Structural and Phylogenetic Analysis of the Actin Gene from the Yeast \textit{Phaffia rhodozyma}

JAN WERY,$^*$ MIEKE J. M. DALDERUP,$^*$ JOSE TER LINDE,$^+$ TEUN BOEKHOUT$^§$ AND ALBERT J. J. VAN OOYEN$^*$$^‡$

$^*\text{Division of Industrial Microbiology, Department of Food Science, Wageningen Agricultural University,}$

$^+\text{PO Box 8129, 6700 EV Wageningen, The Netherlands,}$

$^\S\text{Gist-brocades NV, PO Box 1, 2600 MA Delft, The Netherlands}$

$^\P\text{Central Bureau voor Schimmelcultures, Yeast Division, Julianalaan 67, 2628 BC Delft, The Netherlands}$

Received 4 August 1995; accepted 1 December 1995

The gene coding for actin from \textit{Phaffia rhodozyma} was cloned and sequenced. The \textit{Phaffia} actin gene contains four intervening sequences and the predicted protein consists of 375 amino acids. The structural features of the \textit{Phaffia} actin introns were studied and compared with actin introns from seven fungi and yeasts with ascomycetous and basidiomycetous affinity. It was shown that the architecture of the \textit{Phaffia} introns most resembles that of the basidiomycete \textit{Filobasidiella neoformans} (perfect stage of \textit{Cryptococcus neoformans}), whereas least resemblance occurs with the ascomycetous yeasts. Based on the intron structure, the ascomycetous yeasts can be accommodated in one group in that their splice site sequences are very similar and show less homology with the other fungi investigated, including \textit{Phaffia}. It was demonstrated that the \textit{Phaffia} actin introns cannot be spliced in \textit{Saccharomyces cerevisiae}, which shows that the differences found in intron structure are significant. Alignment of the \textit{Phaffia} actin gene with the actin sequences from the yeasts and fungi investigated showed a high level of homology both on the DNA level and on the protein level. Based on these alignments \textit{Phaffia} showed highest homology with \textit{F. neoformans} and both organisms were accommodated in the same cluster. In addition, the actin gene comparisons also supported the distant relationship of \textit{Phaffia} with the ascomycetous yeasts. These results supported the usefulness of actin sequences for phylogenetic studies! The sequence presented here has been submitted to the EMBL data library under Accession Number X89898.

KEY WORDS — \textit{Phaffia rhodozyma}; actin gene; DNA sequencing; phylogenetic studies.

INTRODUCTION

\textit{Phaffia rhodozyma} is a red yeast that produces astaxanthin as the main carotenoid (Phaff \textit{et al.}, 1972). Astaxanthin is widely distributed in nature giving crustaceans, birds (flamingos), fish (salmon, trout) and many other organisms their distinctive orange-red colour through their diet.

\textit{Phaffia} is one of very few organisms known to synthesize astaxanthin. The growing economic importance of astaxanthin as a fish feed additive for the cultivation of salmon and trout has made \textit{Phaffia} increasingly attractive as a possible profitable source.

Most studies conducted on this yeast have been concerned with general physiological properties, related to taxonomy (Miller \textit{et al.}, 1976) and the physiology of astaxanthin production (Johnson and Lewis, 1979).

Based on classical ultrastructural and chemotaxonomic observations, such as the ability to synthesize carotenoids, the multilayered cell wall, the enteroblastic way of budding and a positive reaction to Diazonium Blue B, \textit{Phaffia} is accommodated in the basidiomycetous yeasts (Miller \textit{et al.}, 1976; Sugiyama \textit{et al.}, 1985).

Very little is known about the genetics of \textit{Phaffia} and until now no genes have been isolated. In order to gain more insight in this field, we have isolated and sequenced the \textit{Phaffia} actin gene.

Actin plays a crucial role in elementary eukaryotic cellular processes such as motility, cell growth regulation, cell differentiation and provides structural stability (Stossel, 1984; Pollard and Cooper, 1986; Pollard, 1990).
Apparently these functions have not allowed actins of different species from a broad range of phyla to diverge significantly, since the actin protein is highly conserved. This feature makes the actin gene useful for phylogenetic analysis, especially in combination with classical methods, which are mainly based on chemotaxonomic and ultrastructural studies. In this context actin gene analysis has been used for the grouping of organisms (Vandekerckhove and Weber, 1984; Fletcher et al., 1994).

In recent years molecular taxonomy of fungi (e.g. rRNA/rDNA comparisons) has been useful for phylogenetic studies (Hendriks et al., 1991, 1992; Berbee and Taylor, 1992a,b; Suh and Sugiyama, 1993; Wilmotte et al., 1993; Kurtzman, 1994). It was demonstrated that ascomycete and basidiomycete genera were well separated by comparing 18s and 26s rDNA sequences. Only a few fungal actin genes have been isolated so far and little is known about the validity of using these genes for phylogenetic studies among fungi.

This report describes the isolation and analysis of the Phaffia actin gene. As a first step to a better understanding of the genetics of Phaffia, the structure of the Phaffia actin gene, including the presence, location, size and sequence of introns, was studied and compared to the structure of actin genes from different fungi. In addition the expression of the intron containing part of the Phaffia actin gene in S. cerevisiae was studied.

Finally we have compared the actin exon sequences in order to validate the use of actin for phylogenetic studies among fungi.

MATERIALS AND METHODS

Strains and media

Escherichia coli strains DH5a and JM 109 were used for transformation and amplification of recombinant plasmids according to standard methods (Sambrook et al., 1989). Phaffia rhodozyma strain CBS 6938 and Saccharomyces cerevisiae strain 703 (ura3, trp1, his3) were used in various experiments. Phaffia was cultivated at 21°C in YepD medium containing 1% yeast extract, 2% bactopeptone, 2% glucose. DNA isolation from Phaffia was done as described by Davis et al. (1980) with modifications (Wery, in preparation).

S. cerevisiae was cultivated at 30°C in YepD medium. S. cerevisiae transformants were primarily cultivated in appropriately supplemented minimal medium (0.67% yeast nitrogen base, 2% glucose) and on YepD plates containing 50 μg/ml Geneticin (G418). Transformation of S. cerevisiae was performed according to the LiAc method of Schiestl and Gietz (1989). Total RNA was isolated from exponentially growing S. cerevisiae transformants mainly according to Lacy and Dickson (1981).

DNA methods

Plasmid pTZ18R (Pharmacia) was used as a starting plasmid for all cloning steps. Insert sequences originating from fractionated total Phaffia DNA were isolated from 0-7% agarose gels by agarase treatment according to the supplier (Boehringer). DNA fragments obtained by polymerase chain reaction (PCR) were purified from low melting point agarose using the Magic PCR Prep kit (Promega). DNA digestions and ligations were performed using enzymes purchased from BRL and applied according to the supplier’s recommendations.

Positive colonies were generally picked up by colony lifting and hybridization using the non-radioactive DIG DNA Labelling and Detection Kit (Boehringer). For DNA hybridizations, digested total Phaffia DNA was transferred to a nitro-cellulose filter and hybridization was performed according to standard protocols (Sambrook et al., 1989).

Polymerase chain reaction

Standard reactions were carried out in an automated thermal cycler (Perkin-Elmer). Conditions were: 5 min at 95°C, followed by 30 repeated cycles; 2 min at 94°C, 2 min at 45°C, 3 min at 72°C; ending with one cycle; 10 min at 72°C.

Oligonucleotides used in this study (inverse PCR):

1. 5'-CGCCATCTTCTATAACAATACC-3'
2. 5'-GCATCAAGGAGAAGCTCTGCTA-3'
3. 5'-GTAGTCTTCATGGCCAGAGCTCATCCTTTTGTTTGTTTCCGGG-3'
4. 5'-TGTATATGAGATATGGAGTCTTGGTATAG-3'

Oligonucleotides used for generation of the 0.2 kb part of the 5' end of the kanamycin resistance (KmR) gene:

1. 5'-CGCATCTTTCTATAACATACC-3'
2. 5'-GCATCAAGGAGAAGCTCTGCTA-3'

Oligonucleotides used for generation of the 0.7 kb alcohol dehydrogenase I (ADH1) promoter from S. cerevisiae (Bennetzen and Hall, 1982):

3. 5'-GTAGTCTTCATGGCCAGAGCTCATCCTTTTGTTTGTTTCCGGG-3'
4. 5'-TGTATATGAGATATGGAGTCTTGGTATAG-3'
Figure 1. (A) Restriction maps of plasmids pGB-Ph9 (Wery et al, in preparation), pSEP3 and pSEP4 (as described in Materials and Methods). act1=Phaffia actin sequences. act1P=actin gene promoter. rDNA=Phaffia ribosomal DNA. adh1P=S. cerevisiae alcohol dehydrogenase promoter. (B) Phaffia actin/ KmR fusion in pSEP3 and expression in S. cerevisiae.

5'.CTATCTCATATACAATGTGGATTGA
ACAAAGATG-3'

6'.GCCTGACTTCTGGCCAGCCACGAT
AGC-3'

Oligonucleotides used for the generation of the intron part of the Phaffia actin gene fused to the 5' portion of the KmR gene from pGB-Ph9:
7. 5'-CTATCTCATATAATGTACGTCGACATGCTTTC-3'
8. As oligonucleotide 6. Fusion PCR reactions were performed as described above, except that two DNA fragments with compatible ends were used as a template in equimolar amounts. (Note: restriction sites are underlined, overlapping sequences are bold.)

Synthesis of cDNA and amplification by PCR was performed essentially according to Kawasaki et al. (1988). Oligonucleotides used for cDNA synthesis and amplification of mRNA originating from pSEP3:
9. 5'-TCAGAACTCAAGCTTTACAATCAA
10. 5'-GGGAATTCGCTTGGACTGAGCCTC

Oligonucleotides used for cDNA synthesis and amplification of S. cerevisiae actin mRNA:
5'-TACTGAATTAACAATGGATTC-3'
5'-GAGCTTCATCACCAACGTA-3'

Construction of plasmids pSEP3 and pSEP4
A 1.4 kb EcoRI fragment containing the S. cerevisiae ARS1TRP1 sequence (Stinchcomb et al., 1979) was inserted in the ribosomal DNA portion of pGB-Ph9 (Wery et al., in preparation). An MscI fragment containing the Phaffia actin promoter, introns and a 5' portion of the KmR gene was deleted. Insertion of an MscI PCR product, in which the 5' portion of the KmR gene (generated using oligonucleotides 5 and 6) was directly fused to the S. cerevisiae ADH1 promoter (generated using oligonucleotides 3 and 4), yielded pSEP4 (Figure 1). PSEP3 was obtained by insertion of an MscI PCR fragment containing the in-frame fusion between a 0.9 kb intron containing part of the Phaffia actin gene coding for 83 amino acids and the 5' portion of the KmR gene (generated using oligonucleotides 7 and 8) driven by the ADH1 promoter (generated using oligonucleotides 3 and 4).

Phylogeny
Phylogenetic relationships were inferred from the coding sequences of the actin genes from Phaffia rhodozyma (this study), Filobasidiella neoformans (EMBL U10867), Thermomyces lanuginosus (EMBL X07463), Aspergillus nidulans (M22869), Schizosaccharomyces pombe (Y00447), Saccharomyces cerevisiae (J01310), Kluyveromyces lactis (M25826), Candida albicans (X16377) and

Absidia glauca (M64729) using neighbour-joining with Kimura 2-parameter distances (Kumar et al., 1993). Robustness of the resulting tree was tested by bootstrap analysis using 1000 replicates. Alternatively, the sequences were analysed using PAUP version 3.0s (Swofford, 1991).

RESULTS
Isolation of the Phaffia actin gene
Total Phaffia DNA was digested with several enzymes and separated by agarose gel electrophoresis. After blotting, the DNA was hybridized using a PCR-amplified 200 bp K. lactis actin DNA fragment. The Phaffia DNA only hybridized weakly as shown by the rather faint signals (Figure 2).

A 3.8 kb EcoRI-SalI Phaffia DNA fragment, hybridizing with the K. lactis actin probe, was cloned in pTZ18R. The resulting plasmid PGB-Ph1 was partly sequenced and based on sequence comparison with the K. lactis and S. cerevisiae actin genes (Deshler et al., 1989; Gallwitz and Sures, 1980). The 5' end of the gene including the promoter region was missing.

The restriction sites around the Phaffia actin gene were mapped by digestion of chromosomal DNA with several enzymes followed by hybridization with a 500 bp SalI-BamHI fragment at the 5' end of the Phaffia actin gene insert from pGB-Ph1 (not shown). A 3.1 kb hybridizing fragment was detected using XhoI. Since it was known that a 1.4 kb SalI-XhoI portion is located at the 5' end of
The actin insert in pGB-Ph1, it was concluded that the hybridizing band contained a 1.7 kb sequence flanking the 5' end of the *Phaffia* actin insert in pGB-Ph1.

This 1.7 kb sequence was isolated using the inverse PCR technique (Ochman et al., 1988). Chromosomal DNA from *Phaffia* was digested with *XhoI* and subsequently ligated. PCR was performed with the ligated material using oligonucleotides 1 and 2 (Materials and Methods), which were designed on the basis of known sequences 60 bp downstream of the *SalI* site and 60 bp upstream of the *XhoI* site. A PCR product with an expected length of 1.7 kb was obtained.

After digestion with *SalI* and *XhoI* the PCR fragment was cloned in the *SalI* site of pTZ18R, yielding pGB-Ph2. A 0.6 kb 3' portion of the *Phaffia* DNA insert in pGB-Ph2 was sequenced and found to contain the 5' end of the *Phaffia* actin gene including the putative ATG start codon. Thus the full-length *Phaffia* actin gene and promoter were obtained.

**Architecture of Phaffia actin intron and comparison with other fungal actin introns**

The 1808 bp *Phaffia* actin gene codes for 375 amino acids (Figure 3). The 5' part of the gene contains four introns, ranging in length from 71 to 338 bp. Sequence elements that play an important role in the splicing process are partly conserved.

In Table 1 the 5' splice sites, the branch sequences and the 3' splice sites are depicted. The

**Figure 3.** Nucleotide sequence of the *Phaffia* actin gene. The amino acid sequences of exons 1 to 5 are shown below their respective DNA sequences. The splice sequences are underlined.
Table 1. Consensus actin intron sequences.

<table>
<thead>
<tr>
<th>Intron</th>
<th>Size (bp)</th>
<th>Pos. 5'</th>
<th>5' Putative branch</th>
<th>3' Putative branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phaffia</td>
<td>1</td>
<td>338</td>
<td>CAT/GTACGT</td>
<td>TATTGAC</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>98</td>
<td>CTC/GTATGT</td>
<td>TTCTTAT</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>177</td>
<td>CCC/GTAAAGT</td>
<td>TGCTGTAGT</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>71</td>
<td>TGG/GTTCGT</td>
<td>TATTAAC</td>
</tr>
<tr>
<td>F. neoformans</td>
<td>1</td>
<td>52</td>
<td>AAG/GTACGT</td>
<td>AGCTAAC</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>45</td>
<td>CCC/GTAAAGT</td>
<td>AACCTAC</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>47</td>
<td>CGG/GTATGT</td>
<td>TATTAAC</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>52</td>
<td>TTT/GTAAAGT</td>
<td>ACCTAAC</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>52</td>
<td>GGA/GTAAAGT</td>
<td>CGCTAAC</td>
</tr>
<tr>
<td>T. lanuginosus</td>
<td>1</td>
<td>148</td>
<td>AAG/GTAAAGC</td>
<td>TGCTAACAT</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>157</td>
<td>TGG/GTATGT</td>
<td>AGCTGAC</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>146</td>
<td>TCC/GTAAAGT</td>
<td>TACTGAC</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>58</td>
<td>TGG/GTAAATT</td>
<td>CTCTAAC</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>57</td>
<td>ATG/GTATAGT</td>
<td>CACTAAC</td>
</tr>
<tr>
<td>A. nidulans</td>
<td>1</td>
<td>111</td>
<td>AAG/GTAAAGG</td>
<td>CGCTAACAT</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>232</td>
<td>TGG/GTATGT</td>
<td>TGCTGAC</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>194</td>
<td>TCC/GTAAAGT</td>
<td>TGCTAAC</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>59</td>
<td>TGG/GTAAAT</td>
<td>CGCTAAC</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>56</td>
<td>ATG/GTATGT</td>
<td>CACTAAC</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>54</td>
<td>1786</td>
<td>AAG/GTATAGA</td>
</tr>
<tr>
<td>S. bayanus</td>
<td>1</td>
<td>348</td>
<td>CTG/GTATGT</td>
<td>TACTAAC</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>304</td>
<td>CTG/GTATGT</td>
<td>TACTAAC</td>
</tr>
<tr>
<td>K. lactis</td>
<td>1</td>
<td>777</td>
<td>CTG/GTATGT</td>
<td>TACTAAC</td>
</tr>
<tr>
<td>C. albicans</td>
<td>1</td>
<td>656</td>
<td>GTC/GTATGT</td>
<td>TACTAAC</td>
</tr>
<tr>
<td>U1SnRNA*</td>
<td></td>
<td></td>
<td></td>
<td>GaCTAAC</td>
</tr>
<tr>
<td>U2SnRNA</td>
<td></td>
<td></td>
<td></td>
<td>PuCTPuAC</td>
</tr>
<tr>
<td>Consensus†</td>
<td></td>
<td></td>
<td></td>
<td>PyAG</td>
</tr>
</tbody>
</table>

Pos. = nucleotide position of the intron within the gene. 5' = 5' splice sequence. Branch = branch sequence. 3' = 3' splice sequence. Exon intron boundaries are indicated by a slash. The number between the branch sequence and the 3' splice sites indicates the distance in nucleotides.

*Jacob and Gallinaro (1989).
†Rambosek and Leach (1987).

5' splice site, consisting of nine nucleotides, includes three nucleotides of the exon adjacent to six nucleotides of the 5' part of the intron. Exon intron boundaries are indicated by a slash. If compared with the consensus 5' splice sequence CAG/GTAAAGT (Jacob and Gallinaro, 1989), the 5' splice sites of the Phaffia actin introns display five to seven complementary matches. In all four introns the first and the last GT in the intron part of the 5' splice site are conserved.

In fungal introns the sequence N PuCTPuAC has been proposed as the consensus branch sequence (Rambosek and Leach, 1987). In this sequence the last A is considered to be important in the splicing process in that it is involved in lariat formation by covalent binding to the first G of the intron (Domdey et al., 1984; Padgett et al., 1984). Based on this criterion putative branch sites can be proposed within all four Phaffia actin introns. In these branch sequences, four to six nucleotides match with the consensus sequence. In all four branch sites, the first and fourth T in the sequence as well as the important last A are conserved.

The 3' splice site (PyAG) is completely conserved in all four Phaffia actin introns. The distance between the putative branch sites and the 3' splice site ranges from 7 to 67 bp. In mammalian introns this spacing varies from 10 to 40 bp (Green, 1986; Padgett et al., 1986), in S. cerevisiae from 3 to 134 bp (Gallwitz et al., 1987; Woolford,
1989) and in most of the intron-containing non-actin genes in *S. pombe* from 3 to 16 bp (Prabhala *et al.*, 1992). Similar to mammalian and *S. pombe* introns (Green, 1986; Padgett *et al.*, 1986; Prabhala *et al.*, 1992), Py stretches can be assigned in this region in the first two introns of the *Phaffia* actin gene. Considering the complete intron sequences, except for the third intron, they all show high Py content (64–68%).

The *Phaffia* actin intron architecture was compared to the actin introns of various ascomycete and basidiomycete fungi, including *Saccharomyces bayanus*, *S. cerevisiae*, *K. lactis*, *C. albicans*, *F. neoformans*, *T. lanuginosus* and *A. nidulans*. The actin sequences were collected from the EMBL/GenBank database. Based on this comparison the investigated species are divided into two groups. Group 1 accommodates the filamentous fungi *T. lanuginosus* and *A. nidulans*, and the basidiomycetous yeasts *F. neoformans* (the perfect stage of *Cryptococcus neoformans*) and *P. rhodozyma*.

Accommodated in group 2 are the ascomycetous yeasts *S. cerevisiae*, *S. bayanus*, *K. lactis* and *C. albicans*. In group 2 the actin genes contain only one intron, whereas group 1 actins contain four to six introns. There is also a size difference. Group 2 introns, ranging in size from 304 to 777 bp, are on average significantly longer than the 52–337 bp group 1 introns.

Furthermore, large differences exist between the consensus sequence elements involved in splicing. In group 2 all 5′ splice sites conform to the consensus sequence G\_G/TG/GTAGT. The branch sequences in group 2 introns are all in accordance with the sequence TACTAAC. Finally, high homology also exists between the 3′ splice sites of group 2 members. They all conform to PyAG/APuG.

In group 1 little of the homology displayed in group 2 remains. The exon part of the 5′ splice sequence shows no consensus at all, whilst in the intron part only the first GT is maintained throughout the group. Other bases in the 5′ splice site are more or less variable. The branch sequences (TACTAAC box) in group 1 introns also show a high degree of variability with only a well-conserved second T and last A (branch point) and a Py as a last base. The intron part of the 3′ splice site is also well conserved in group 1 introns; however, opposite to group 2 the exon part does not comply with any consensus.

At the splice site consensus level, *P. rhodozyma* and *F. neoformans* share a number of similarities that do not exist among the other members of group 1.

*Phaffia* intron 3 resembles *F. neoformans* intron 2 in that they both contain CCC/GTAAGT as the 5′ splice site. In addition the sequences GTACGT and GTATGT occur in both organisms as the intron part of the 5′ splice site. Furthermore, *Phaffia* intron 4 and *F. neoformans* intron 3 contain identical branch sites (TATTAAC) and 3′ splice sites (CAG/IGA).

*Phaffia* actin introns cannot be spliced in *S. cerevisiae*

From the foregoing it can be concluded that the *Phaffia* actin introns differ from the *S. cerevisiae* actin intron. We questioned to what extent these differences in intron architecture have an influence on splicing efficiency and examined whether the *Phaffia* actin introns could be correctly spliced by *S. cerevisiae*.

Obviously at least some impairment of the splicing efficiency could be anticipated. For example, the branch site sequences of the *Phaffia* actin introns show little similarity with the *S. cerevisiae* TACCTAAC box. In addition the intron part of the 5′ splice sites differs up to 2 bp.

However, it can be argued that there also exist similarities in the intron architecture of both organisms. First, like in *S. cerevisiae*, all *Phaffia* introns are located near the 5′ end of the actin gene, opposite to *A. nidulans* and *F. neoformans*. It was shown previously by Yoshimatsu and Nagawa (1994) that a 5′ end positioning of introns is important for correct splicing in *S. cerevisiae*. Second, the intron part of the 3′ splice site is highly conserved, whilst the maximum distance between the branch site and the 3′ splice site in *Phaffia* is 67 bp. It was shown previously that a 66 bp extension of this spacer in the *S. cerevisiae* actin gene to 106 bp still allowed relatively efficient splicing (Cellini *et al.*, 1986).

Finally, only the first *Phaffia* actin intron exceeds the size of the actin intron in *S. cerevisiae* marginally, while the other *Phaffia* actin introns are smaller.

To investigate the possible splicing of the introns by *S. cerevisiae*, we have constructed plasmid pSEP3 and pSEP4 (Figure 1). In pSEP3 the Km\(^\text{R}\) gene is used as a reporter gene for the correct splicing of *Phaffia* actin introns by *S. cerevisiae*. It is known that this transposon Tn5 encoded gene confers kanamycin resistance to *E. coli* and G418
Plasmid pSEP3 contains the ARS1 sequence and the TRP1 gene of *S. cerevisiae* (Stinchcomb *et al.*, 1979). It also harbours the KmR gene fused in-frame to an 83 amino acids coding intron containing part of the *Phaffia* actin gene under control of the strong *S. cerevisiae* ADH1 promoter. Plasmid pSEP4 differs from pSEP3 in that the KmR coding sequence is directly fused to the ADH1 promoter (Figure 1B).

As pSEP3 and pSEP4 were constructed using PCR techniques, three of each plasmid, obtained from independent *E. coli* clones, were transformed to *S. cerevisiae*, giving rise to the same results.

Transformants were initially screened for a TRP 1 phenotype. Transformant colonies were subsequently resuspended and dilutions containing 100–1000 cells were plated on YepD plates containing G418. Upon transformation with pSEP4, transformants were observed that were resistant to high G418 concentrations exceeding 1 mg/ml. However, pSEP3 transformants did not show any resistance at 50 µg/ml, a concentration at which adaptation is observed in non-transformed cells. This result indicated that the *Phaffia* actin introns in pSEP3 were not spliced efficiently. However, the event of partial splicing could not be ruled out, since some *Phaffia* actin introns share more homologies with the *S. cerevisiae* actin intron than others. Therefore it was decided to analyse the RNA of pSEP3 transformants by means of reverse transcriptase-PCR (Kawasaki *et al.*, 1988). Copy DNA prepared from total RNA from a pSEP3 transformant was amplified by PCR using oligo primers that were designed on regions at both sides of the intron containing part of pSEP3. No splicing products could be detected; only a 0·9 kb PCR product representing amplified unspliced cDNA and contaminating plasmid DNA was visible (Figure 4, lane f).

As a control on the procedure, PCR was also performed on the cDNA using oligonucleotides that were designed on sequences of the *S. cerevisiae* actin gene at both sides of the intron. Besides a 0·5 kb DNA fragment originating from contaminating chromosomal DNA and unspliced RNA, a 180 bp DNA fragment was also obtained, which is exactly the size of correctly spliced actin pre-mRNA (lane b). From these results it can be concluded that none of the *Phaffia* actin introns can be spliced in *S. cerevisiae*.

**Alignment of actin sequences shows a close phylogenetic relationship between Phaffia and Cryptococcus**

It was investigated if phylogenetic studies based on actin homology were consistent with the phylogenetic background of the fungi involved.

Therefore the actin sequences of *S. bayanus*, *S. cerevisiae*, *K. lactis*, *C. albicans*, *F. neoformans*, *T. lanuginosus*, *A. nidulans*, as well as *Sz. pombe* and the zygomycete *Absidia glauca*, which were not included in the intron architecture comparisons, were collected from the EMBL/GenBank database.

After removal of non-coding stretches (e.g. introns) the nucleotide sequences were aligned and phylogenetic relationships were inferred using neighbour-joining with Kimura 2-parameter distances (Kumar *et al.*, 1993). Robustness of the resulting tree was tested by bootstrap analysis using 1000 replicates (Figure 5). The actin gene from the zygomycete *A. glauca* was included as an outgroup member.

The nucleotide sequence of the actin gene of *Phaffia* was found to be most similar to that of the other basidiomycetous yeast, *F. neoformans*. These
Figure 5. Phylogenetic tree using nucleotides of the coding regions of the actin gene of some ascomycetous and basidiomycetous fungi, including some yeasts. The tree was generated using neighbour-joining with Kimura 2-parameter distance (Kumar et al., 1993). Bootstrap values are indicated on the branches and the scale of the branch lengths is indicated by the bar.

Two species formed a well-supported terminal branch (bootstrap value 95%). Two filamentous ascomycetous species, *T. lanuginosus* and *A. nidulans*, forming a sister group with the aforementioned basidiomycetes, are another well-supported terminal branch. The ascomycetous budding yeasts *S. bayanus*, *S. cerevisiae*, *K. lactis* and the anamorph *C. albicans* formed a well-supported cluster, with the fission yeast *S. pombe* only remotely related. Essentially the same phylogenetic relationships were obtained using parsimony methods (data not shown).

The amino acid sequences of the protein resulted in the same phylogenetic tree (data not shown). The actin protein of *Phaffia* was found to be most similar to that of *F. neoformans*. Surprisingly, the actin of *S. pombe* displayed a very high homology to the *Phaffia* and *F. neoformans* actin proteins.

DISCUSSION

In this report the isolation, cloning and analysis of the actin gene from *Phaffia rhodozyma* is described. This is the first *Phaffia* gene to be isolated and its architecture, involving the presence of introns and their structure, was shown to be quite different from actin genes studied so far from genetically well-studied yeasts such as *K. lactis*, *S. pombe* and *S. cerevisiae*. Since the phylogenetic relationships of these organisms are in agreement with the large differences in actin gene architecture, it was investigated whether this trend was also present if more actin genes from various fungi were analysed and compared.

A comparison made on the level of splice site consensus indeed supported the differences that exist between the species investigated. It was shown that the ascomycetous yeasts are clearly separated as a group from the filamentous ascomycetes and the basidiomycetous yeasts including *Phaffia* and *F. neoformans*. Moreover, the common basidiomycetous background of the latter two taxa was demonstrated by a number of similarities in splice site consensi, that were not shared with the other species.

The large phylogenetic distance between the ascomycetous yeast *S. cerevisiae* and the basidiomycete *Phaffia*, as confirmed by actin intron comparison, was emphasized by the fact that the introns of the latter were not spliced out by the first. Splicing of the *K. lactis* actin intron and the Candida maltosa FDH1 intron in *S. cerevisiae* have been studied (Deshler et al., 1989; Sasnaukas et al., 1992). Despite the fact that these organisms belong to different genera, the splice sequence element homologies were sufficient for efficient splicing. We demonstrated that none of the *Phaffia* actin introns are spliced in *S. cerevisiae*. Apparently the splice site sequences, in particular the branch sites, diverge too much. In this context it is noteworthy to mention that it was shown previously that the introns of the xylanase gene of the basidiomycetous yeast *Cryptococcus albidos* were not spliced in *S. cerevisiae* (Moreau et al., 1992). The putative branch sites in the introns of this gene (Boucher et al., 1988) also differ significantly from the *S. cerevisiae* TACTAAC consensus.

The results obtained from the actin exon alignments are in accordance with earlier findings based on rRNA/rDNA sequencing studies. It has been reported that molecular phylogenetic studies using 18S rRNA sequences divide basidiomycete and ascomycete taxa (Hendriks et al., 1991; Suh and Sugiyama, 1993; Wilmutte et al., 1993). This is supported by our actin gene comparisons. In addition, in our phylogenetic tree the ascomycetous yeasts, except *S. pombe*, form a cluster distinct from the filamentous ascomycetous fungi.

Our study suggests a distant relationship between *S. pombe* and the ascomycetous yeasts investigated. Partial and complete 18S rRNA
studies demonstrated that Sz. pombe is only remotely related to budding ascomycetous yeasts (Kurtzman, 1989; Sogin et al., 1989; Eriksson et al., 1993).

In conclusion, the results presented in this report show that Phaffia differs greatly from the ascomycetous budding yeasts, not only based on classical taxonomic features but also on the level of actin gene architecture. This knowledge is very useful in future genetic research on this yeast. Furthermore, it was shown that actin gene comparisons among fungi are useful for phylogenetic studies.

ACKNOWLEDGEMENTS

This work was supported by Gist-brocades. We thank Jacobus Bos (Gist-brocades) for performing primer synthesis and DNA sequencing.

REFERENCES


Kumar, S., Tamura, K. and Nei, M. (1993). MEGA: Molecular evolutionary genetics analysis, version 1.01. The Pennsylvania State University, University Park, PA 16802, USA.


