥ 70% are shown. Branches that were well supported by at least one locus but not contradicted by another locus in Fig. 2 are shown in bold. L, tree length; CI, consistency index; RI, retention index; RC, rescaled consistency index; MP, most parsimonious. (b) Neighbour-joining (NJ) tree from analysis of the combined data of the four loci. Branch lengths are proportional to the Kimura's two-parameter distance. Bootstrap values are also indicated. ▲, Isolates of *Histoplasma capsulatum* var. *farciminosum*; ●, lone lineages. The aligned data set is available at http://www.treebase.org; Study accession number 5960.

to endogenous reactivation of a latent infection acquired overseas in endemic areas (Manfredi *et al*. 1994). The disease is endemic to southeast Asia and India but the incidence and prevalence of histoplasmosis have not been extensively described. Clinical isolates, one from England (H142), two from China (H177 and H178), six from Thailand (H205–H210) and two from India (H192 and H204), formed a homogeneous monophyletic group with a bootstrap support of 98% within the LAm A clade (Fig. 4a). This Eurasian group corresponds to the Asian type *H. capsulatum* which was identified by Tamura *et al*. (2002) in their phylogeny based on the internal transcribed spacer region. One Chinese isolate (H178, Beijing) and four of the six Thai isolates shared one multilocus genotype. Judging from its close genetic distance to Indian isolates, and histoplasmosis not being endemic to England, the English isolate H142 obtained in 1940 was likely to have been acquired in India. A single Chinese isolate (H181, Beijing) was located in the NAm 2 clade, despite the fact that the patient from whom H181 was isolated had never been out of China. Similarly, the Thai isolate H211, which had an RAPD pattern unlike other Thai isolates (Poonwan *et al*. 1998), was found in the LAm A clade but outside the Eurasian subclade. H181 and H211 might represent cases of indirect acquisition of a foreign fungus from infected zoo animals or contaminated body parts of New World animals used in Asian medicine, e.g. bear gall bladders.

North American populations

Two discrete phylogenetic species, NAm 1 and NAm 2, have been reported in North America (Spitzer *et al*. 1990; Kasuga *et al*. 1999). Our previous study showed that these two phylogenetic species were as distant from each other as from any other *Histoplasma* clades in the world. Genetic diversities observed within NAm 1 and NAm 2 clades are much smaller than that observed in the LAm A clade. In this study, two lone lineages of Latin American isolates from Argentina (H167) and Mexico (EH315) were found to be distantly associated with the NAm 1 and NAm 2, respectively.

African population

On the African continent, two clinically distinct forms of histoplasmosis are known, one caused by *Hc* var. *duboisii* and the other by *Hc* var. *capsulatum* (Kwon-Chung & Bennett 1992; Rippon 1988). The disease histoplasmosis duboisii is characterized by cutaneous and subcutaneous lesions whereas histoplasmosis capsulatii is characterized by infection of the lung. A single African clade was formed from clinical isolates of *Hc* var. *duboisii* from the Guinea–Liberian border (H87 and H91), Zaire (H137), Belgium (H88, probably from a former Belgian colony), Senegal (H147), Nigerian soil (H187) (Gugnani *et al*. 1994) and a *Hc* var. *capsulatum* isolate from South Africa (H143). The finding of one *Hc* var. *capsulatum* isolate (H143) in the *Hc* var. *duboisii* clade contradicts the traditional view that two distinct forms of histoplasmosis exist in Africa. It is likely that histoplasmosis in Africa is caused by a monophyletic group of *H. capsulatum* isolates which included isolates presently assigned to *Hc* var. *duboisii* and *Hc* var. *capsulatum*.

Australian and Netherlands (Indonesian?) populations

Histoplasmosis in Oceania is rare and poorly known. Incidences seem to be largely restricted to people who have visited bat-infested caves (Isbister *et al*. 1976; Harden & Hunt 1985). One soil and four clinical isolates from Australia were very homogeneous genetically, with only one polymorphic site in the four gene regions despite their diverse geographical origins. Two clinical isolates received from the Netherlands (of Indonesian origin?) formed a distinct clade which was the sister group to the Australian clade with bootstrap support of 92% (Fig. 4a). The two Dutch isolates H144 and H176 had been deposited in the Centraalbureau voor Schimmelcultures (CBS) in the Netherlands in 1965 and 1969, respectively. No information on the history of patients or the geographical origin of the isolates is available. As histoplasmosis is not endemic to the Netherlands, it is likely that H144 and H176 originated in former Dutch colonies, possibly in Indonesia. On the contrary, an Indonesian clinical isolate (H141), deposited in the CBS in 1955, was found to belong to the LAm A clade. Isolate H141 had a multilocus genotype identical to that of a Panamanian soil isolate (H188, LAm A), suggesting that the infection caused by H141 might have originated in Panama.

Histoplasma capsulatum *var*. farciminosum

Equine histoplasmosis is not caused by a monophyletic group of *Hc* var. *farciminosum* isolates. Four multilocus genotypes were found among 13 isolates obtained from cases of histoplasmosis farciminosi and identified as *Hc* var. *farciminosum* from various geographical locations. One isolate was found in the African clade (H189), another in the NAm 2 clade (H173) and the other 11 isolates, of which 10 had identical multilocus genotypes, were found in the Eurasian clade (filled triangles in Figs 2–4). No other isolates had the multilocus genotypes of H189 or H173 so the possibility of cross contamination in our laboratory can be excluded. It appears that the disease histoplasmosis farciminosi is just a form of histoplasmosis affecting horses rather than humans and may be caused by isolates originating independently from at least three *Histoplasma* clades.

Dating the divergence time of Histoplasma capsulatum

Assuming that *H. capsulatum* forms a star phylogeny, when did the radiation of *H. capsulatum* start? In order to date the radiation event, we have to rely on the molecular clock hypothesis and DNA mutation rates extrapolated from other systems because no palaeontological data are available for *H. capsulatum*. Under the neutral theory of evolution, mutations in any given DNA sequence accumulate at an approximately constant rate as long as the DNA sequence retains its original function. In our data set, only a small portion of substitutions (28 of 296 substitutions) represents nonsynonymous substitutions. Moreover, in any of the four loci, nonsynonymous substitutions per site were always significantly fewer ($P < 0.05$) than synonymous substitutions per site as judged by pairwise comparisons between isolates, suggesting that the nonsynonymous polymorphisms in any of the four loci were not under positive Darwinian selection (Nei & Kumar 2000). Two coalescent theory-based tests, Tajima's test (Tajima 1989) and the HKA test (Hudson *et al*. 1987), were used to detect natural selection at the four genetic loci of the four largest phylogenetic species, LAm A, LAm B, NAm 2 and Africa. The HKA test failed to detect deviation from the neutral mutation hypothesis in any of the four loci in the four clades (Rozas & Rozas 1999). Tajima's test also did not detect deviation from neutral evolution with one exception at the ole locus in the LAm A clade ($P < 0.05$). Overall DNA substitutions in the four loci did not deviate significantly from neutral evolution.

We have previously estimated the divergence time of *H. capsulatum* and *Blastomyces dermatitidis* from a small subunit rRNA gene tree (Kasuga *et al*. 2002). The value was strongly dependent on the algorithm used to estimate divergence time and the calibration time points. When the divergence of Eurotiomycetes (plectomycetes) and Sordariomycetes (pyrenomycetes) was set to 400 Ma (T.N. Taylor *et al*. 1999) and either of two methods of divergence estimation were used, the Langley Fitch algorithm, which assumes rate constancy, or a nonparametric rate-smoothing algorithm, which does not assume rate constancy (Sanderson 1997), divergence times for *Histoplasma* and *Blastomyces* were 32 and 128 Ma, respectively. To estimate the nucleotide substitution rates at the two protein loci for which we have data for both *H. capsulatum* and *B. dermatitidis*, arf and tub1, the genetic diversity was compared with these divergence times. For arf, nucleotide substitution was estimated to be between 0.86×10^{-9} and 3.43×10^{-9} substitutions per base per year and tub1 was estimated to be between 1.63×10^{-9} and 6.56×10^{-9} substitutions per base per year. Absolute substitution rates at the ole and H-anti loci could not be estimated due to the unavailability of corresponding gene sequences in *B. dermatitidis*. To estimate the time of the radiation of *Histoplasma* species, we needed to estimate the amount of DNA substitution that had accumulated among the populations. We used the nucleotide diversity (π), which is the average pairwise distance between isolates (see the footnote to Table 3) (Li 1997). Inclusion of individuals with identical genotypes leads to the underestimation of population richness, which is a concern because *H. capsulatum* propagates clonally as

Table 3 Estimation of substitution rates from nucleotide diversity values

	Length/bp	Nucleotide diversity π*	π/μ	DNA substitution rate μ**
arf	470	0.0147	1.96×10^{7}	(0.86×10^{-9})
tub1	278	0.0343	2.20×10^{7}	(1.63×10^{-9})
H-anti	412	0.0223		1.17×10^{-9}
ole	425	0.0165		0.87×10^{-9}
4 genes combined	1585			Ave. 1.08×10^{-9} (4.32×10^{-9})***

*We used the nucleotide diversity (π), which is the average pairwise distance between isolates (Li 1997):

$$\pi = 2/n(n-1) \times \sum_{ij} \pi_{ij}$$

where n is the number of isolates and π_{ij} is the number of nucleotide differences per base between *t*he ith and *j*th isolates and $n(n-1)/2$ is the number of possible pairwise comparisons.

**Only DNA substitution rates for arf and tub1 loci have been estimated (Kasuga *et al*. 2002). The shown DNA substitution rates are based on one divergence time of *Histoplasma capsulatum* and *Blastomyces dermatitidis*, which is 127.8 Ma. Rates for H-anti and ole were estimated from the average value of π/μ at arf and tub1 loci together with π values at H-anti and ole. Average substitution rates were calculated by summing up the products of length and π at each locus and then divided by the total length of the concatenated genes (1585 bp).

***DNA substitution rates based on two *Histoplasma-Blastomyces* divergence times of 127.8 and 31.8 Ma are shown, respectively.

well as sexually. Suspected clonal isolates are as follows: 10 of 13 isolates of *Hc* var. *farciminosum* (H90, H95, H96, H148, H174, H175, H190, H191, H193 and H194) and the two Peruvian DNA samples (H140 and H185) had identical multilocus genotype in each of the groups. These isolates were obtained from animals kept in crowded maintenance facilities and probably represent clonal forms of *H. capsulatum* which spread from host to host. Two isolates of *Hc* var. *duboisii* (H87 and H91) and three Panamanian isolates (H82, H83 and probably H81) were isolated from single patients and showed identical multilocus genotypes in each of the groups. Therefore, they are very likely to be clones. These duplicated clonal isolates were excluded from the data set. There are several other cases where individuals share an identical multilocus genotype, e.g. H5, 6, 139, 179 and 181 and EH332, 374, 379, 383, 391, 394 and 408 (see Fig. 4). The probability of sampling a particular genotype more than once in the data set can be calculated using a binomial expression using allele frequencies assuming that (i) different genotypes arise by recombination and not mutation; (ii) mating is random and (iii) loci are at linkage equilibrium (Fisher *et al*. 2000). The probabilities of observing the H5 genotype five times or more and the EH332 genotype seven times or more are 0.55 and 0.18, respectively. These fungal isolates were obtained from different locations or from different noncaptive host individuals. These isolates were left in the data set because the evidence for their clonality was weak. Table 3 shows values for π at each of the four loci. Under a molecular clock and assuming no intralocus recombination, the nucleotide diversity π for each locus would be approximately proportional to the DNA substitution rate μ at the locus; μ for H-anti and ole were calculated from the population diversities estimated for these loci and the averages of π/μ for arf and tub1 (Table 3).

We calculated the average genetic distance for all four loci among eight lineages (NAm 1, NAm 2, LAm A, LAm B, Australia + Netherlands (Indonesia?), Africa, H81 and H153) to be 2.80%. This value and the average substitution rate for the four loci (1.08×10^{-9}–4.32×10^{-9}) placed the radiation of *Histoplasma* at approximately 3.2–13.0 Ma, mirroring the range of the *Histoplasma* and *Blastomyces* divergence values, 32–128 Ma (Fig. 5). By either estimate, the radiation of *Histoplasma* is one-tenth as old as the divergence of *Histoplasma* and *Blastomyces*.

Comparing population diversities

The extent of population diversity varies among populations. For example, isolates in each of the NAm 1, NAm 2, LAm B and Australian clades were homogeneous and coalesced during the late Pliocene to Pleistocene epochs (Figs 3 and 5). On the other hand, genetic diversities in the LAm A and African clades seem much larger and coalesced in the Miocene to Pliocene epochs. Apparent differences in genetic diversity might be attributable to sample sizes as we have the largest collection of LAm A isolates ($n = 55$). To account for sample size difference, we resampled five isolates from 55 LAm A isolates randomly with replacement and calculated nucleotide diversity of the five isolates using the nucleotide diversity π; this procedure was repeated 10 000 times. The resampled distance distribution was compared with diversities of several populations (Fig. 6). The statistical resampling demonstrates that the observed small population diversities of LAm B and NAm 2 were not due to sampling error ($P < 0.002$) whereas the population diversity of the African clade appears to be comparable to the LAm A clade. The LAm A clade contains isolates from Mexico, Guatemala, Panama, Colombia, Surinam and Brazil. Genetic diversities observed in subpopulations of the LAm A from Brazil and Colombia themselves had diversities as large as the entire LAm A clade. The genetic diversity found in the Mexican (including Guatemala) population is significantly smaller than in the Brazilian and Colombian populations but still larger than in LAm B and NAm 2 clades (Fig. 6). Thus, the size of the endemic area or sampling area does not correlate with the genetic diversity of each breeding population.

Inference of population history

The large numbers of isolates belonging to NAm 2 and LAm A clades enabled us to analyse their population history. Templeton's NCA is suited for this analysis because this method does not require presumption about the underlying population process; instead, a historical reconstruction is derived using inference key (Templeton 1998; Knowles & Maddison 2002). The NCA can detect geographical association and discriminate between phylogeographical associations due to recurrent but restricted gene flow vs. historical events operating at the population level such as past fragmentation, colonization or range expansion events (Templeton 1998). A haplotype genealogy was estimated for each population at each of the four loci by statistical parsimony (Templeton *et al*. 1992; Clement *et al*. 2000). The geographical location of each of the haplotypes and the haplotype genealogy were then used to detect geographical associations and infer population history. Table 4 summarizes the outcome of the NCA. Among the LAm A isolates, geographical differentiation caused by restricted gene flow was detected in three of the four loci. When Eurasian isolates were included in the LAm A, a long-distance colonization event from Latin America to Eurasia and past fragmentation were detected for arf and tub1 loci, respectively. On the other hand, geographical differentiation was not detected in the NAm 2 for any of the four loci, unlike the differentiation between Alabama and Indiana isolates found using single nucleotide polymorphism (SNP) and microsatellite markers (Carter *et al*. 2001).

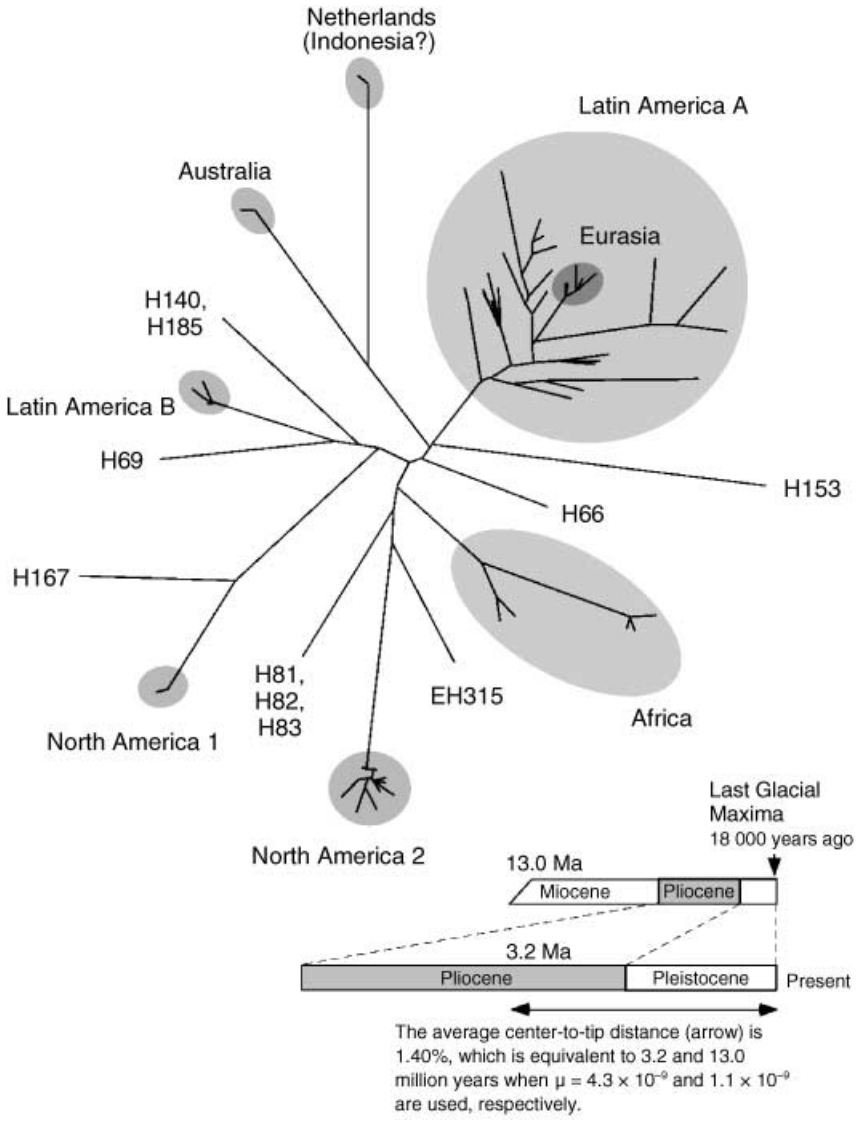

Fig. 5 An unrooted representation of the neighbour-joining tree showing the star phylogeny. The average genetic distance among distinct lineages [NAm 1, NAm 2, LAm A, LAm B, Australia + Netherlands (Indonesia?) group, Africa, H81 and H153] was 2.8%. Half of this value corresponds to the time point of the radiation, which is 3.2 and 13.0 million years ago (Ma) when the DNA substitution rates of 4.3×10^{-9} and 1.1×10^{-9} are used, respectively.

Discussion

Radiation of Histoplasma capsulatum

The clues that we can use to try to explain the present distribution of *Histoplasma* species concern the phylogeographical pattern of genetic variation and the rate of molecular evolution in this fungus. There are seven clades, leaving aside for the moment the Eurasian clade. There are six lone lineages, all from Latin America. There is almost no resolution of the relationships among the clades and lineages, the Australian and Netherlands (Indonesia?) clades being the sole exception. Two of the clades, LAm A and Africa, harbour diverse genotypes, the rest have relatively little genetic variation. It seems reasonable to assume that *Histoplasma* experienced a radiation 3.2–13 Ma (Pliocene to Miocene) in Latin America when the global climate was warmer than at present (Chandler 1999; Brining *et al*. 2002) and that the clades in Africa, Australasia and North America are the result of dispersal. The advent of the Pleistocene, 1.8 Ma, brought a period of intense cold, subjecting much of the Earth's temperate zones to repeated glaciations. Current patterns of temperate flora and fauna have been attributed to the effects of these glaciations and the resulting biotic refugia (Willis & Whitaker 2000). Modern *Histoplasma* populations are endemic in temperate forests and tropical rainforests (Furcolow 1958; Fonseca 1971). The low genetic variation found in modern temperate populations and the high variation found in tropical regions may be explained by such glacial refugia. Equatorial populations would not suffer the migration and genetic loss associated with refugia but temperate populations would lose genetic variation during the range reduction as refugia form. In temperate areas, relatively few genotypes

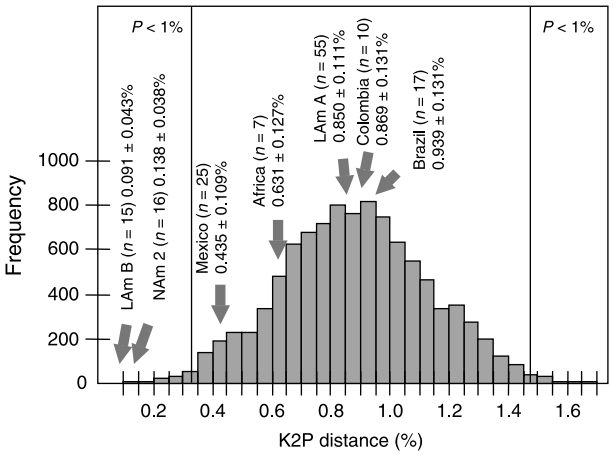

Fig. 6 A distribution of resampled LAm A population diversity. Five isolates were resampled 10 000 times with replacement from a population of 55 Latin American group A isolates and mean pairwise distance for each resample was calculated. Arrows point to population diversities calculated from the original data; four of the geographical populations as well as subpopulations of LAm A, i.e. from Mexico, Colombia and Brazil, are shown.

would then be available for recolonization as the frozen habitat thawed. Geological events in the New World are consistent with this scenario. At the last glacial maximum (LGM), 18 000 years ago, abundant tropical rainforest remained in Central America and the Amazon Basin (Colinvaux *et al*. 1996; Willis & Whitaker 2000), the present site of the most diverse clade, LAm A. In this clade, there is no evidence that alleles in both the arf and tub1 loci coalesce any more recently than the *Histoplasma* radiation (Fig. 3a and d). Conversely, in temperate Latin America, at the LGM, e.g. Argentina where LAm B is pervasive, deserts and semideserts predominated. Today, this region hosts LAm B, in which alleles do coalesce well after the *Histoplasma* radiation.

Similarly, at the LGM in North America the present endemic area was covered with a thick ice sheet and taiga and temperate forests were restricted to the southernmost part of the present endemic area (Adams 1997). Again, in NAm 1 and NAm 2, alleles coalesce well after the *Histoplasma* radiation. Note that the large genetic distance between the NAm 1, NAm 2, LAm B clades and other clades shows that they were part of the original radiation of *Histoplasma* species and thus do not represent migrations from the diverse LAm A clade following the LGM.

At the LGM, modern Australian *Histoplasma* sites were all semideserts or arid scrub and the nearly identical genotypes of the five isolates could be explained by a severe population bottleneck at that time.

African *Histoplasma* was part of the original Pliocene radiation of species and is genetically diverse and, therefore, the diversity was maintained through the LGM. Although much of Africa was arid and cooler at the LGM, tropical rainforest surrounded by savanna persisted in Central Africa (Adams 1997), probably providing the habitat in which *Histoplasma* survived.

Eurasia harbours the clade that is the most difficult to understand. The Eurasian clade arises from within LAm

Table 4 Summary of the nested clade analysis of LAm A, Eurasian, and NAm 2 clades

Loci	Statistically significant historical events
	Within LAm A*
arf	Restricted gene flow with isolation by distance
H-anti	Inconclusive outcome
ole	Restricted gene flow with isolation by distance
tub1	Restricted gene flow with isolation by distance
	Northward contiguous range expansion to Mexico
	Between LAm A and Eurasia
arf	Long distance colonization event from Colombia or Mexico to Eurasia
H-anti	Inconclusive outcome
ole	Restricted gene flow with isolation by distance
tub1	Past fragmentation
	Within NAm 2**
arf	No population-level historical event detected
H-anti	No population-level historical event detected
ole	No population-level historical event detected
tub1	No population-level historical event detected

*LAm A isolates were grouped into four geographical populations; *Mexico*, *Colombia*, *São Paulo* and *Rio De Janeiro*.
**NAm 2 isolates were grouped into three geographical populations; *Midwest*, *South* and *South East* (see Materials and methods).

A and the genotypes of the individuals in the clade are very homogeneous, notwithstanding their having been collected from the Far East to Europe. Nested clade analysis of the arf locus suggested a long distance colonization event from Latin America to Eurasia and NCA of the tub1 locus suggested a past fragmentation event, both at unspecified times. The Eurasian clade originated between 1.7 and 6.8 Ma, based on the estimated percentage of DNA substitutions per nucleotide per 1 million years of 0.16–0.65% at tub1 locus and the maximum pairwise distance among isolates in the Eurasian clade of 1.1% at tub1 locus (Eurasian isolates are monomorphic at arf). This estimate provides an upper limit for the immigration of LAm A individuals to Eurasia but the event could have been more recent if several individuals with different genotypes were involved in the initial dispersal. For example, one cargo of domesticated horses or donkeys infected with multiple individuals of *H. capsulatum*, transported as recently as 500 years ago, could have initiated the Eurasian clade.

Histoplasma capsulatum var. *farciminosum* individuals were found in three clades, African (H189), NAm 2 (H173) and Eurasian, the latter of which accommodated 11 of the 13 individuals. It is clear that *Hc* var. *farciminosum* is not a monophyletic group and that individuals have acquired the ability to cause superficial disease in horses and other equidae more than once. Therefore, *Hc* var. *farciminosum* is not a valid taxon, it is a disease. The 10 of 11 individuals of *Hc* var. *farciminosum* from Eurasia had identical alleles at all four loci, indicating that they represent one clone, ranging from Poland to Egypt to India.

Do lone lineages represent cryptic species?

Seven evolutionary lineages were represented by single individuals or single genotypes that did not belong to any of the seven phylogenetic species. Can these lone lineages be considered as cryptic species? Our sampling favoured human clinical isolates, and the lone lineages were biased against this trait, so the lineages may represent larger populations of fungi in nature. For example, EH315 was recovered from a wild bat and the two Peruvian individuals with identical genotypes (H140 and H185) were recovered from owl monkeys. The Peruvian individuals did not form mycelium in the laboratory and could not be cultivated (Miller & Owens 1999), forcing us to use DNA from infected liver and spleen for our PCR amplification. The Peruvian individuals had larger yeast cells than typical *H. capsulatum* and reminded mycologists of another fungal pathogen, *Lacazia loboi*, but clearly belong in the genus *Histoplasma*. Isolate H153 was also phenotypically distinct, having unusually large macroconidia, not converting to yeast at 37 °C and causing an atypical, disseminated cutaneous histoplasmosis (Lacaz *et al*. 1999). Other lone lineages, e.g. H66, H69 from Colombia, H167 from Argentina and H81 from Panama, showed no phenotypic differences, either in morphology or ecology. On balance, it seems likely that some of our lone lineages represent fungi that are not likely to be collected by clinicians, might not be recognized as *Histoplasma* or cannot be cultivated by methods routinely used in clinical laboratories. If this thinking is correct, the lone lineages may represent natural populations that we have not sampled adequately. The dramatic increase in recovery of NAm 1 individuals correlated with the AIDS pandemic provides support for this idea. Prior to the pandemic, which began in the early 1980s, NAm 2 was predominant and NAm 1 was represented by only two individuals, H9 [Downs, obtained in 1968 from an 86-year-old woman (Gass & Kobayashi 1969)] and H79 [obtained from a striped skunk in the 1940s (Emmons *et al*. 1949)]. As AIDS spread, NAm 1 became common in clinics (Spitzer *et al*. 1990), not due to an increase in NAm 1 in nature but to an increase in susceptible hosts.

Conclusion

Histoplasma capsulatum comprises at least seven phylogenetic species, one in each of Africa, Australia and the Netherlands (Indonesia?) and two species each in North America and Latin America. The Eurasian population originated from within one of the Latin American species. In addition, seven distinct lineages represented by single isolates or genotypes were identified in Latin America. Each of these lineages potentially represents an independent phylogenetic species. Judging from the observed genetic diversity and DNA substitution rates, the radiation of *Histoplasma* started between 3 and 13 Ma in Latin America. The present day population structure of *Histoplasma* can be explained by refugial populations in the last glacial maxima.

Acknowledgements

We thank E. Keath, G. Kobayashi, L. Wheat, P. Connolly, S. Moser, B. Hines, W. Dismukes and D. Muir for supplying isolates, DNA and associated clinical information; M. Sugiyama and E. Jensen for sharing unpublished results and Rachel Whitaker and Jeremy Dettman for useful comments on the manuscript. Financial support for this work was provided by the National Institutes of Health (grants HL55953 and AI37232 to J.W.T.).

References

Adams J (1997) Global land environments since the last interglacial. Oakridge National Laboratory, TN, USA. http://www.esd.ornl.gov/ern/qen/nerc.html.

Avise JC, Ball RM Jr (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology*, **7**, 45–67.

Bauder B, Kuebber-Heiss A, Steineck T, Kuttin ES, Kaufman L (2000) Granulomatous skin lesions due to histoplasmosis in a badger (*Meles meles*) in Austria. *Medical Mycology*, **38**, 249–253.

Baum DA, Shaw KL (1995) Genealogical perspectives on the species problem. In: *Experimental and Molecular Approaches to Plant Biosystematics* (eds Hoch PC, Stephenson AG.), pp. 289–303. Missouri Botanical Garden, St Louis, MO.

Berliner MD (1968) Primary subcultures of *Histoplasma capsulatum*, I. Macro and micro-morphology of the mycelial phase. *Sabouraudia*, **6**, 111–118.

Brining L, Chan V, Choi E, De Sosa M, Lee C (2002) *The Miocene Epoch*. Museum of Paleontology, University of California, Berkeley, CA, USA. http://www.ucmp.berkeley.edu/tertiary/mio.html.

Carr J, Shearer GJ (1998) Genome size, complexity, and ploidy of the pathogenic fungus *Histoplasma capsulatum*. *Journal of Bacteriology*, **180**, 6697–6703.

Carter DA, Taylor JW, Dechairo B *et al*. (2001) Amplified single-nucleotide polymorphisms and a $(GA)_n$ microsatellite marker reveal genetic differentiation between populations of *Histoplasma capsulatum* from the Americans. *Fungal Genetics and Biology*, **34**, 37–48.

Chandler MA (1999) *The Climate of the Pliocene: Simulating Earth's Last Great Warm Period*. Goddard Institute for Space Studies, Columbia University, NY, NY, USA. http://www.giss.nasa.gov/research/paleo/pliocene/index.html.

Clement M, Derington J, Posada D (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Colinvaux PA, De Oliveira PE, Moreno JE, Miller MC, Bush MB (1996) A long pollen record from lowland Amazonia: Forest and cooling in glacial times. *Science*, **275**, 85–88.

Crandall KA (1996) Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type sequences. *Molecular Biology and Evolution*, **13**, 115–131.

Deepe GS Jr, Durose GG (1995) Immunobiological activity of recombinant H antigen from *Histoplasma capsulatum*. *Infection and Immunity*, **63**, 3151–3157.

Dettman J, Jacobson D, Taylor JW (2003a) A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. *Evolution*, in press.

Dettman J, Jacobson D, Turner E, Pringle A, Taylor JW (2003b) Reproductive isolation and phylogenetic divergence in Neurospora: comparing methods of species recognition in model eukaryote. *Evolution*, in press.

Dress A, Huson D, Moulton V (1996) Analyzing and visualizing sequence and distance data using SPLITSTREE. *Discrete Applied Mathematics*, **71**, 95–109.

Emmons CW, Morlan HB, Hill EL (1949) Histoplasmosis in rats and skunks in Georgia. *Public Health Report*, **64**, 1423–1430.

Felsenstein J (1985) Confidence limits on phylogenies: an approach to using the bootstrap. *Evolution*, **39**, 783–791.

Fisher MC, Koenig GL, White TJ, Taylor JW (2000) Pathogenic clones versus environmentally driven population increase: analysis of an epidemic of the human fungal pathogen *Coccidioides immitis*. *Journal of Clinical Microbiology*, **38**, 807–813.

Fonseca JC (1971) Analisis estadistico y ecologio-epidemiologico de la sensibilidad a la histoplasmina en Colombia, 1950–1968. *Antioquia Medica*, **21**, 109–154.

Furcolow ML (1958) Recent studies on the epidemiology of histoplasmosis. *Annals of the New York Academy of Science*, **72**, 127–164.

Gargano S, Di Lallo G, Kobayashi GS, Maresca B (1995) A temperature-sensitive strain of *Histoplasma capsulatum* has an altered delta 9-fatty acid desaturase gene. *Lipids*, **30**, 899–906.

Gass M, Kobayashi GS (1969) Histoplasmosis: an illustrative case with unusual vaginal and joint involvement. *Archives of Dermatology*, **100**, 724–727.

Gugnani HC, Muotoe-Okafor FA, Kaufman L, Dupont B (1994) A natural focus of *Histoplasma capsulatum* var. *duboisii* is a bat cave. *Mycopathologia*, **127**, 151–157.

Harden TJ, Hunt PJ (1985) Histoplasmosis and Australian cave environments. *Helictite*, **23**, 23–26.

Harris GS, Keath EJ, Medoff J (1989) Characterization of alpha and beta tubulin genes in the dimorphic fungus *Histoplasma capsulatum*. *Journal of General Microbiology*, **135**, 1817–1832.

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**, 182–192.

Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.

Isbister J, Elliott M, Nogrady S (1976) Histoplasmosis: an outbreak occurring among young men who visited one cave. *Medical Journal of Australia*, **2**, 243–248.

Kasuga T, Taylor JW, White TJ (1999) Phylogenetic relationships of varieties and geographical groups of the human pathogenic fungus *Histoplasma capsulatum* Darling. *Journal of Clinical Microbiology*, **37**, 653–663.

Kasuga T, White TJ, Taylor JW (2002) Estimation of nucleotide substitution rates in Eurotiomycete fungi. *Molecular Biology and Evolution*, **19**, 2318–2324.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623–2635.

Kwon-Chung KJ, Bennett JE (1992) *Medical Mycology*. Lea & Febiger, Malvern, PA.

Lacaz CdS, Del Negro GMB, Vidal MSM *et al*. (1999) Atypical disseminated cutaneous histoplasmosis in an immunocompetent child, caused by an 'aberrant' variant of *Histoplasma capsulatum* var. *capsulatum*. *Revista do Instituto de Medicina Tropical de São Paulo*, **41**, 195–202.

Li W-H (1997) *Molecular Evolution*. Sinauer Associates, Sunderland, MA.

Lodge JK, Johnson RL, Weinberg RA, Gordon JI (1994) Comparison of myristoyl-CoA: protein N-myristoyltransferases from three pathogenic fungi: *Cryptococcus neoformans*, *Histoplasma capsulatum*, and *Candida albicans*. *Journal of Biological Chemistry*, **269**, 2996–3009.

Manfredi R, Mazzoni A, Nanetti A, Chiodo F (1994) Histoplasmosis capsulati and duboisii in Europe: the impact of the HIV pandemic, travel and immigration. *European Journal of Epidemiology*, **10**, 675–681.

Mayden RL (1997) A hierarchy of species concepts: The denouement in the saga of the species problem. In: *Species: the Units of Biodiversity* (eds Claridge MF, Dawah HA, Wilson MR), pp. 381–424. Chapman & Hall, London.

de Medeiros Muniz M, Pizzini CV, Peralta JM, Reiss E, Zancope-Oliveira RM (2001) Genetic diversity of *Histoplasma capsulatum* strains isolated from soil, animals, and clinical specimens in Rio de Janeiro State, Brazil, by a PCR-based random amplified polymorphic DNA assay. *Journal of Clinical Microbiology*, **39**, 4487–4494.

Miller GF, Owens JW (1999) Ultrastructural characterization of the agent of systemic yeast infection of owl monkeys. *Medical Mycology*, **37**, 139–145.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Poonwan N, Imai T, Mekha N *et al.* (1998) Genetic analysis of *Histoplasma capsulatum* strains isolated from clinical specimens in Thailand by a PCR-based random amplified polymorphic DNA method. *Journal of Clinical Microbiology*, **36**, 3073–3076.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotype. *Molecular Ecology*, **9**, 487–488.

Rannala B, Yang Z (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution*, **43**, 304–311.

Reyes-Montes MR, Bobadilla-Del Valle M, Martínez-Rivera MA *et al.* (1999) Relatedness analyses of *Histoplasma capsulatum* isolates from Mexican patients with AIDS-associated histoplasmosis by using histoplasmin electrophoretic profiles and randomly amplified polymorphic DNA patterns. *Journal of Clinical Microbiology*, **37**, 1404–1408.

Rippon JW (1988) *Medical Mycology*, 3rd edn. W.B. Saunders, Philadelphia.

Rozas J, Rozas R (1999) DnaSP, Version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, **14**, 1218–1231.

Sebghati TS, Engle JT, Goldman WE (2000) Intracellular parasitism by *Histoplasma capsulatum*: fungal virulence and calcium dependence. *Science*, **290**, 1368–1372.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Spitzer ED, Keath EJ, Travis SJ *et al.* (1990) Temperature-sensitive variants of *Histoplasma capsulatum* isolated from patients with acquired immunodeficiency syndrome. *Journal of Infectious Diseases*, **162**, 258–261.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tamura M, Kasuga T, Watanabe K *et al.* (2002) Phylogenetic characterization of *Histoplasma capsulatum* strains based on ITS region sequences, including two new strains from Thai and Chinese patients in Japan. *Nippon Ishinkin Gakkai Zasshi*, **43**, 11–19.

Taylor JW, Jacobson JD, Kroken S *et al.* (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology*, **31**, 21–32.

Taylor ML, Chávez-Tapia CB, Vargas-Yañez R *et al.* (1999) Environmental conditions favoring bat infection with *Histoplasma capsulatum* in Mexican shelters. *American Journal of Tropical Medicine and Hygiene*, **61**, 914–919.

Taylor TN, Hass H, Kerp H (1999) The oldest fossil ascomycetes. *Nature*, **399**, 648.

Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplogypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Willis KJ, Whitaker RJ (2000) The refugial debate. *Science*, **287**, 1407–1407.

This work is a part of collaborations among 20 laboratories in nine countries to understand the biogeography of the fungal pathogen *H. capsulatum*.
