JOURNAL OF BACTERIOLOGY, Apr. 1990, p. 2036–2045
0021-9193/90/042036-10$02.00/0

Vol. 172, No. 4

Sequence Analysis and Expression of the Two Genes for Elongation Factor 1α from the Dimorphic Yeast *Candida albicans*

PAULA SUNDSTROM,* DEBORAH SMITH, AND PAUL S. SYPHERD

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, Irvine, California 92717

Received 10 August 1989/Accepted 8 January 1990

Two *Candida albicans* genes that encode the protein synthesis factor elongation factor 1α (EF-1α) were cloned by using a heterologous *TEF1* probe from *Mucor racemosus* to screen libraries of *C. albicans* genomic DNA. Sequence analysis of the two clones showed that regions of DNA flanking the coding regions of the two genes were not homologous, verifying the presence of two genes, called *TEF1* and *TEF2*, for EF-1α in *C. albicans*. The coding regions of *TEF1* and *TEF2* differed by only five nucleotides and encoded identical EF-1α proteins of 458 amino acids. Both genes were transcribed into mRNA in vivo, as shown by hybridization of oligonucleotide probes, which bound specifically to the 3′ nontranslated regions of *TEF1* and *TEF2*, respectively, to *C. albicans* total RNA in Northern (RNA) blot analysis. The predicted EF-1α protein of *C. albicans* was more similar to *Saccharomyces cerevisiae* EF-1α than to *M. racemosus* EF-1α. Furthermore, codon bias and the promoter and termination signals of the *C. albicans* EF-1α proteins were remarkably similar to those of *S. cerevisiae* EF-1α. Taken together, these results suggest that *C. albicans* is more closely related to the ascomycete *S. cerevisiae* than to the zygomycete *M. racemosus*.

Candida albicans is a dimorphic fungal pathogen that acquires virulence and the ability to adhere to host tissues as it undergoes the transition from yeast to mycelial forms (26, 27, 37, 45, 61). The phylogenetic relationship of *C. albicans* to other fungi is uncertain because of the lack of a sexual cycle, resulting in its being placed in the form subdivision *Deuteromycotina* (Fungi Imperfecti) along with a diverse group of 15,000 other species (48). To clarify the relationship of *C. albicans* to other fungi, an analysis of macromolecules with conserved function such as rRNA or proteins could provide valuable information on interrelationships among species (K. H. Schliefer and W. Ludwig, *in* B. Fernholm, K. Bremer, and H. Jornvall, ed., *The Hierarchy of Life*, in press).

Our long-term interest in gene expression during fungal dimorphism in the zygomycete *Mucor racemosus* has resulted in extensive analysis of the highly conserved component of the translational apparatus, elongation factor 1α (EF-1α). In addition to belonging to the class of proteins that bind GTP (G proteins). EF-1α (which is analogous to the bacterial protein EF-Tu) fulfills an essential cellular function in protein synthesis in that it binds charged tRNA molecules and transports them to the acceptor site on the ribosome adjacent to a growing polypeptide chain. Studies of mutant EF-1α proteins in *Saccharomyces cerevisiae* have shown that EF-1α can also control gene expression through its influence on translational accuracy (55), as has been shown in *Escherichia coli* (53, 65, 66). We have shown that in *M. racemosus*, EF-1α undergoes substantial posttranslational modification during the course of mycelium formation, resulting in the methylation of about 20% of its lysine residues (18). Such posttranslational changes make possible a role for EF-1α in the regulation of gene expression during morphogenesis and has raised questions about other aspects of regulation of EF-1α expression that might also contribute to morphogenetic changes. This consideration led us to study the genes for EF-1α from *M. racemosus*. We have shown that *M. racemosus* has three genes for EF-1α, denoted *TEF1*, *-2*, and *-3* (40, 42), which are transcribed at different levels, and that the ratio of transcription from each gene does not vary significantly with stage of growth (41).

We have extended our studies on EF-1α to *C. albicans* to determine whether any of the characteristics of fungal EF-1α proteins could be used to delineate the phylogenetic relationship of *C. albicans* to other fungi and to determine whether there were multiple genes that were differentially expressed during the two growth phases. Using an *M. racemosus* gene probe, we have identified and cloned two genes for EF-1α from *C. albicans*, which we named *TEF1* and *-2*, in keeping with the convention of naming *C. albicans* genes after corresponding genes in *S. cerevisiae* (31). The predicted EF-1α protein of *C. albicans* is more similar to that of the ascomycete *S. cerevisiae* than to that of the zygomycete *M. racemosus*. This conclusion is further supported by an analysis of the transcriptional initiation and termination/processing regions of the genes.

MATERIALS AND METHODS

Bacterial and yeast strains and plasmids. *E. coli* JM109 was the host strain for bacteriophage M13mp18 (71), which was used for subcloning and sequencing. Replicative-form and single-stranded DNA were purified by the procedures recommended by Bethesda Research Laboratories, Inc. (Gaithersburg, Md.,) (5). *E. coli* LE 392 (46) was used as a host for bacteriophage lambda. RNA and DNA were prepared from *C. albicans* SC5314, a clinical isolate (29). The pH regimen developed by Brummel and Soll (9) was used to prepare log-phase yeast and germ tube-forming cells. One colony from a 4-day-old plate was placed in 100 ml of Lee medium (41) (pH 4.5) and grown overnight at room temperature with shaking until a density of 66 Klett units (early log phase) was reached. The culture was divided into two 50-ml samples and centrifuged for 5 min at 2,000 × *g* at room temperature.

* Corresponding author.

Samples were suspended to concentration of 10 Klett units (3×10^5 to 5×10^5 cells per ml) in Lee medium, prewarmed to 37°C, at pH 4.5 to prepare yeasts or in pH 6.5 Lee medium to prepare germ tube-forming cells. Cultures were incubated 3 to 3.5 h with shaking (200 rpm) until the organisms grown at pH 6.5 had germ tubes three to four times longer than the mother cell. Cultures reached a density of 55 Klett units at the time of harvest. Under these conditions, 100% of the cells growing at pH 6.5 had germ tubes. Less than 5% of organisms grown at pH 4.5 had initiated germ tube formation, and these tubes remained only 0.5 times the length of the mother cell throughout the 37°C incubation period.

Library screening. Two genomic libraries containing *C. albicans* SC5314 DNA were generously provided by Myra Kurtz and Donald Kirsch of Squibb Pharmaceuticals. One library consisted of DNA completely digested with *Eco*RI and inserted into the lambda vector 590 (13, 28). The other library consisted of DNA digested to completion with *Hin*dIII and inserted into the vector lambda 607 (13; Myra Kurtz, personal communication). The libraries were each plated on two 150-mm-diameter petri dishes, each containing approximately 2,000 plaques. Plaques were transferred to duplicate nitrocellulose filters and screened with one of two probes, which had been radiolabeled to 5×10^6 cpm by the random primer method (14, 15). The probe used for the *Hin*dIII library was a 1.2-kb *Hin*dIII fragment, released from a *Hin*dIII digest of a pUC9 clone containing a *Hin*dIII site in the vector just 5′ to the *TEF1* insert and a *Hin*dIII site at nucleotide 1232 in the *M. racemosus TEF1* coding region (42). This fragment was purified on an agarose gel and included the entire *M. racemosus TEF1* gene except for the 142 nucleotides at the 3′ end of the gene, and was used for screening the *Hin*dIII library. A fragment containing a region from the *C. albicans TEF1* gene identified in the *Hin*dIII library was subsequently used to probe the *Eco*RI library. This fragment was an 870-kilobase-pair (kb) *Eco*RI/*Hin*dIII fragment (see Fig. 2) internal to the *C. albicans TEF* coding region and was also gel purified before being used to probe the *Eco*RI library. Positive plaques were cut out of the plate and plaque purified by standard methods (11). Bacteriophage DNA was prepared according to standard methods (46).

Restriction endonuclease analysis and agarose gel electrophoresis. Restriction endonuclease digestions of DNA using *Hin*dIII (Boehringer Mannheim Biochemicals, Indianapolis, Ind.), *Eco*RI, *Bgl*II (New England Bio-Labs, Beverly, Mass.), and *Eco*RV (Bethesda Research Laboratories) were performed as instructed by the manufacturers. DNA fragments were resolved through 1% agarose gels containing 0.5 μg of ethidium bromide per ml, using a Tris-acetate buffer system (50 mM Tris [pH 8.0], 20 mM sodium acetate, 2 mM EDTA, pH 8.0). Total RNA was prepared and separated on denaturing formaldehyde-agarose gels (32).

Southern blot analysis. Genomic DNA prepared by the procedure of Boeke et al. (7) or bacteriophage DNA prepared as described above was resolved by electrophoresis through agarose gels and further analyzed by transfer to nitrocellulose filters by the procedure of Southern (62). The DNA transfer buffer was 20× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate, pH 7.0). Passive transfer was allowed to proceed over a 24-h period, after which the nitrocellulose filter was baked at 80°C under vacuum for 1.5 h before hybridization. Nitrocellulose filter hybridization conditions were essentially as previously described (40) except that filters were not prehybridized. Filters were wetted in 6× SSC. Wetting solution was removed and replaced with hybridization solution consisting of 50% deionized formamide, 5× Denhardt solution (1× Denhardt solution is 0.1% Ficoll, 0.1% bovine serum albumin, and 0.1% polyvinylpyrrolidone), 5× SSPE (20× SSPE is 3.6 M NaCl plus 200 mM NaH_2PO_4, pH 7.4), 20 mM EDTA, 0.1% sodium dodecyl sulfate, and 100 μg of denatured salmon sperm DNA per ml plus 5×10^6 cpm of an [α-^{32}P]dGTP-random-primer-labeled fragment of *M. racemosus TEF1* or an *Eco*RI-*Hin*dIII fragment internal to the *C. albicans TEF* coding region as described for library screening. Hybridizations were allowed to proceed for 24 h at 37°C in a shaking air incubator. The filters were washed twice for 15 min each in 0.2× SSC–0.1% sodium dodecyl sulfate. Lambda DNA digested with *Hin*dIII served as molecular weight markers.

DNA sequencing. DNA sequences were determined by the dideoxy-chain termination method (56), using the Sequenase kit (U.S. Biochemical Corp., Claremont, Calif.). Sequencing was initiated from the universal primer region of M13 clones and from a degenerate oligonucleotide primer prepared to bind to base pairs (bp) 31 to 51 of the *M. racemosus TEF1* coding strand (40) that is in a highly conserved, GTP-binding region of EF-1α. As *C. albicans* EF-1α sequences were identified, new oligonucleotide primers were prepared (Operon, San Pablo, Calif.) to extend the sequence.

Northern (RNA) blot analysis. RNA separated on denaturing agarose gels was transferred to nitrocellulose (36). Filters were wetted with 6× SSC as described above and then prehybridized in a solution containing 5× Denhardt solution, 5× SSPE, 20 mM EDTA, 0.1% sodium dodecyl sulfate, and 100 μg of denatured salmon sperm DNA per ml. Prehybridization solution was replaced with hybridization solution consisting of the above components and 10^6 cpm/20 pmol of gene-specific 26-mer oligonucleotides (synthesized by Operon; see Fig. 4B for exact sequence), which had been end labeled with [γ-^{32}P]dATP (Dupont NEN Research Products, Claremont, Calif.) and polynucleotide kinase (U.S. Biochemical Corp.) (1). Synthetic oligonucleotide probes were synthesized and purified by high-pressure liquid chromatography by Operon. After hybridization overnight at 42°C, blots were washed three times for 15 min each in 2× SSC at 42°C. Migration of rRNA bands was used to determine the molecular weights of hybridizing bands.

Filters were exposed to Kodak XAR-5 film at −70°C with a Cronex Lightning-Plus intensifier screen for 24 h.

RESULTS

Detection and cloning of *C. albicans* EF-1α genes. To identify the EF-1α genes of *C. albicans*, we used the *TEF1* gene of *M. racemosus* (42) to probe a Southern blot of *C. albicans* SC5314 genomic DNA digested to completion with *Hin*dIII. The *M. racemosus TEF1* probe hybridized to two fragments of 3.4 and 2.3 kb under conditions of high stringency (Fig. 1). The presence of two bands suggested that *C. albicans* had either two genes for EF-1α or one gene that was cut with *Hin*dIII. Subsequent sequencing of cloned copies of the two bands showed that *C. albicans* had two genes for EF-1α (see below). To clone the *C. albicans* EF-1α genes, the *M. racemosus TEF* probe was used to screen a library of SC5314 genomic DNA that had been digested to completion with *Hin*dIII and inserted into phage lambda 590 (see Materials and Methods). Of the 6,000 plaques screened, 14 hybridized strongly to the *M. racemosus TEF1* probe. After plaque purification, DNA from the positive clones was isolated, digested with *Hin*dIII, and analyzed by Southern

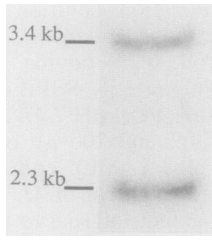


FIG. 1. Southern blot analysis of genomic DNA (10 μg) from *C. albicans* SC5314 digested with *Hin*dIII and probed with *TEF1* of *M. racemosus*.

blotting. Of the 14 positive clones, two contained inserts of 2.3 kb, six contained inserts of 3.4 kb, and the remaining clones had multiple inserts. Two fragments that comigrated with the *M. racemosus TEF*-hybridizing bands from Southern blot analysis of *C. albicans* genomic DNA were cloned into M13mp18 for DNA sequencing. The 2.3-kb clone was designated *TEF1*, and the 3.4-kb clone was designated *TEF2*.

Initial sequence analysis of the DNA adjacent to the 3′ ends of the two clones showed that these clones were truncated by 150 nucleotides, as judged by sequence comparison with the *M. racemosus TEF* sequence (64). To clone the remaining portion of the genes, we screened a library containing *C. albicans* SC5314 genomic DNA digested to completion with *Eco*RI and inserted into the vector bacteriophage lambda 607. This library was screened with the 0.87-kb *Eco*RI-*Hin*dIII fragment containing the central portion of the *C. albicans TEF* genes (Fig. 2). Nine plaques hybridized with this probe and were purified. Southern blot analysis of DNA from these clones showed that one clone had a 2.6-kb insert and five had 4.9-kb inserts. Fragments of the same size were also detected in Southern blot analysis of *C. albicans* genomic DNA digested with *Eco*RI when probed with the 0.87-kb *Eco*RI-*Hin*dIII fragment (not shown), indicating that the 3′ ends of both genes had been cloned. Insert DNA from the clone containing the 2.6-kb insert, as well as from one of the five clones containing the 4.9-kb insert, were cloned into M13mp18 in preparation for DNA sequencing.

To determine which of the *Eco*RI clones overlapped which of the *Hin*dIII clones, genomic DNA was digested with *Bgl*II, an enzyme that does not cut in the coding region of either *TEF* gene, and subjected to Southern blot analysis (not shown), using probes consisting of DNA flanking the coding regions of the genes (Fig. 2). Both the 1.33-kb *Hin*dIII-*Eco*RI region of the 3.4-kb *Hin*dIII clone (probe B; Fig. 2) and the 1.7-kb *Hin*dIII-*Eco*RI region of the 2.4-kb *Eco*RI clone (probe D) hybridized to a 9.4-kb *Bgl*II fragment, indicating that DNA containing probes B and D overlapped. Conversely, the 650-bp *Hpa*I fragment from the 5′ region of the 2.3-kb *Hin*dIII fragment (probe A) hybridized to a 2.5-kb *Bgl*II fragment. The 2.6-kb *Eco*RI fragment (probe C) hybridized only weakly to the 2.5-kb *Bgl*II fragment because of the presence of a *Bgl*II site 20 bp after the 3′ end of the gene. From these results, it was apparent that the 2.3-kb *Hin*dIII fragment overlapped the 4.5-kb *Eco*RI fragment (*TEF1*) and that the 3.4-kb *Hin*dIII fragment overlapped the 2.5-kb *Eco*RI fragment (*TEF2*) (Fig. 2).

Sequence analysis of coding regions. The reading frame of the *C. albicans TEF* genes was established by alignment with other EF-1α proteins. Sequences of the cloned inserts showed that the 5′-flanking regions of *TEF1* and *TEF2*, as well as the 3′-flanking regions, were not homologous, verifying that *C. albicans* has two genes for EF-1α. The two genes did not have introns and coded for identical proteins of 458 amino acids with a molecular weight of 49,940. The location of the initial ATG was based on alignment with *S. cerevisiae* and *M. racemosus* proposed proteins. Although these proteins have not been sequenced, the five amino-terminal amino acid residues match exactly those found in the corresponding *E. coli* protein EF-Tu, and the amino terminus of EF-Tu has been sequenced (23), strengthening the argument the amino-terminal residue of the predicted EF-1α proteins for the three fungi is the indicated methionine. The predicted proteins were compared with EF-1α from *S. cerevisiae* (50) and *M. racemosus* (40) (Fig. 3). Since *C. albicans* was found to have exactly the same number of amino acids for EF-1α as did the other two fungi, it was not necessary to introduce gaps to align the sequences. The *C. albicans* EF-1α gene was more like the *S. cerevisiae* than the *M. racemosus* EF-1α gene in that there were 42 differences between *S. cerevisiae* and *C. albicans* EF-1α but 61 differences between *M. racemosus* and *C. albicans* EF-1α. There were 71 amino acid differences between *M. racemosus* and *S. cerevisiae* predicted proteins.

Although the two EF-1α genes encoded identical putative proteins, there were five nucleotide differences in the coding sequences of the two genes (Fig. 4A). This resulted in two differences in the restriction maps in that only *TEF1* had a

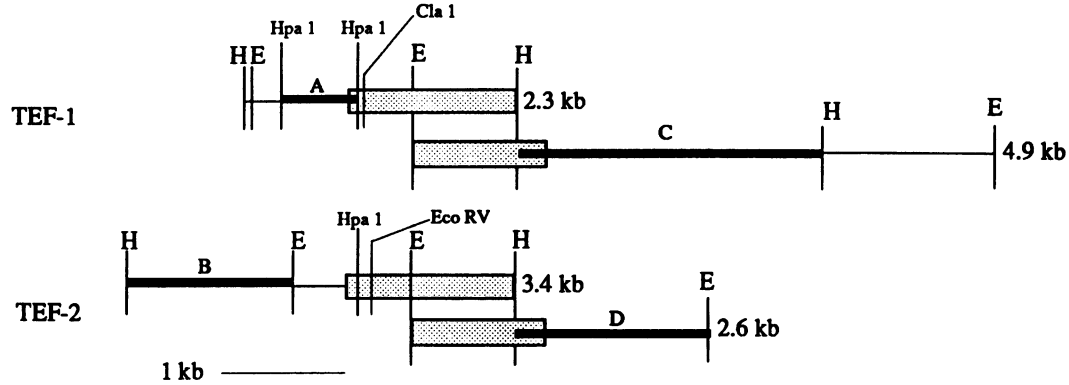


FIG. 2. Genomic organization of *C. albicans TEF1* and *TEF2*. Symbols: ▨, coding regions; ■, flanking probes used to identify overlapping *Hin*dIII and *Eco*RI fragments. *TEF1* has a *Cla*I site at position 100 that is not present in *TEF2*. *TEF2* has an *Eco*RI site at nucleotide 220 that is not present in *TEF1*.

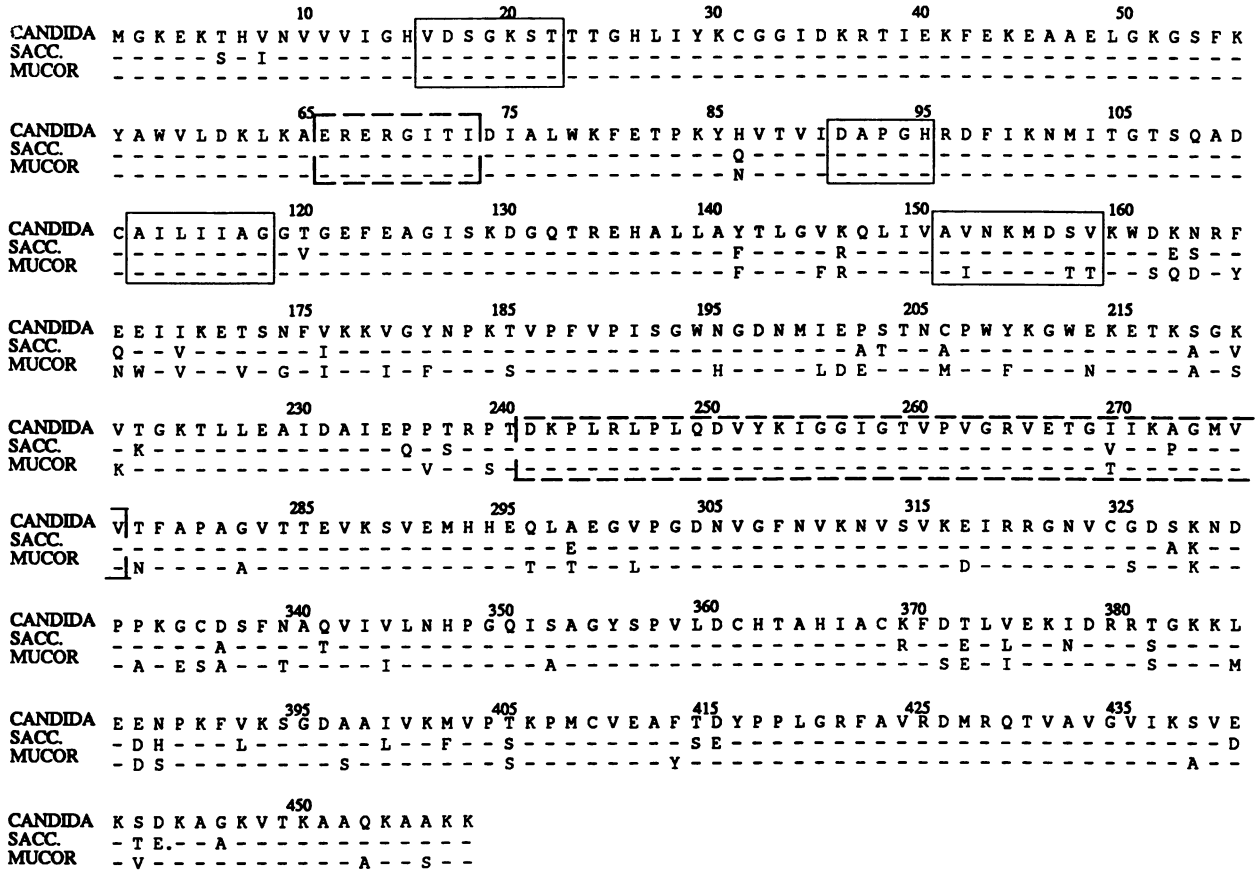


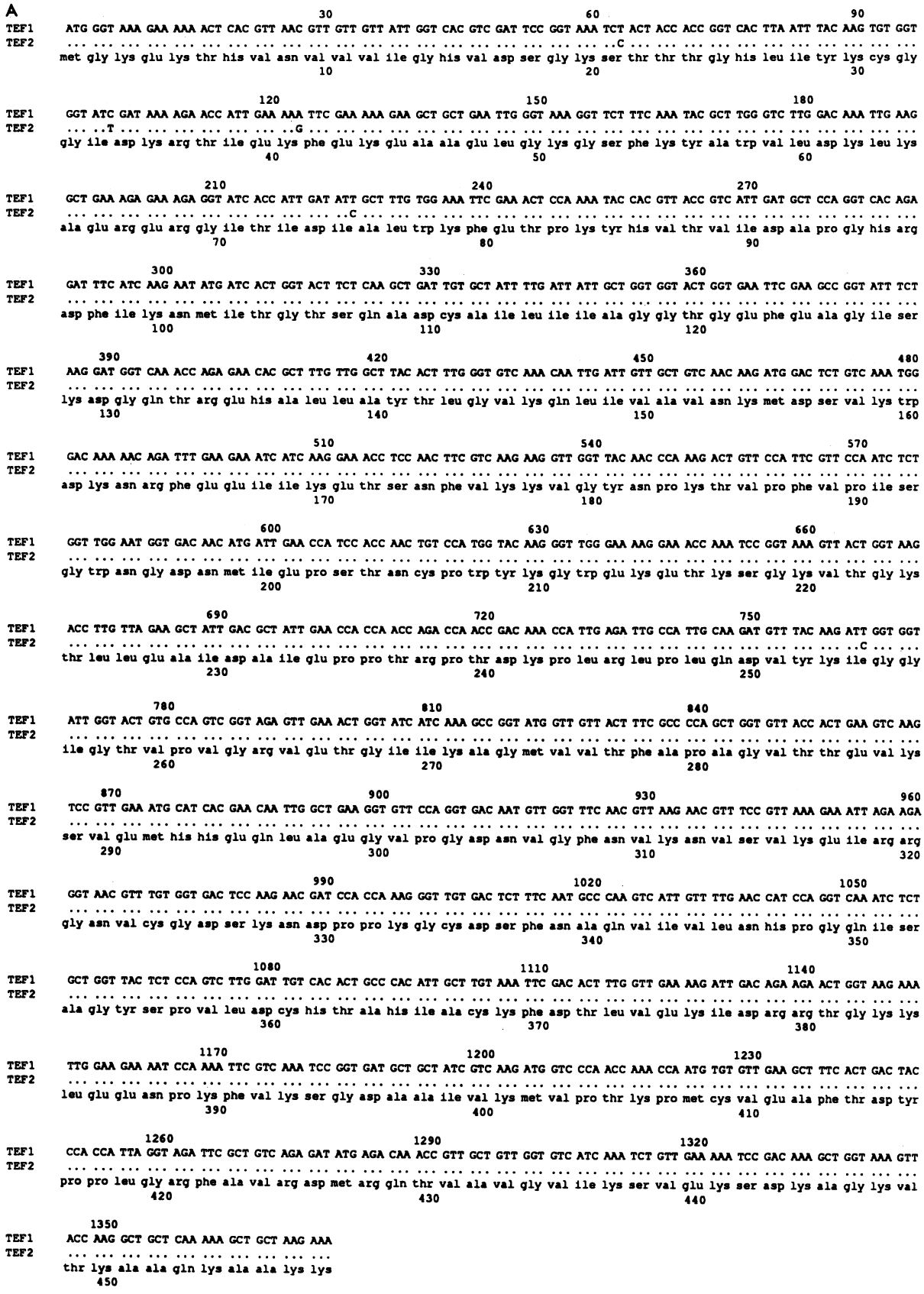
FIG. 3. Alignment of EF-1α predicted proteins from *C. albicans*, *S. cerevisiae*, and *M. racemosus*. Since all three proteins had 458 amino acids, it was not necessary to introduce gaps to align sequences. There were 42 amino acid differences between *C. albicans* and *S. cerevisiae* proteins, 61 differences between *C. albicans* and *M. racemosus* proteins, and 71 differences between *S. cerevisiae* and *M. racemosus* proteins. Solid boxes indicate GDP-binding regions; dashed boxes indicate tRNA-binding regions, determined on the basis of analogy with EF-Tu from *E. coli* (23). Dashes indicate amino acids identical to those of the *C. albicans* EF-1α predicted protein.

*Cla*I site at position 100 and only *TEF2* had an *Eco*RV site at position 220 relative to the ATG start codon (Fig. 2).

The codon usage of the two *C. albicans* EF-1α genes was nearly identical, as expected given the similarity of the two genes (Table 1), and was highly biased. A comparison of codon usages of EF-1α genes from *M. racemosus* (64), *S. cerevisiae* (50), and *C. albicans* showed that all three organisms had the same codon preferences for the amino acids valine, tyrosine, threonine, serine, phenylalanine, isoleucine, glutamine, and cysteine. In contrast, markedly different codon preferences were seen for the amino acids arginine and proline, where *C. albicans* and *S. cerevisiae* used AGA and CCA, respectively, whereas *M. racemosus* used exclusively CGT for arginine and three different codons for proline, the primary codon being CCC. *TEF* genes from *M. racemosus* in general displayed less codon bias than did *C. albicans* and *S. cerevisiae* genes. For example, *M. racemosus* used three different codons for leucine, glycine, and alanine, whereas *C. albicans* and *S. cerevisiae* used only one codon, GGT, for glycine, and only two codons each for leucine and alanine. In addition, for glutamate, *M. racemosus* used both available codons equally, whereas *S. cerevisiae* and *C. albicans* preferred GAA. In the case of asparagine and lysine, *S. cerevisiae* codon preferences were similar to those of *M. racemosus*, although both codons were used by all three organisms. In one case, histidine, the codon preference of *C. albicans* was more similar to that of *M. racemosus* than to that of *S. cerevisiae*, but overall, the codon preferences of *C. albicans* were more similar to those of *S. cerevisiae* than to those of *M. racemosus*.

Sequence analysis of flanking regions. Sequence analysis showed that overall, *TEF1* and *TEF2* had different 5′-flanking regions as well as different 3′-flanking regions. These flanking sequences were further compared for the presence of short sequences that might be important in regulating gene expression. The 3′-flanking regions of both genes possessed similar AT (ATTATATT...AAATATA[A] TATATATTTT) sequences 70 and 33 bp after the stop codon (Fig. 4B). On the basis of the size of the EF-1α mRNA (1.5 kb) and preliminary data suggesting a transcription initiation site 20 to 25 bp upstream of the ATG start codon (not shown), these A+T-rich sequence motifs are within the region where termination-polyadenylation signals would be expected to occur. Both genes had sequences similar to those found to be important for transcription termination in *S. cerevisiae* (72) (Fig. 4B). A sequence suggested to direct termination-polyadenylation in yeast cells, TAAATAAG (3), was found 60 nucleotides after the stop codon in *TEF2*. In addition, slightly further downstream from the stop codon, both genes had a sequence AATAAA similar to the polyadenylation signal found in mammalian mRNA (39).

More than 600 bp of 5′-flanking sequences from both genes were sequenced, and several features were found to be common to the two genes (Fig. 4C). First, potential TATAA


A
TEF1
TEF2

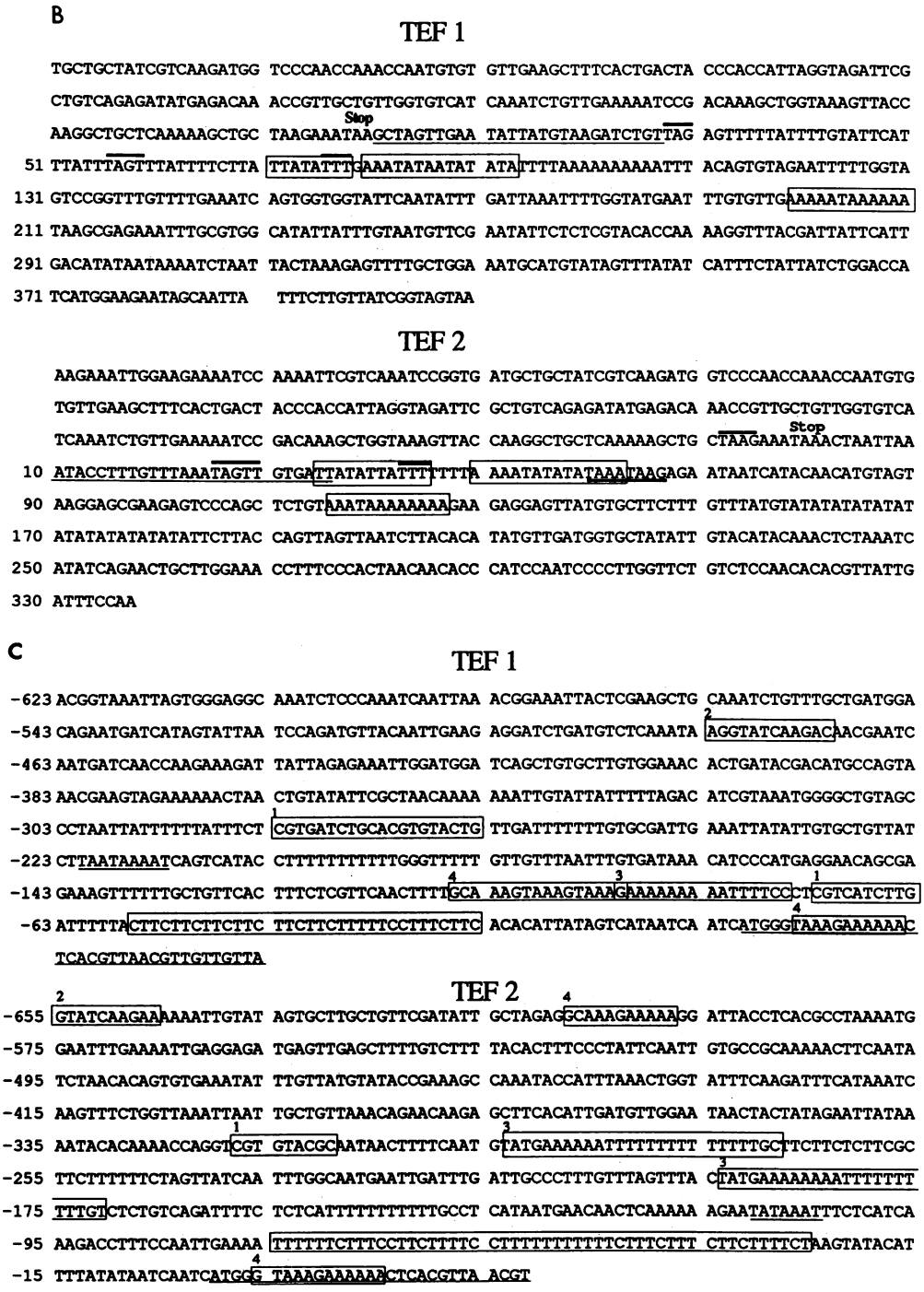


FIG. 4. DNA sequences of *TEF1* and *TEF2* of *C. albicans*. (A) Coding sequences. Dots indicate identical nucleotides. The five nucleotide differences between *TEF1* and *TEF2* are indicated. (B) 3′-Flanking sequences of *C. albicans TEF1* and *TEF2*. TAA stop codons show locations of the ends of EF-1α coding sequences. Sequences complementary to gene-specific oligonucleotide probes are underlined. Boxes indicate common features between the two genes. Underbars and overbars indicate postulated polyadenylation-termination sequences found in *S. cerevisiae* (72). (C) 5′-Flanking sequences of *C. albicans TEF1* and *TEF2*. The coding sequences and putative TATAA elements are underlined. Numbered boxed regions indicate common sequence motifs in the two genes. Unnumbered boxes enclose pyrimidine-rich sequences.

sequences were found 200 bp in *TEF1* and 92 bp in *TEF2* before the ATG start codon. Second, pyrimidine-rich tracts were found 56 to 24 bp in *TEF1* and 75 to 26 bp in *TEF2* before the ATG start codons. Third, both genes had several short sequence motifs in common (Fig. 4B and C) as well as stretches of T residues. 5′ sequences were also searched for the presence of RPG and HOMOL motifs, which activate transcription in *S. cerevisiae* (70) and are found upstream of genes whose protein products are involved in translation such as EF-1α and ribosomal protein genes (69). Although

TABLE 1. Comparison of frequency of codon usage of EF-1α genes from *C. albicans*, *M. racemosus*, and *S. cerevisiae*[a]

Amino acid	Codon	Codon usage				Amino acid	Codon	Codon usage				Amino acid	Codon	Codon usage			
		CA1	CA2	MUC	SAC			CA1	CA2	MUC	SAC			CA1	CA2	MUC	SAC
Ala	GCT	28	28	27	30	Gly	GGT	43	43	37	42	Pro	CCT	0	0	5	0
	GCC	5	5	9	7		GGC	0	0	2	0		CCC	0	0	16	0
	GCA	0	0	1	0		GGA	0	0	1	0		CCA	24	24	1	23
	GCG	0	0	0	0		GGG	0	0	1	0		CCG	0	0	0	0
Arg	CGT	0	0	17	0	His	CAT	2	2	1	5	Ser	TCT	10	9	9	14
	CGC	0	0	0	0		CAC	9	9	10	6		TCC	9	10	16	6
	CGA	0	0	0	0								TCA	0	0	0	0
	CGG	0	0	0	0	Ile	ATT	20	19	15	17		TCG	0	0	0	0
	AGA	16	16	0	18		ATC	12	13	19	13		AGT	0	0	0	1
	AGG	0	0	0	0		ATA	0	0	0	0		AGC	0	0	0	0
Asn	AAT	5	5	1	0	Leu	TTA	3	3	0	3	Thr	ACT	17	17	12	14
	AAC	12	12	15	16		TTG	18	18	5	21		ACC	16	16	18	14
							CTT	0	0	6	0		ACA	1	1	1	0
Asp	GAT	12	12	19	8		CTC	0	0	11	0		ACG	0	0	0	0
	GAC	13	13	7	16		CTA	0	0	0	0						
							CTG	0	0	0	0	Trp	TGG	6	6	6	6
Cys	TGT	8	8	6	6												
	TGC	0	0	0	1	Lys	AAA	30	30	3	3	Tyr	TAT	0	0	1	0
							AAG	22	22	46	46		TAC	9	9	7	8
Gln	CAA	9	9	8	12												
	CAG	0	0	0	0	Met	ATG	9	9	11	8	Val	GTT	29	29	15	26
													GTC	16	16	26	19
Glu	GAA	30	30	17	30	Phe	TTT	1	1	1	2		GTA	0	0	0	0
	GAG	0	0	12	1		TTC	14	14	16	15		GTG	0	0	0	1

[a] Codon usages of *C. albicans TEF* genes are shown in columns CA1 and CA2; those for *M. racemosus* and *S. cerevisiae* genes are shown in columns MUC and SAC, respectively.

no exact matches to the RPG and HOMOL consensus sequences were found, several sequences that were 70% similar to the consensus sequences were present (Fig. 4C). Finally, the 10 bp 5′ to the ATG start codons were identical for both genes.

mRNA levels of the two genes. To determine whether both genes were transcribed, gene-specific oligonucleotide probes complementary to 3′ nontranslated regions (Fig. 4B) were end labeled to the same specific activity and used to probe Northern blots of total RNA from yeast and mycelial growth phases of *C. albicans*. Both genes were expressed during yeast and mycelial growth, with no apparent quantitative differences (Fig. 5).

DISCUSSION

Analysis of the *C. albicans TEF* genes has provided several pieces of evidence which, when taken together, suggest that *C. albicans* is more related to the ascosporogenous yeast *S. cerevisiae* (subdivision *Ascomycotina*) than to *M. racemosus* (subdivision *Zygomycotina*). First, the predicted EF-1α proteins of *C. albicans* and *S. cerevisiae* were found to have fewer amino acid differences than those of *C. albicans* and *M. racemosus*. These results support those of a recent study which showed that the *C. albicans* β-tubulin gene is more related to the β-tubulin gene of *S. cerevisiae* than to that of either *Aspergillus nidulans* or *Neurospora crassa*, which are both ascomycetes but in different classes than *S. cerevisiae* (48, 59). Second, the *TEF* gene families of *C. albicans* and *S. cerevisiae* are each composed of two nearly identical genes lacking introns. In contrast, the *M. racemosus TEF* gene family consists of three members, all of which have introns, and which have two to three times more nucleotide differences between individual genes than have individual members of the *C. albicans* or *S. cerevisiae TEF* gene families. This suggests that *TEF* gene duplication occurred at an earlier time in *M. racemosus* than in *C. albicans* or *S. cerevisiae* and that *C. albicans* diverged more recently from *S. cerevisiae* than from *M. racemosus*.

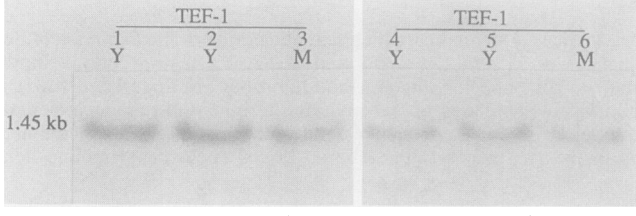


FIG. 5. Expression of both *C. albicans TEF* genes during yeast and mycelial growth. Shown is a Northern blot analysis using total RNA (20 μg) from log-phase cells (Y) grown at 25°C (lanes 1 and 4) or 37°C (lanes 2 and 5) and from organisms with germ tubes (M) (lanes 3 and 6). Nitrocellulose filters were probed with gene-specific oligonucleotides from *TEF1* (lanes 1 to 3) or *TEF2* (lanes 4 and 6).

Further evidence for the relatedness of *C. albicans* and *S. cerevisiae* was found in the comparison of codon usages of their EF-1α genes. The codon usage profiles of *C. albicans* and *S. cerevisiae* were remarkably similar to each other and different from that of *M. racemosus* for the amino acids arginine, which has six possible codons, and proline, which has four possible codons. The codon bias we found for *TEF1* and *TEF2* of *C. albicans* is in agreement with that found for the β-tubulin gene of *C. albicans* (another highly expressed gene) (59). The fact that codon usage was highly biased for the EF-1α genes from all three organisms was not surprising

given the fact that EF-1α genes are highly expressed. A correlation between biased codon usage of highly expressed genes and tRNA abundance is found for both *E. coli* and *S. cerevisiae* (4, 21, 53, 65, 66). By analogy, it is likely that the codon bias seen in *C. albicans* and *M. racemosus* EF-1α proteins also represents the predominant isoacceptor tRNA species in these organisms.

The fact that *C. albicans* is able to produce mRNA from both *TEF1* and *TEF2* indicates that both genes possess functional promoters. The 5′-flanking regions of the *C. albicans* EF-1α genes have several features that are found in *S. cerevisiae* promoters (44, 47, 63), including TATAA elements, T-rich sequences, and pyrimidine-rich sequences, and possibly upstream activation sequences (UASs) (51, 63). We found sequences similar to HOMOL and RPG UASs (10, 20, 67), which have been found in upstream regions from ribosomal protein genes of several different yeast species (49, 69), as well as a few examples of repeated sequences that might be important in transcription of the EF-1α genes. Clearly, positive identification of UASs in the *C. albicans* EF-1α genes will require further studies. The similarities between promoter regions of *C. albicans* and *S. cerevisiae* genes support the findings that several *C. albicans* genes (16, 24, 28–30, 31, 58) are expressed in *S. cerevisiae*. The alternative question of whether *S. cerevisiae* genes can be expressed in *C. albicans* is more difficult to answer because of the limited number of *C. albicans* mutants for which the biochemical lesion is known. To date, *LEU2* is the only known gene of *S. cerevisiae* that is expressed in *C. albicans* (24).

The 3′-flanking regions of the *C. albicans TEF* genes also have features in common with those of *S. cerevisiae* flanking regions. Both of the *C. albicans TEF* genes were found to have sequences similar to the *S. cerevisiae* termination sequence described by Zaret and Sherman (72). However, in addition to having the *S. cerevisiae*-like sequences, both *C. albicans TEF* genes were found to have another common sequence that may play a role in termination.

EF-1α proteins from all three fungi were more homologous to each other than to the EF-1α genes of higher organisms. The homologies of *C. albicans* EF-1α at the amino acid level to those of *S. cerevisiae* and *M. racemosus* were 90 and 85%, respectively, whereas *S. cerevisiae* and human EF-1α proteins are only 81% homologous (8). Likewise, *S. cerevisiae* and *Artemia* EF-1α proteins are only 79% homologous (40). In addition, the EF-1α predicted proteins from all three fungi are 458 amino acids in length, whereas EF-1α proteins from higher eucaryotic organisms are slightly larger, human and mouse proteins having 462 amino acids (8, 43), the *Artemia* protein having 459 (17), and the two *Drosophila* proteins having 463 and 464 (19). Tomato EF-1α has only 448 amino acids (52). Procaryotic genes for EF-Tu are shorter, ranging from 393 amino acids in *E. coli* EF-Tu (22), 405 in *Thermus thermophilus* (57), 400 in *Thermotoga maritima* (2), and 428 in the archaebacterium *Methanococcus vannielii* (33).

The presence of multiple genes for EF-1α in eucaryotes or EF-Tu in procaryotes is a feature of most organisms, the only known exceptions being *Bacillus* spp., *Chlamydia* spp., and archaebacteria (16a, 33, 60). The function of multiple genes for EF-1α is unknown. In the case of *S. cerevisiae*, deletion of either gene allows normal growth (12), suggesting that the presence of multiple genes is a safeguard against gene loss or inactivation. In other cases, multiple genes are regulated differently; for example, the *Drosophila* gene F2 is strongly expressed in the pupal developmental phase (19). In addition, the two genes for EF-Tu in *E. coli* are regulated differently (68). Since the *C. albicans* genes were expressed at equivalent levels, one would expect either gene to support growth, as is the case in *S. cerevisiae*.

ACKNOWLEDGMENTS

This work was supported by a Bank of America Gianinni Foundation fellowship to Paula Sundstrom and by Public Health Service grants GM23999 and AI25788 from the National Institutes of Health.

LITERATURE CITED

1. **Asubel, F., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl** (ed.) 1989. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
2. **Bachleitner, M., W. Ludwig, K. O. Stetter, and K. H. Schleifer.** 1989. Nucleotide sequence of the gene coding for the elongation factor Tu from the extremely thermophilic eubacterium *Thermotoga maritima*. FEMS Microbiol. Lett. **57**:115–120.
3. **Bennetzen, J. L., and B. D. Hall.** 1982. The primary structure of the *Saccharomyces cerevisiae* gene for alcohol dehydrogenase I. J. Biol. Chem. **257**:3018–3025.
4. **Bennetzen, J. L., and B. D. Hall.** 1982. Codon selection in yeast. J. Biol. Chem. **257**:3026–3031.
5. **Bethesda Research Laboratories, Inc.** Bethesda Research Laboratories, Inc., Gaithersburg, Md.
6. **Birnstiel, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41**:349–359.
7. **Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink.** 1985. Ty elements transpose through an RNA intermediate. Cell **40**:491–500.
8. **Brands, H. G. M., J. A. Maassen, F. J. Van Hemert, R. Amons, and W. Möller.** 1986. The primary structure of the α subunit of human elongation factor 1α. Structural aspects of guanine-nucleotide-binding sites. Eur. J. Biochem. **155**:167–171.
9. **Brummel, M., and D. R. Soll.** 1982. The temporal regulation of protein synthesis during synchronous bud or mycelium formation in the dimorphic yeast *Candida albicans*. Dev. Biol. **89**:211–224.
10. **Buchman, A. R., N. F. Lue, and R. D. Kornberg.** 1988. Connections between transcriptional activators, silencers, and telomeres and revealed by functional analysis of a yeast DNA-binding protein. Mol. Cell. Biol. **8**:5086–5099.
11. **Carlock, L. R.** 1986. Analyzing lambda libraries. Focus **8**:6–8.
12. **Cottrelle, P., D. Thiele, V. Price, S. Memet, J. Micoun, C. March, J. Buhler, A. Sentenac, and P. Fromageot.** 1985. Cloning, nucleotide sequence and expression of one of two genes coding for yeast elongation factor 1a. J. Biol. Chem. **260**: 3090–3096.
13. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. A manual for genetic engineering. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
15. **Feinberg, A. P., and B. Vogelstein.** 1984. Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137**:266–267.
16. **Gillum, A. M., E. Y. H. Tsay, and D. R. Kirsch.** 1984. Isolation of the *Candida albicans* gene for orotidine-5′-phosphate decarboxylase by complementation of *S. cerevisiae* ura3 and *E. coli* pyrF mutations. Mol. Gen. Genet. **198**:179–182.

16a. **Goldstein, B. P., G. Zaffaroni, O. Tiboni, B. Amiri, and M. Denaro.** 1989. Determination of the number of tuf genes in *Chlamydia trachomatis* and *Neisseria gonorrhoea*. FEMS Microbiol. Lett. **51**:305–309.

17. **Hemert, V. F. J., R. Amons, W. J. M. Pluijms, H. van Ormondt, and W. Moller.** 1984. The primary structure of elongation factor EF-1α from the brine shrimp *Artemia*. EMBO J. **3**:1109–1113.

18. **Hiatt, W. R., R. Garcia, W. C. Merrick, and P. S. Sypherd.** 1982. Methylation of elongation factor 1α from the fungus *Mucor racemosus*. Proc. Natl. Acad. Sci. USA **79**:3433–3437.
19. **Hovemann, B., S. Richter, U. Walldorf, and C. Cziepluch.** 1988. Two genes encode related cytoplasmic elongation factors 1α (EF-1α) in *Drosophila melanogaster* with continuous and stage specific expression. Nucleic Acids Res. **16**:3175–3194.
20. **Huet, J., and A. Sentenac.** 1987. TUF, the yeast DNA-binding factor specific for UAS rpg upstream activating sequences: identification of the protein and its DNA-binding domain. Proc. Natl. Acad. Sci. USA **84**:3648–3652.
21. **Ikemura, T.** 1982. Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. J. Mol. Biol. **158**:573–597.
22. **Jones, M. D., T. Peterson, K. Nielson, S. Magnusson, L. Sottrup-Jensen, K. Gausing, and B. Clark.** 1980. The complete amino acid sequence of elongation factor Tu from *E. coli*. Eur. J. Biohem. **108**:507–526.
23. **Jurnak, F.** 1985. Structure of the GDP domain of EF-Tu and location of the amino acids homologous to ras oncogene proteins. Science **230**:32–36.
24. **Kelly, R., S. M. Miller, and M. B. Kurtz.** 1988. One step gene disruption by cotransformation to isolate double auxotrophic mutants in *Candida albicans*. Mol. Gen. Genet. **214**:24–31.
25. **Kelly, R., S. M. Miller, M. B. Kurtz, and D. R. Kirsch.** 1987. Directed mutagenesis in *Candida albicans*: one-step gene disruption to isolate *ura3* mutants. Mol. Cell. Biol. **7**:199–207.
26. **Kimura, L. H., and N. N. Pearsall.** 1980. Relationship between germination of *Candida albicans* and increased adherence to human buccal epithelial cells. Infect. Immun. **28**:464–468.
27. **King, R., J. Lee, and A. Morris.** 1980. Adherence of *Candida albicans* and other *Candida* species to mucosal epithelial cells. Infect. Immun. **27**:667–674.
28. **Kirsch, D. R., M. H. Lai, and J. O'Sullivan.** 1988. Isolation of the gene for cytochrome P450 L1A1 (lanosterol 14α-demethylase) from *Candida albicans*. Gene **68**:229–237.
29. **Kurtz, M. B., M. W. Cortelyou, and D. R. Kirsch.** 1986. Integrative transformation of *Candida albicans*, using a cloned *Candida ADE2* gene. Mol. Cell. Biol. **6**:142–149.
30. **Kurtz, M. B., and J. Marrinan.** 1989. Isolation of *Hem3* mutants from *Candida albicans* by sequential gene disruption. Mol. Gen. Genet. **217**:47–52.
31. **Kurtz, M. B., D. R. Kirsch, and R. Kelly.** 1988. The molecular genetics of *Candida albicans*. Microbiol. Sci. **5**:58–63.
32. **Langford, C. J., and D. Gallwitz.** 1983. Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell **33**:519–527.
33. **Lechner, K., and A. Bock.** 1987. Cloning and nucleotide sequence of the gene for an archaebacterial protein synthesis elongation factor Tu. Mol. Gen. Genet. **208**:523–528.
34. **Lee, K. L., H. R. Buckley, and C. C. Campbell.** 1975. An amino acid liquid synthetic medium for the development of mycelial and yeast forms of *Candida albicans*. Sabouraudia **13**:148–153.
35. **Leer, R. J., M. M. C. Van Raamsdonk-Duin, W. H. Mager, and R. J. Planta.** 1985. Conserved sequences upstream of yeast ribosomal protein genes. Curr. Genet. **9**:273–277.
36. **Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker.** 1977. RNA molecular weight determinations of gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry **16**:4743–4751.
37. **Lehrer, N., E. Segal, R. Cihlar, and R. Calderone.** 1986. Pathogenesis of vaginal candidiasis: studies with a mutant which has reduced ability to adhere *in vitro*. J. Med. Vet. Mycol. **24**:127–131.
38. **Lenstra, J. A., A. V. Vliet, A. C. Arnberg, F. J. Van Hemert, and W. Möller.** 1986. Genes coding for the elongation factor EF-1α in *Artemia*. J. Biochem. **155**:475–483.
39. **Lewin, B.** 1983. Genes, p. 424–425. John Wiley & Sons, Inc., New York.
40. **Linz, J. E., L. M. Lira, and P. S. Sypherd.** 1986. The primary structure and the functional domains of an elongation factor-1α from *Mucor racemosus*. J. Biol. Chem. **261**:15022–15029.
41. **Linz, J. E., and P. S. Sypherd.** 1987. Expression of three genes for elongation factor 1α during morphogenesis of *Mucor racemosus*. Mol. Cell. Biol. **7**:1925–1932.
42. **Linz, J. E., and P. S. Sypherd.** 1987. Three genes for the elongation factor EF-1α in *Mucor racemosus*. Mol. Cell. Biol. **6**:593–600.
43. **Lu, X., and D. Werner.** 1989. The complete cDNA sequence of mouse elongation factor 1α (EF-1α) mRNA. Nucleic Acids Res. **17**:442.
44. **Lue, N. F., A. R. Buchman, and R. D. Kornberg.** 1989. Activation of yeast RNA polymerase II transcription by a thymidine-rich upstream element in vitro. Proc. Natl. Acad. Sci. USA **86**:486–490.
45. **Maisch, P., and R. Calderone.** 1980. Adherence of *Candida albicans* to a fibrin platelet matrix formed *in vitro*. Infect. Immun. **27**:650–656.
46. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 504. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
47. **McNeil, J. B.** 1988. Functional characterization of a pyrimidine-rich element in the 5′ noncoding region of the yeast iso-1-cytochrome *c* gene. Mol. Cell. Biol. **8**:1045–1054.
48. **Moore-Landecker, E.** 1982. Fundamentals of the Fungi, 2nd ed. Prentice-Hall, Inc., Englewood Cliffs, N.J.
49. **Nagashima, K., M. Kasai, S. Nagata, and Y. Kaziro.** 1986. Structure of the two genes coding for polypeptide chain elongation factor 1α (EF-1α) from *Saccharomyces cerevisiae*. Gene **45**:265–273.
50. **Nagata, S., K. Nagashima, Y. Tsunetsugu-Yokota, K. Fujimura, M. Miyazaki, and Y. Kaziro.** 1984. Polypeptide chain elongation factor 1α (EF-1α) from yeast: nucleotide sequence of one of the two genes for EF-1α from *Saccharomyces cerevisiae*. EMBO J. **3**:1825–1830.
51. **Pfeifer, K., K. Kim, S. Kogan, and L. Guarente.** 1989. Functional dissection and sequence of yeast HAP1 activator. Cell **56**:291–301.
52. **Pokalsky, A. R., W. R. Hiatt, N. Ridge, R. Rasmussen, C. M. Houck, and C. K. Shewmaker.** 1989. Isolation and characterization of a cDNA for tomato elongation factor 1α. Nucleic Acids Res. **17**:4661–4673.
53. **Rosset, R., and L. Gorini.** 1969. A ribosomal ambiguity mutation. J. Mol. Biol. **39**:95–112.
54. **Rotenberg, M. O., and J. L. Woolford.** 1986. Tripartite upstream promoter element essential for expression of *Saccharomyces* ribosomal protein genes. Mol. Cell. Biol. **6**:674–687.
55. **Sandbaken, M. G., and M. R. Culbertson.** 1988. Mutations in elongation factor EF-1α affect the frequency of frameshifting and amino acid misincorporation in *Saccharomyces cerevisiae*. Genetics **8**:923–934.
56. **Sanger, F., and A. R. Coulson.** 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94**:441–448.
57. **Seidler, L., M. Peter, F. Meissner, and M. Sprinzl.** 1987. Sequence and identification of the nucleotide binding site for the elongation factor Tu from Thermus thermophilus HB8. Nucleic Acids Res. **15**:9263–9277.
58. **Singer, S. C., C. A. Richards, R. Ferone, D. Benedict, and P. Ray.** 1989. Cloning, purification, and properties of *Candida albicans* thymidylate synthase. J. Bacteriol. **171**:1372–1378.
59. **Smith, R. A., H. S. Allaudeen, M. H. Whitman, Y. Koltin, and J. A. Gorman.** 1988. Isolation and characterization of a β-tubulin gene from *Candida albicans*. Gene **63**:53–63.
60. **Smith and Paress.** 1978. Genetic and biochemical characterization of kirromycin resistance mutations in *Bacillus subtilis*. J. Bacteriol. **135**:1107–1117.
61. **Sobel, J., G. Muller, and H. Buckley.** 1984. Critical role of germ tube formation in the pathogenesis of candidal vaginitis. Infect. Immun. **44**:576–580.
62. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**: 503–517.
63. **Struhl, K.** 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell **49**:295–297.
64. **Sundstrom, P., L. M. Lira, D. Choi, J. E. Linz, and P. S.**

Sypherd. 1987. Sequence analysis of the EF-1α gene family of Mucor racemosus. Nucleic Acids Res. **15**:9997–10006.
65. **Tapio, S., and C. G. Kurland.** 1982. Mutant EF-Tu increases missense error *in vitro*. Mol. Gen. Genet. **205**:186–188.
66. **Thompson, R. C.** 1988. EFTu provides an internal kinetic standard for translational accuracy. Trends Biochem. Sci. **13**: 91–93.
67. **Vignais, M. L., L. P. Woudt, G. M. Wassenaar, W. H. Mager, A. Sentenac, and R. J. Planta.** 1987. Specific binding of TUF factor to upstream activation sites of yeast ribosomal protein genes. EMBO J. **6**:1451–1457.
68. **Vijgenboom, E., L. Nilsson, and L. Bosch.** 1988. The elongation factor EF-Tu from *E. coli* binds to the upstream activator region of the TufA-TufB operon. Nucleic Acids Res. **16**:10183–10197.
69. **Woudt, L. P., W. H. Mager, R. T. M. Nieuwint, G. Wassenaar, A. C. van der Kuyl, J. J. Murre, M. F. M. Hoekman, P. G. M. Brockhoff, and R. J. Planta.** 1987. Analysis of upstream activation sites of yeast ribosomal protein genes. Nucleic Acids Res. **15**:6037–6048.
70. **Woudt, L. P., A. B. Smit, H. Mager, and R. J. Planta.** 1986. Conserved sequence elements upstream of the gene encoding yeast ribosomal protein L25 are involved in transcription activation. EMBO J. **5**:1037–1040.
71. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
72. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28**:563–573.