Letter to the Editor

Estimation of Nucleotide Substitution Rates in Eurotiomycete Fungi

Takao Kasuga,* Thomas J. White,† and John W. Taylor*

*Department of Plant and Microbial Biology, University of California–Berkeley; and †Celera Diagnostics, Alameda, California

In the entire fungal kingdom, only DNA substitution rates in the SSU rRNA gene (Berbee and Taylor 1993, 2001) and amino acid substitution rates (Heckman et al. 2001) have been estimated and used to date fungal divergences. However, these molecules are not sufficiently variable to date events at or below the genus level. DNA sequences of protein-coding genes and the internal transcribed spacer (ITS) region are sufficiently variable, but their substitution rates are not known. In this article, we investigated the DNA substitution rates in the protein-coding genes and the ITS region by comparison with the SSU rDNA divergence in a representative fungal lineage, Eurotiomycetes (=plectomycetes) (Eriksson and Winka 1998). The Eurotiomycete lineage is a monophyletic class of Ascomycota (Berbee and Taylor 1992) and includes many economically important fungi, such as Penicillium chrysogenum (antibiotic production) and Aspergillus oryzae (soy sauce production), as well as many human pathogens such as Coccidioides immitis (lung disease), Histoplasma capsulatum (lung disease), and Trichophyton rubrum (athlete's foot). Estimating DNA substitution rates in quickly evolving molecules such as the ITS and protein-coding genes will not only provide a means to estimate the divergence times between lineages at and below the genus level, but in conjunction with coalescent theory, it may provide information for estimating epidemiological parameters such as effective population size (Watterson 1975) and recombination rates (Hey and Wakeley 1997).

With the aid of published Eurotiomycete phylogenies based on *SSU rDNA* (Ogawa, Yoshimura, and Sugiyama 1997; Sugiyama, Ohara, and Mikawa 1999), we searched for pairs of closely related species which had not only *SSU rDNA* sequences but also sequences of homologous protein-coding genes. Next, we compared the extent of differentiation at synonymous sites of protein-coding genes between sister taxa. It was found that saturation at synonymous sites of protein-coding loci is attained before divergence of the *SSU rDNA* sequences reaches 1%–2%. A 1%–2% divergence usually corresponds to divergences seen within a genus or among closely related genera. Therefore, the protein genes should be well suited to dating divergences at and below the genus level.

Mol. Biol. Evol. 19(12):2318–2324. 2002

© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

In this paragraph we briefly outline our methodology. The first step of our approach was to estimate the divergence times between sister species from the SSU rDNA phylogenetic tree (Ochman and Wilson 1987; Ochman, Elwyn, and Moran 1999). Twenty-seven Eurotiomycete species and two out-group taxa, i.e., the Sordariomycete Neurospora crassa and the loculoascomycete Capronia pilosella, were subjected to distance analysis. The regularity of nucleotide substitution rates at the SSU rDNA gene was evaluated by the two-cluster test and the branch length test (Takezaki, Rzhetsky, and Nei 1995; Nei and Kumar 2000). Rates were not found to be constant for six species and four nodes using the branch length test and the two-cluster test, respectively (P < 0.05, data not shown). Apparently, rate constancy does not hold throughout this system. Because rates may not be constant, we also considered a model in which molecular evolutionary rates vary across lineages instead of remaining constant. We compared the Langley Fitch (LF) algorithm, which assumes rate constancy, with the nonparametric rate-smoothing (NPRS) algorithm, which does not (Langley and Fitch 1974; Sanderson 1997). All methods of phylogenetic inference depend heavily on their underlying models. For this reason we used a hierarchical likelihood ratio test to search for the DNA substitution model that best fit our SSU rDNA data set (Posada and Crandall 1998). Phylogenetic relationships were then inferred using the selected evolutionary model and the heuristic search option for maximum likelihood (ML) implemented in PAUP 4.0b8a (Swofford 2001). The ML tree obtained was then used to estimate the divergence times between closely related taxa using the NPRS and LF methods. A simple and widely used Kimura two-parameter model was also used to construct a neighbor-joining (NJ) tree; the NJ tree was also subjected to the NPRS and LF methods as a comparison. Both NPRS and LF methods require at least one calibration point to fit branch lengths to the geological time scale. We used the divergence of Eurotiomycetes (=plectomycetes) and Sordariomycetes (=pyrenomycetes) for calibration. Initially, the divergence was set to 100 arbitrary units, and the relative divergence times between taxa were estimated (fig. 1; details are available in Supplementary Material at http:// www.molbiolevol.org/). For all pairs of taxa, the NPRS algorithm gave 15.2% to 87.5% larger values for divergence times than did the LF algorithm. For both NPRS and LF algorithms, most of the divergence values calculated from the NJ tree with a Kimura two-parameter algorithm were larger than those from the ML tree; NJ values were up to 48.6% larger than ML values. A simulation study indicated that divergence times are estimated more accurately using NPRS rather than LF when (1) there are enough data, (2) evolution is not clock-

Key words: synonymous substitution rates, molecular clock, plectomycete, fungi, evolution, *ITS*.

Address for correspondence and reprints: Takao Kasuga, Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, California 94720. E-mail: kasugat@uclink4.berkeley.edu.



FIG. 1.—Estimation of divergence times according to the NPRS (a) and LF (b) methods. Branch lengths are proportional to time. Either of the three published values (310, 400, or 670 Myr) for the ES divergence time can be used to calibrate the trees. The tree topology was obtained through the heuristic search for ML implemented in PAUP 4.0b8a (Swofford 2001) using the Tamura Nei DNA substitution model with equal base frequencies, among-site rate variation Gamma (0.6431), and the proportion of sites unable to accept substitutions (0.6425). Aligned sequences have been deposited in TreeBase (study accession number S804, www.treebase.org/treebase/).

like, and (3) levels of rate autocorrelation are moderate to high (Sanderson 1997). Unfortunately, although we believe we have sufficient data (1,631 bp), we have no way of assessing the other parameters, so we have used the divergence times of both the NPRS and the LF methods, hoping to account for the ambiguity of the past.

As mentioned above, the NPRS and LF algorithms give only relative divergence times, which may be understood as a percentage of the time since the divergence of Eurotiomycetes and Sordariomycetes (ES). This ES divergence time can be inferred from the fossil record or estimated by molecular clock approaches, both of which again require calibration points from fossils. The oldest well-documented ascomycete fossils are the perithecia, asci, and ascospores that resemble those of extant Sordariomycetes and are found in the 400-Myr-old Rhynie chert (Taylor, Hass, and Kerp 1999). Thus, 400 Myr for the ES divergence would seem to provide a conservative date; however, earlier dates can be expected. The ES divergence time estimated from a molecular clock approach largely depends on a fundamental calibration point: the fungi-animal or fungi-animal-plant divergence time. Berbee and Taylor (Berbee and Taylor 2001) calibrated an SSU rDNA phylogenetic tree using a fungianimal divergence time of 965 Myr (Doolittle et al. 1996) and estimated the ES divergence to be 310 Myr. Heckman et al. (2001) used a fungi-animal-plant divergence time of 1,576 Myr, and their protein diversity analyses yielded an ES divergence of 670 Myr. The 310-Myr date is probably too recent, judging from the fossil evidence, and the 670-Myr date provides a more ancient estimate. We used three values, 310, 400, and 670 Myr,

as calibration points for estimation of the divergence times between closely related Eurotiomycete species.

Our goal was to estimate DNA substitution rates at six independent protein-coding loci and the ITS of the rDNA repeat unit. To obtain substitution rates, we required estimates for divergence times and genetic distances between sister species. Divergence times between sister species of Eurotiomycete fungi were estimated using NPRS and LF algorithms based on three calibration times corresponding to the three estimates of ES divergence; hence, a total of six rate estimates were calculated for each of the species pairs (table 1). For class 1 chitin synthase (CHS1), class 2 chitin synthase (CHS2), zinc finger protein (creA), and orotidine-5'-phosphate decarboxylase (pyrG), four different measures of genetic distance were calculated: DNA substitutions in synonymous sites (K_S) , nonsynonymous sites (K_A) , the third bases of codons (K_{3rd}) , and exons. Introns were not included in these data sets. The modified Nei-Gojobori method implemented in MEGA2.0 (Kumar et al. 2000) was used to estimate K_S and K_A . The Kimura two-parameter model was used for all the other distance measures. Large parts of the data sets for ADP-ribosylation factor (arf) and alpha tubulin (tub1) are intron sequences, so for these loci four measures of genetic distance were calculated: DNA substitutions in exons, introns, K_{3rd} , and K_s . There were no nonsynonymous substitutions in the arf or tub1 loci.

The most data were available for *CHS1*, which was found in 4 out of 10 pairs of sister species. We used these data to evaluate how estimates of nucleotide substitution varied in different regions of the Eurotiomycete

$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$						RAI	TE ESTIMATES A	THE GIVEN E	S CALIBRATIO	IN TIME POINT	SI		
							(10	⁻⁹ substitutions	site per year)		2		
Special plane Acto Direct Along GO Myr Restance Program Margin control Margin control Margin control 244 Myr 244 Myr 313 Myr 333 Myr Berner at L(1992) Program Margin control Margin control 100 123 Myr 244 Myr 244 Myr 313 Myr 333 Myr Program Margin control Margin control 100 Myr 100 Myr 240 Myr 248 Myr 313 Myr 348 Myr Program Margin control Margin control 100 Myr 100 Myr </th <th>Locus and</th> <th>Accesson No.</th> <th></th> <th></th> <th>DNA Dis-</th> <th></th> <th>NPRS</th> <th></th> <th></th> <th>LF</th> <th></th> <th></th>	Locus and	Accesson No.			DNA Dis-		NPRS			LF			
Given the relation. Given the relation. 91.1 Mp 91.1 Mp 91.3 Mp <th< th=""><th>SPECIES PAIR</th><th>AND LOCUS</th><th>SITE</th><th>L (bp)</th><th>TANCE^b</th><th>310 Myr</th><th>400 Myr</th><th>670 Myr</th><th>310 Myr</th><th>400 Myr</th><th>670 Myr</th><th>REFERENCE</th></th<>	SPECIES PAIR	AND LOCUS	SITE	L (bp)	TANCE ^b	310 Myr	400 Myr	670 Myr	310 Myr	400 Myr	670 Myr	REFERENCE	
Histophotone coportinue, Basonyces channels, Basonyces channels, Basonyces familie, Basonyces familie,	CHSI Divergence time estimates					99.1 Mvr	127.8 Mvr	214.1 Mvr	24.6 Mvr	31.8 Mvr	53.3 Mvr		
Interviewer derentinds. MS242 Suspringens Energene time suspringenses MS242 Suspringenses Suspringenses Suppringenses Suppringense Suppringenses Suppringenses Suppringense	Histoplasma capsulatum,	M82947,	I		1								
Interfactor	Blastomyces dermatitidis	M82942	Exons	561 197	0.15	0.75	0.58	0.35	3.00	2.33	1.39	Bowen et al. (1992)	
Decomposition in the number of the			Inird bases Synonymous	18/ 153	7C.U	20.2 2 49	2.04 2.70	1.22	10.01	8.19 10.86	4.89 6.49	NINO-Vega et al.	
			Nonsynonymous	408	0.02	0.09	0.07	0.04	0.34	0.27	0.16	(0007)	
$ \begin{array}{cccc} Trippoprint Inform. Trippopri Trippoprint Inform. Tripp$	Divergence time estimates	A D/07507				26.8 Myr	34.6 Myr	58.0 Myr	11.5 Myr	14.8 Myr	24.8 Myr		
$ \begin{array}{ccccc} \mbox{Trunc} $	Armvoaerma mcurvatum, Trichophyton ruhrum	AB005362, AB005793	Exons	561	0.12	2,31	1.79	1 07	5 41	4.19	2.50		
$ \begin{array}{ccccc} \label{eq:constraints} \\ \mbox{functuations} & 123 & 0.4 & 0.14 & 708 & 4.23 & 21.6 & 165.5 & 9.88 \\ \mbox{functuations} & 1.28071, \\ \mbox{functuations} & 1.280, \\ \mbox{functuations} & 1.21, \\ functuat$			Third bases	187	0.39	7.23	5.61	3.35	16.91	13.11	7.83		
$ \begin{array}{c cccc} \mbox{Divergance time estimates} & May any monome estimates May any many and monome estimates 128071. Evens 564 My 644 My 160 My 20.60 My 345 My 345 My My finamente estimates 128071. Evens 561 0.61 1.82 0.43 3.73 0.23 1.32 0.65 My 20.60 My$			Synonymous	153	0.49	9.14	7.08	4.23	21.36	16.55	9.88		
			Nonsynonymous	408	c0.0	0.47	00.0	0.22	1.09	0.84	00.0		
$ \begin{array}{ccccc} \mathematical matrix and the set of the set$	Divergence time estimates Mathematical Journation	1 78071				43.7 Myr	56.4 Myr	94.4 Myr	16.0 Myr	20.6 Myr	34.5 Myr		
$ \begin{array}{cccc} \mbox{Final base} & Final bas$	M. filamentosa	L28072	Exons	561	0.61	1.82	1.41	0.84	4.98	3.86	2.30		
$ \begin{array}{cccc} \mbox{Pixements} & \mbox{Final constraints} & Final constr$			Third bases	187	0.42	4.84	3.75	2.24	13.25	10.27	6.13		
$ \begin{array}{cccc} \mbox{Divergence time estimates} \\ Divergence time$			Synonymous Nonsvnonvmous	153 408	0.62 0.04	$7.14 \\ 0.48$	5.54 0.37	3.30 0.22	19.54 1.32	15.15 1.02	9.04 0.61		
$ \begin{array}{cccc} \mbox{Divergence time estimates}\\ \mbox{Divergence time estimates}\\ \mbox{L28067}\\ \mbox{Unincarpus result}\\ \mbox{L28067}\\ \mbox{Tubit bases}\\ $											14.1 Myr		
$ \begin{array}{cccc} \mbox{Containants}, Containan$	Divergence time estimates	1 78067				52.0 Myr	67.1 Myr	112.4 Myr	6.5 Myr	8.4 Myr			
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	Uncinocarpus reesti	L28069	Exons	561	0.23	2.20	1.71	1.02	17.59	13.63	8.14		
$ \begin{array}{c} CH32 \\ CH32 \\ CH32 \\ Divergence time estimates \ldots \\ Wasynonymous \\ B. dermatifidis \ldots \\ Wasynonymous \\ Mag949, \\ B. dermatifidis \ldots \\ M82949, \\ B. dermatifidis \ldots \\ M82949, \\ Wasynonymous \\ Mag9419, \\ Wasynonymous \\ Mag941, \\ Wasynonyma \\ Mag941, \\ Mag941, \\ Wasynonymous \\ Ma$			Third bases	187	0.88	8.43	6.53	3.90	67.28	62.14	31.13		
CH32 CH32 CH32 CH3 CH3 <th colspa<="" td=""><td></td><td></td><td>Synonymous</td><td>153 408</td><td>1.20</td><td>11.53 0.49</td><td>8.94 0.38</td><td>5.34 0.23</td><td>92.09 3 97</td><td>71.37 3.04</td><td>42.61 1 81</td><td></td></th>	<td></td> <td></td> <td>Synonymous</td> <td>153 408</td> <td>1.20</td> <td>11.53 0.49</td> <td>8.94 0.38</td> <td>5.34 0.23</td> <td>92.09 3 97</td> <td>71.37 3.04</td> <td>42.61 1 81</td> <td></td>			Synonymous	153 408	1.20	11.53 0.49	8.94 0.38	5.34 0.23	92.09 3 97	71.37 3.04	42.61 1 81	
$ \begin{array}{ccccc} \mbox{Divergence time estimates} \\ H. copsulatum, M82949, \\ H. copsulatum, M82949, \\ H. copsulatum, M82949, \\ B. dermatifidis \\ M82949, \\ B. dermatifidis \\ M82949, \\ Third bases \\ Synonymous \\ Synonymous \\ 154.7 \\ 0.57 \\ 0.27 \\ 0.57 \\ 0.57 \\ 0.57 \\ 0.57 \\ 0.28 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.12 \\ 0.12 \\ 0.12 \\ 0.03 \\ 0.12 \\ 0.12 \\ 0.12 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.11 \\ 0.03 \\ 0.12 \\ 0.12 \\ 0.12 \\ 0.12 \\ 0.12 \\ 0.12 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.11 \\ 0.08 \\ 0.05 \\ 0.11 \\ 0.01 \\ $	CHS2			P	0.0	1	0000	1	1	5			
H. capsulatum, t. capsulatum,M82949, m82949,Exons Third bases5670.140.690.530.322.762.141.28Bowen et al. (1992)B. dermatifidisM82943Exons154.70.572.860.151.049.037.004.18Nino-Vega et al.Synonymous154.70.572.862.121.3211.508.925.32(2000)Synonymous154.70.572.860.110.080.050.430.330.20(2000)Vorega et al.Synonymous412.30.020.110.080.050.430.330.20(2000)LosterAAA73.6 Myr73.6 Myr73.6 Myr73.6 Myr73.0 Myr20.2 MyrDrystale et al.Livergence time estimatesL03811,AA73.6 Myr73.6 Myr73.6 Myr71.44.26(1993)Aspergillus niger,L03811,A1.57 Myr73.6 Myr73.6 Myr71.44.26(1993)Aspergillus niger,L03811,A1.57 Myr73.6 Myr73.6 Myr73.6 Myr20.2 MyrDrystale et al.Aspergillus niger,L03811,A0.294.253.201.9622.17Drystale et al.Aspergillus niger,Nonsynonus358.61.4721.6216.7610.0046.9236.3621.71NonsynonusSynonymous904.40.081.180.900.542.52<	Divergence time estimates					99.1 Myr	127.8 Myr	214.1 Myr	24.6 Myr	31.8 Myr	53.3 Myr		
D. dermands Mo.2945 Exons Trind bases 500 1847 0.14 0.57 $0.022.25 1.741.74 1.041.04 9.030.33 7.00 4.181.8 Nino-Vega et al. Synonymous 154.7 0.57 2.86 2.22 1.74 9.03 7.00 4.18 Nino-Vega et al. Synonymous 154.7 0.57 2.86 2.22 1.74 1.04 9.03 7.00 4.18 Nino-Vega et al. Synonymous 412.3 0.02 0.11 0.03 0.20$	H. capsulalum,	M82949,		101	110	02.0	0 67	, , ,		- - -	96.1	(1001) Is to see the	
$ \begin{array}{c} \label{eq:constraint} \mbox{Tonymous} & 154.7 & 0.57 & 2.86 & 2.22 & 1.32 & 11.50 & 8.92 & 5.32 & (2000) \\ \mbox{Nonsynonymous} & 412.3 & 0.02 & 0.11 & 0.08 & 0.05 & 0.43 & 0.33 & 0.20 & (2000) \\ \mbox{Nonsynonymous} & 412.3 & 0.02 & 0.11 & 0.08 & 0.05 & 0.43 & 0.33 & 0.20 & (2000) \\ \mbox{Divergence time estimates} & \mbox{Lossel} & 12.8 & 0.29 & 4.25 & 1.30 & 15.7 \text{Myr} & 20.2 \text{Myr} & 23.9 \text{Myr} & 15.7 \text{Myr} & 20.2 \text{Myr} & 21.64 & 0.29 & 4.25 & 13.06 & 7.80 & 36.56 & 28.33 & 16.92 & (1993) \\ \mbox{Aspergillus niger, } & \mbox{Lossele} & 1.47 & 21.62 & 16.76 & 10.00 & 46.92 & 36.36 & 21.71 & 0.003 & 0.54 & 2.52 & 1.95 & 1.17 \\ \mbox{Nonsynonymous} & 904.4 & 0.08 & 1.18 & 0.90 & 0.54 & 2.52 & 1.95 & 1.17 & 0.00 & 0.54 & 0.55 & 0.54 & 0.55 & 0.54 & 0.55 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.55 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.55 & 0.54 & 0.54 & 0.55 & 0.54 & 0.54 & 0.55 & 0.54 & 0.55 & 0.54 & 0.54 & 0.55 & 0.54 & 0.55 & 0.54 & 0.54 & 0.54 & 0.55 & 0.54 & 0.54 & 0.55 & 0.54 & 0.55 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.55 & 0.54 & 0.54 & 0.55 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.54 & 0.55 & 0.54 $	b. aermannais	C46201VI	EXOIIS Third bases	70C	0.14	0.0 25 C	CC.U 174	1.04	2.70 9.03	7.00	1.20 4 18	Bowen et al. (1992) Nino-Veσa et al	
Nonsynonynous 412.3 0.02 0.11 0.08 0.05 0.43 0.33 0.20 cred cred 34.0 Myr 43.9 Myr 73.6 Myr 15.7 Myr 20.2 Myr 33.9 Myr Divergence time estimates Aspergillus niger, L03811, 34.0 Myr 43.9 Myr 73.6 Myr 15.7 Myr 20.2 Myr 33.9 Myr Aspergillus niger, L03811, Aspergillus niger, 15.7 Myr 20.2 Myr 33.9 Myr Divysdale et al. Aspergillus niger, L03811, Aspergillus niger, 15.7 Myr 20.2 Myr 42.6 (1993) Aspergillus niger, L0381 1.5 16.85 13.06 7.80 36.56 28.33 16.92 Synonymous 358.6 1.47 21.62 16.76 10.00 46.92 36.36 21.71 Nonsynonymous 904.4 0.08 0.54 2.52 1.95 1.17			Synonymous	154.7	0.57	2.86	2.22	1.32	11.50	8.92	5.32	(2000)	
<i>creA</i> Divergence time estimates <i>Aspergillus niger</i> , L03811, A. <i>orizae</i>			Nonsynonymous	412.3	0.02	0.11	0.08	0.05	0.43	0.33	0.20		
Divergence time estimates 34.0 Myr 43.9 Myr 73.6 Myr 15.7 Myr 20.2 Myr 33.9 Myr Aspergillus niger, L03811, Drysdale et al. Drysdale et al. Drysdale et al. Aspergillus niger, L03811, Drysdale et al. Drysdale et al. Drysdale et al. A. orizae More in the second state 1284 0.29 4.25 3.29 1.96 9.21 7.14 4.26 (1993) A. orizae Mirch bases 4.28 1.15 16.85 13.06 7.80 36.56 28.33 16.92 Synonymous 358.6 1.47 21.62 16.76 10.00 46.92 36.36 21.71 Nonsynonymous 904.4 0.08 1.18 0.90 0.54 2.52 1.17	creA												
A. orizae JJ272151 Exons 1284 0.29 4.25 3.29 1.96 9.21 7.14 4.26 (1993) A. orizae Third bases 428 1.15 16.85 13.06 7.80 36.56 28.33 16.92 Synonymous 358.6 1.47 21.62 16.76 10.00 46.92 36.36 21.71 Nonsynonymous 904.4 0.08 1.18 0.90 0.54 2.52 1.95 1.17	Divergence time estimates	1.03811				34.0 Myr	43.9 Myr	73.6 Myr	15.7 Myr	20.2 Myr	33.9 Myr	Drvsdale et al.	
Third bases 428 1.15 16.85 13.06 7.80 36.56 28.33 16.92 Synonymous 358.6 1.47 21.62 16.76 10.00 46.92 36.36 21.71 Nonsynonymous 904.4 0.08 1.18 0.90 0.54 2.52 1.95 1.17	A. orizae	AJ272151	Exons	1284	0.29	4.25	3.29	1.96	9.21	7.14	4.26	(1993)	
Nonsynonymous 904.4 0.08 1.18 0.90 0.54 2.52 1.95 1.17			Third bases	428 358 6	1.15	16.85 21.62	13.06 16.76	7.80	36.56 46 97	28.33 36 36	16.92 21.71		
			Nonsynonymous	904.4	0.08	1.18	0.90	0.54	2.52	1.95	1.17		

2320 Kasuga et al.

Table 1 Estimate

Table 1 Continued												
					RAT	TE ESTIMATES / (10)	AT THE GIVEN E.	S CALIBRATIO site per year)	n time Point	s		
	ACCESSON NO			DNA Dis.		NPRS			LF			
SPECIES PAIR	AND LOCUS	SITE	L (bp)	TANCE ^b	310 Myr	400 Myr	670 Myr	310 Myr	400 Myr	670 Myr	REFERENCE	
<i>pyrG</i> Divergence time estimates					34.0 Myr	43.9 Myr	73.6 Myr	15.7 Myr	20.2 Myr	33.9 Myr		
Aspergillus fumigatus, A. niger ^a	Y11303, X96734	Exons Third bases	834 278 200 1	0.24 0.79	3.51 11.60	2.72 8.99	1.62 5.37	7.62 25.18	5.90 19.52	3.52 11.65	Weidner et å (1998)	al.
		Synonymous Nonsynonymous	229.7 601.3	0.07 0.07	14.69 1.03	0.08	0.80 0.48	31.88 2.23	24.70 1.73	14.75 1.03		
arf Divergence time estimates					90.1 Mvr	177 & Mvr	214.1 Mvr	74.6 Mvr	31.8 Myr	53.3 Mur		
Divergence unite estimates H. capsulatum,	AF072341,				1711 IVI	121.0 MJ	214.1 MJ	24.0 MJ	14M 0.1C	IVINI C.CC	Kasuga et al.	
B. dermatitidis	AY013310	Exons Introns	280 174	0.06 0.48	0.28 2.42	0.22 1.88	0.13 1.12	1.13 9.74	0.88 7.55	0.52 4.51	(1999)	
		Third bases Synonymous	94 78.8	0.15	0.78	0.60	0.36 0.51	3.12 4.46	2.42 3.46	1.45		
tbul												
Divergence time estimates					99.1 Myr	127.8 Myr	214.1 Myr	24.6 Myr	31.8 Myr	53.3 Myr	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
H. capsulatum, B. dermatitidis.	AY013312, AY013313	Exons	52	0.10	0.53	0.41	0.24	2.13	1.65	0.98	Kasuga et al. (1999)	
		Introns	213	0.49	2.49	1.93	1.15	10.01	7.76	4.63		
		I nird Dases Synonymous	18 15.6	0.28	1.58 1.58	1.08	0.04 0.73	ес.с 6.37	4.55 4.94	2.95 2.95		
SLI												
Divergence time estimates	A E 2 7 7 2 7 7				99.1 Myr	127.8 Myr	214.1 Myr	24.6 Myr	31.8 Myr	53.3 Myr		
B. dermatitidis.	AF322388 AF322388	ITS1 and ITS2	327	0.21	1.08	0.83	0.50	4.33	3.35	2.00		
		5.85 rDNA	529	0.14	0.73	0.57	0.34	2.84	2.28	1.36		
Divergence time estimates	A 171 66120				26.8 Myr	34.6 Myr	58.0 Myr	11.5 Myr	14.8 Myr	24.8 Myr		
Artnroaerma ıncurvanum, T. rubrum	AF100129, AF170472	ITS1 and ITS2	360	0.25	4.58	3.55	2.12	10.71	8.30	4.95		
		5.85 rDNA	518	0.16	3.06	2.37	1.41	7.15	5.54	3.31		
Divergence time estimates					106.0 Myr	136.8 Myr	229.1 Myr	45.8 Myr	59.0 Myr	98.9 Myr		
Eremascus albus, Ascosphaera apis	U18359, U18362	ITS1 and ITS2	337	0.28	1.34	1.04	0.62	3.10	2.40	1.44		
		5.85 rDNA	494	0.18	0.87	0.67	0.40	2.01	1.55	0.93		
Divergence time estimates	A E001067				4.5 Myr	5.8 Myr	9.6 Myr	1.7 Myr	2.2 Myr	3.8 Myr		
r encuum cu ysogenum, Eupenicillium crustaceum	AFU3466 AF033466	ITS1 and ITS2	342	0.01	1.32	1.02	0.61	3.40	2.63	1.57		
		5.8S rDNA	531	0.01	0.85	0.66	0.39	2.18	1.69	1.01		

Locus and Accesson No.				RA	TE ESTIMATES 4 (10	AT THE GIVEN E	S CALIBRATIO site per year)	n time Point	S	
			DNA Dis-		NPRS			LF		
SPECIES PAIR AND LOCUS	SITE	L (bp)	TANCE ^b	310 Myr	400 Myr	670 Myr	310 Myr	400 Myr	670 Myr	REFERENCE
Divergence time estimates				23.8 Myr	30.7 Myr	51.5 Myr	11.5 Myr	15.0 Myr	25.1 Myr	
Latronum turn, Uteol. , AF109330 II Aspergillus fumigatus AF109330 II	TS1 and ITS2	302	0.38	7.92	6.14	3.66	16.21	12.56	7.50	
T	5.8S rDNA	459	0.23	4.77	3.69	2.21	9.76	7.56	4.52	
Divergence time estimates				80.8 Myr	104.2 Myr	174.5 Myr	36.9 Myr	50.2 Myr	84.0 Myr	
Dysocotanty need, 010001, Thermoascus crustaceus U18353 II	TS1 and ITS2	335	0.16	1.02	0.79	0.47	2.12	1.84	0.98	
	5.8S rDNA	492	0.11	0.71	0.55	0.33	1.48	1.14	0.68	

ger

^o DNA distance was estimated using the Kimura two-parameter method

clade and how estimates were influenced by assumptions of rate constancy or rate variability (LF vs. NPRS), the substitution model (ML vs. NJ), or the ES divergence value (310, 400, or 670 Myr). Using data for all four pairs of species, K_S for CHS1 varied 4.7-fold between the NPRS and LF algorithms, 1.5-fold between the ML-Tamura Nei and NJ-Kimura two-parameter models, and 2.2-fold over the three ES calibration points. The average synonymous substitution rates K_s estimated with NPRS and the three ES values were 7.8 \pm 3.4 × 10⁻⁹ (310 Myr), 6.1 \pm 2.6 × 10⁻⁹ (400 Myr), and $3.0 \pm 1.3 \times 10^{-9}$ (670 Myr) ($n = 4; \pm$ values are one standard deviation). The average K_S value estimated with LF and an ES of 400 Myr was 28.5 \pm 28.7 \times 10^{-9} . Most of the variation could be explained by use of the LF algorithm or the NPRS algorithm and its underlying assumptions of rate constancy or variability. In spite of the several sources of variation, K_S estimates ranged only from 10^{-9} to 10^{-8}

Only one pair of sister species could be used for substitution rate estimation in the remaining proteincoding loci, CHS2, creA, pyrG, arf, and tub1. K_s values for these genes are also in the range of 10^{-9} - 10^{-8} . The $K_{\rm S}$ values for the four species pairs at the CHS1 and one value for each of the CHS2, creA, pyrG, arf, and tub1 loci gave the average of $6.3 \pm 5.4 \times 10^{-9}$ (n = 9) when the ES divergence of 400 Myr and NPRS were used. On average, in CHS1, CHS2, creA, and pyrG, synonymous substitutions were found to accumulate about 22 times faster ($K_S/K_A = 22.2 \pm 8.4, n = 7$) than nonsynonymous substitutions. This value is higher than the average ratio of *Drosophila* genes ($K_S/K_A = 8.2$) but is not an extreme value (Li 1997). The K_S/K_A ratio is positively influenced by both the intensity of negative selection and the population size.

The *arf* and *tub1* data sets consist of exons and introns, for which rates were estimated separately. At *arf* and *tub1* loci, introns mutate 2.2 and 1.6 times faster than synonymous sites and 8.6 and 4.7 times faster than all sites in exons, respectively.

The *ITS* in the *rDNA* repeat unit have been widely used for phylogenetic studies (e.g., Berbee et al. 1995; Cullings, Szaro, and Bruns 1996) and population studies (e.g., O'Donnell 1992; Kasuga et al. 1993). DNA substitution rates were estimated using six pairs of sister species. Unlike with protein-coding genes such as CHS1 (Bowen et al. 1992) and RPB2 (Liu, Whelen, and Hall 1999), multiple alignment of DNA sequences from species belonging to different genera was impossible due to the frequent occurrence of indels. The length at the spacer regions (ITS1 + ITS2) varied greatly between the species used in this study, ranging from 302 to 372 bp. DNA substitution rates at the ITS region (ITS1, ITS2 + 5.8S rDNA) vary extensively between lineages, and the average was $1.4 \pm 1.3 \times 10^{-9}$ (n = 5) when the NPRS and an ES divergence of 400 Myr were used. Despite the large functional difference between the ITS region and the protein-coding genes, DNA substitution rates are found to be comparable. In a wide range of plants, substitution rates at the ITS region fall between 1.72×10^{-9} and 7.83 \times 10⁻⁹ (Richardson et al. 2001). Therefore,

L

1

estimates for Eurotiomycetes are comparable to the lower values for plants.

Neutral Mutation Rates Across Kingdoms

Overall, neutral mutation rates in protein-coding genes, which were measured as synonymous substitution rates or approximated as substitution rates at the third bases of codons, ranged from 0.9×10^{-9} to 16.7 \times 10⁻⁹ substitutions per site per year (NPRS and ES = 400 Myr). These values are in the range of DNA substitution rates in most of the protein-coding genes in plants (Gaut et al. 1996), animals (Li 1997), and bacteria (Ochman, Elwyn, and Moran 1999). For example, synonymous substitution rates in cereals are in the range of 5.1×10^{-9} to 7.1×10^{-9} (Wolfe, Sharp, and Li 1989). In mammals, rates were between 1.6×10^{-9} and $6.4 \times$ 10^{-9} , and in *Drosophila*, rates were between 3.7×10^{-9} and 30.0×10^{-9} (Li 1997; Moriyama and Powell 1997). In Escherichia coli and Salmonella typhimurium, substitution rates varied from 0.14 imes 10⁻⁹ to 5.6 imes 10⁻⁹ (Ochman and Wilson 1987). Although neutral mutation rates vary across loci, their range is surprisingly constant among the four major clades plants, animals, bacteria, and fungi in spite of the enormous differences in cellular organization, body size, generation time, and ecology of these organisms.

Acknowledgments

The authors are grateful to Michael Sanderson at the University of California at Davis, Chen Su and Helen Piontkivska at Pennsylvania State University, and Naoko Takezaki at the Max Planck Institute for helping with data analysis; Masato Sugiyama at Mitsubishi Chemical for sharing unpublished data with us; and Anne Pringle for comments on the manuscript. Financial support for this work was provided by the National Institutes of Health (grant A1 43491 to J.W.T.).

LITERATURE CITED

- BERBEE, M. L., and J. W. TAYLOR. 1992. Two ascomycete classes based on fruiting-body characters and ribosomal DNA sequence. Mol. Biol. Evol. **9**:278–284.
- ——. 1993. Dating the evolutionary radiations of the true fungi. Can. J. Bot. 71:1114–1127.
- ——. 2001. Fungal molecular evolution: gene trees and geologic time. Pp. 229–246 *in* D. J. McLAUGHLIN, E. Mc-LAUGHLIN, and P. A. LEMKE, eds. The Mycota. Springer, Berlin.
- BERBEE, M. L., A. YOSHIMURA, J. SUGIYAMA, and J. W. TAY-LOR. 1995. Is *Penicillium* monophyletic? An evaluation of phylogeny in the family Trichocomaceae from 18S, 5.8S, and ITS ribosomal DNA sequence data. Mycologia **87**:210– 222.
- BOWEN, A. R., J. L.-P. CHEN-WU, M. MOMANY, R. YOUNG, P. J. SZANISZLO, and P. W. ROBBINS. 1992. Classification of fungal chitin synthases. Proc. Natl. Acad. Sci. USA 89:519– 523.
- CULLINGS, K. W., T. M. SZARO, and T. D. BRUNS. 1996. Evolution of extreme specialization within a lineage of ectomycorrhizal epiparasites. Nature 379:63–66.

- DOOLITTLE, R. F., D. F. FENG, S. TSANG, G. CHO, and E. LIT-TLE. 1996. Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 271: 470–477.
- DRYSDALE, M. R., S. E. KOLZE, and J. M. KELLY. 1993. The Aspergillus niger carbon catabolite repressor encoding gene, creA. Gene 130:241–245.
- ERIKSSON, O. E., and K. WINKA. 1998. Families and higher taxa of Ascomycota. Myconet 1:17–24.
- GAUT, B. S., B. R. MORTON, B. C. MCCAIG, and M. T. CLEGG. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc. Natl. Acad. Sci. USA **93**:10274–10279.
- HECKMAN, D. S., D. M. GEISER, B. R. EIDELL, R. L. STAUFFER, N. L. KARDOS, and S. B. HEDGES. 2001. Molecular evidence for the early colonization of land by fungi and plants. Science 293:1129–1133.
- HEY, J., and J. WAKELEY. 1997. A coalescent estimator of the population recombination rate. Genetics 145:833–846.
- KASUGA, T., C. WOODS, S. WOODWARD, and K. MITCHELSON. 1993. *Heterobasidion annosum* 5.8s ribosomal DNA and internal transcribed spacer sequence: rapid identification of European intersterility groups by ribosomal DNA restriction polymorphism. Curr. Genet. 24:433–436.
- 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.
- KUMAR, S., K. TAMURA, I. JAKOBSEN, and M. NEI. 2000. MEGA: molecular evolutionary genetics analysis. Arizona State University, Tempe.
- LANGLEY, C. H., and W. FITCH. 1974. An estimation of the constancy of the rate of molecular evolution. J. Mol. Evol. 3:161–177.
- LI, W.-H. 1997. Rates and patterns of nucleotide substitution. Pp. 177–213 in Molecular evolution. Sinauer Associates, Sunderland, Mass.
- LIU, Y. J., S. WHELEN, and B. D. HALL. 1999. Phylogenetic relationships among ascomycetes: evidence from an RNA polymerase II subunit. Mol. Biol. Evol. 16:1799–1808.
- MORIYAMA, E. N., and J. R. POWELL. 1997. Synonymous substitution rates in *Drosophila*: mitochondrial versus nuclear genes. J. Mol. Evol. 45:378–391.
- NEI, M., and S. KUMAR. 2000. Molecular clocks and linearized trees. Pp. 187–206 *in* Molecular evolution and phylogenetics. Oxford University Press, New York.
- NINO-VEGA, G. A., C. A. MUNRO, G. SAN-BLAS, G. W. GOO-DAY, and N. A. R. GOW. 2000. Differential expression of chitin synthase genes during temperature induced dimorphic transitions in *Paracoccidioides brasiliensis*. Med. Mycol. 38:31–39.
- OCHMAN, H., S. ELWYN, and N. A. MORAN. 1999. Calibrating bacterial evolution. Proc. Natl. Acad. Sci. USA 96:12638– 12643.
- OCHMAN, H., and A. C. WILSON. 1987. Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. J. Mol. Evol. 26:74–86.
- O'DONNELL, K. 1992. Ribosomal DNA internal transcribed spacers are highly divergent in the phytopathogenic ascomycete *Fusarium sambucinum* (*Gibberella pulicaris*). Curr. Genet. **22**:213–220.
- OGAWA, H., A. YOSHIMURA, and J. SUGIYAMA. 1997. Polyphyletic origin of species of the anamorphic genus *Geosmithia* and the relationships of the cleistothecial genera: evidence from 18S, 5S, and 28S rDNA sequence analysis. Mycologia **89**:756–771.

- POSADA, D., and K. A. CRANDALL. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14**:817– 818.
- RICHARDSON, J. E., R. T. PENNINGTON, T. D. PENNINGTON, and P. M. HOLLINGSWORTH. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. Science 293:2242–2245.
- SANDERSON, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14:1218–1231.
- SUGIYAMA, M., A. OHARA, and T. MIKAWA. 1999. Molecular phylogeny of onygenalean fungi based on small subunit ribosomal DNA (SSU rDNA) sequences. Mycoscience 40: 251–258.
- SWOFFORD, D. L. 2001. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Mass.
- TAKEZAKI, T., A. RZHETSKY, and M. NEI. 1995. Phylogenetic test of the molecular clock and linearized tree. Mol. Biol. Evol. **12**:823–833.

- TAYLOR, T. N., H. HASS, and H. KERP. 1999. The oldest fossil ascomycetes. Nature **399**:648.
- WATTERSON, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7:256–276.
- WEIDNER, G., C. D'ENFERT, A. KOCH, P. C. MOL, and A. A. BRAKHAGE. 1998. Development of a homologous transformation system for the human pathogenic fungus *Aspergillus fumigatus* based on the pyrG gene encoding orotidine 5'monophosphate decarboxylase. Curr. Genet. 33:378–385.
- WOLFE, K. H., P. M. SHARP, and W. H. LI. 1989. Rates of synonymous substitution in plant nuclear genes. J. Mol. Evol. 29:208–211.

JULIAN ADAMS, reviewing editor

Accepted July 25, 2002