Genes in S. cerevisiae Encoding Proteins with Domains Homologous to the Mammalian ras Proteins

S. Powers,* T. Kataoka,* O. Fasano,* M. Goldfarb,* J. Strathern,* J. Broach,† and M. Wigler*.

* Cold Spring Harbor Laboratory
† Department of Microbiology
Cold Spring Harbor, New York 11724

Summary

The ras genes, which were first identified by their presence in RNA tumor viruses and which belong to a highly conserved gene family in vertebrates, have two close homologs in yeast, detectable by Southern blotting. We have cloned both genes (RAS1 and RAS2) from plasmid libraries and determined the complete nucleotide sequence of their coding regions. They encode proteins with nearly 90% homology to the first 80 positions of the mammalian ras proteins, and nearly 50% homology to the next 80 amino acids. Yeast RAS1 and RAS2 proteins are more homologous to each other, with about 90% homology for the first 180 positions. After this, at nearly the same position that the mammalian ras proteins begin to diverge from each other, the two yeast ras proteins diverge radically. The yeast ras proteins, like the proteins encoded by the mammalian genes, terminate with the sequence cysAAK, where A is an aliphatic amino acid. Thus the yeast ras proteins have the same overall structure and interrelationship as the family of mammalian ras proteins. The domains of divergence may correspond to functional domains of the ras proteins. Monoclonal antibody directed against mammalian ras proteins immunoprecipitates protein in yeast cells containing high copy numbers of the yeast RAS2 gene.

Introduction

The ras genes were first identified as the oncogenes of Harvey (v-H-ras) and Kirsten (v-K-ras) sarcoma viruses (Ellis et al., 1981). Certain tumor cells contain structurally mutated ras genes, which are capable of the tumorigenic transformation of NIH3T3 cells upon DNA-mediated gene transfer (Tabin et al., 1982; Reddy et al., 1982; Taparowsky et al., 1982; Yuasa et al., 1983; Taparowsky et al., 1983; Shimizu et al., 1983a; Capon et al., 1983b). At least three ras genes exist in mammals, the H-, K- and N-ras, which encode highly related proteins of 188-189 amino acids (Capon et al., 1983a; Shimizu et al., 1983; McGrath et al., 1983; Taparowsky et al., 1983). Homologous genes have been identified in Drosophila (Shilo and Weinberg, 1981). Since little is known about the normal biochemical or physiological function of the ras proteins, or the consequence of mutations upon these functions, we sought evidence for the existence of homologous genes in yeast Saccharomyces cerevisiae, a simple eucaryotic organism that can be subjected to powerful genetic analysis. Recently, Defeo-Jones et al. (1983) published the predicted partial amino acid sequence of a yeast gene that has considerable homology to mammalian ras proteins. Gallwitz et al. (1983) have also recently published the complete nucleotide and predicted amino acid sequence of another yeast gene, called YP2. The YP2-encoded protein has much weaker yet significant homology to the mammalian ras proteins. We demonstrate here two yeast genes that encode proteins with amino acid sequences highly related to the mammalian ras proteins. The gene that Defeo-Jones et al. have called c-ras<sup>−</sup>1 corresponds to the gene we have called yeast RAS1.

We present here the complete nucleotide and predicted amino acid sequence of this gene and that of a highly related companion gene, called yeast RAS2. We have tentatively identified the yeast RAS2 protein by immunoprecipitation with an anti-mammalian ras protein monoclonal antibody from extracts of yeast cells harboring RAS2 on a high copy number plasmid. Comparison of the encoded amino acid sequences of the known ras genes suggests the existence of functional domains of the encoded proteins.

Results

Hybridization of Yeast DNA with Mammalian ras Probes

DNA was prepared from yeast, digested with various restriction endonucleases, electrophoresed in agarose gels in triplicate, transferred to nitrocellulose filter paper, and hybridized according to the method of Southern using three different nick-translated probes: DNA fragments from the viral H-ras, the viral K-ras, and the human N-ras genes. Two yeast DNA restriction fragments with strong homology to all three probes were observed. Additional DNA fragments displayed a weak hybridization signal. Results with the K-ras probe are shown in Figure 1a.

Cloning Yeast Homologs to Mammalian ras Genes

The complete yeast S. cerevisiae genome has been cloned into bacterial/yeast shuttle vectors as Sau 3A partial digests (Broach et al., 1979; Nasmyth and Reed, 1980). This library was screened on nitrocellulose filters using as probes fragments of the mammalian ras genes. A number of candidate plasmid clones were isolated and analyzed by restriction endonuclease digestion and Southern blotting. Each of these plasmids contained one of two non-overlapping regions of DNA. The composite restriction endonuclease maps of these regions are shown in Figure 2. We call the homologous genes in yeast RAS1 and RAS2. The restriction endonuclease sites present on these genes identify them as the two major bands seen on Southern blot analysis. We verified this for yeast RAS2 by using it as a probe in Southern hybridization with nitrocel-
Figure 1. blot Hybridization Analysis of the Yeast Genome with ras Probes.
Yeast DNA from the DC5 haploid strain was digested with the indicated
restriction endonucleases, electrophoresed in 1.0% agarose, transferred to
nitrocellulose, and hybridized with 32P-labeled probes. (A) Nitrocellulose
filter hybridized with the 1.0 kb HinII fragment of the viral K-ras gene nick-
translated to 1.0 x 10^6 cpm per microgram. Exposure time was 24 h. (B)
The identical filter used above was rehybridized with the 3.0 kb Eco RI-
Hind III fragment of yeast RAS2, nick-translated to 1.0 x 10^8 cpm per
microgram. Exposure time was also 24 hr. Restriction endonuclease diges-
tions were: 1) Eco RI-Bam HI; 2) Eco RI; 3) Eco RI-Hind III; 4) Hind III; 5)
Bam HI; 6) Bam HI-Hind III; 7) Bam HI-Sal I; 8) Sal I. Numbers on the right
are size markers in kb.

DNA Sequence Analysis of Yeast ras Genes
To determine the precise relatedness of the yeast and
mammalian ras genes, we sequenced the yeast genes as
shown in Figure 2. The nucleotide sequences and pre-
dicted amino acid sequences are shown in Figure 3. There
is an open reading frame encoding exactly 309 amino
acids for yeast RAS1 and exactly 322 amino acids for
yeast RAS2. The initial methionine shown is the first me-
thionine in the open reading frame. The predicted amino
acid sequences show striking homology to each other,
through position 180. After this they diverge radically, until
homology is observed again for the last eight amino acids.

For purposes of comparison to the yeast ras proteins,
the amino acid sequence of the human H-ras protein is
also shown in Figure 3. Homology between the mammalian
and yeast ras proteins begins at position 10 for the yeast
and position 3 for the human H-ras protein. For the next
80 amino acids, there is nearly 90% homology between
any pair of the three. Patchy homology to H-ras is observed
thereafter. Homology to H-ras is increased if we introduce
an insertion at position 126 of the human H-ras protein.
Significantly, the yeast RAS1 and RAS2 proteins show
homology to positions 12, 13, 59, 61, and 63 of the H-ras
protein. It is amino acid substitutions at these positions
that activate the transforming potential of the mammalian
ras proteins (Tabin et al., 1982; Reddy et al., 1982; Tapa-
rowsky et al., 1982, 1983; Yuasa et al., 1983; Fasano et
al., submitted). By contrast, the YP2 protein of Gallwitz et
al. (1983) is not homologous to positions 12 or 63 of the
mammalian ras proteins.

To visualize better the relation of the three proteins, we
have plotted in Figure 4 three cumulative difference func-
tions, D(i), where D(i) = D(i-1) if two proteins under
comparison are identical at the equivalent i^th position; D(i)
= D(i-1) + 1/2 if there is a conservative amino acid
change (glycine and alanine; glutamic and aspartic acid;
lysine and arginine; or leucine, valine, and isoleucine); and
D(i) = D(i-1) + 1 if there is no homology at this position
or a frame-shift. For further comparison, we have plotted
a modified difference function, D'(i), of the three human
ras genes. D'(i) = D'(i-1) + 1 when there is no amino
acid consensus at the i^th position. D'(i) = D'(i-1) other-
wise. For ease of comparison, the D'(i) function is appro-
Figure 3. Yeast ras Sequence Comparisons

Shown at left are the nucleotide and predicted amino acid sequences of the yeast RAS1 and RAS2 genes. For comparison, the predicted amino acid sequence of the human H-ras gene is shown.

The numbers in the right margin indicate the position of the last amino acid in each respective line. Positions are underlined when identical amino acids are encoded and dotted when similar amino acids are encoded (see text). Gaps in nucleotide and amino acid sequence reflect attempts to align better the respective proteins. Asterisks in the H-ras amino acid sequence indicate positions where amino acid substitutions are known to activate the transforming potential of the H-ras protein.
Figure 4. Cumulative Difference Functions of the Various ras Genes
The cumulative difference functions $D(i)$ are plotted for (A) H-, K- and N-ras; (B) RAS1 vs RAS2; (C) H-ma vs RAS2; and (D) H-ma vs RAS1. For ease of visual comparison, the initial values, $D(0)$, are set to 0.0 for (A); 10.0 for (B); 20.0 for (C); and 30.0 for (D). Tick marks on abscissa (amino acid position, yeast ras coordinates) and on ordinate (cumulative difference values) are in units of 10. Functions are plotted up to position 197 in the yeast coordinates. See text for an explanation of the difference functions.

Generally frame-shifted to align with yeast ras sequences. In this representation, regions of amino acid homology are indicated by plateaus, regions of divergence by increasing slopes. At the amino acid level, yeast RAS1 and yeast RAS2 are equally diverged from H-ras. Although the data is not represented, the yeast ras proteins show no more homology to H-ras than to any of the other known mammalian ras proteins, RAS1 and RAS2 are clearly more closely related to each other than they are to mammalian ras. Our data also indicate that the yeast ras proteins diverge from the human ras proteins in roughly the same regions as the human ras proteins have begun to diverge from each other. Significantly, the yeast ras proteins diverge radically from each other in the domain corresponding to the C-terminal variable domain of the mammalian ras proteins (see below).

Immunoprecipitation of Yeast ras2
The extraordinary conservation of the yeast and human ras genes prompted us to attempt immunoprecipitation of the yeast products with a monoclonal antibody (Y13-259) raised initially to the Harvey sarcoma virus ras protein, but which has broad reactivity with mammalian ras proteins (Furth et al., 1982). For this purpose we performed immunoprecipitations and mock immunoprecipitations on 3S-methionine-labeled extracts from yeast cells containing high copy number, autonomously replicating plasmids containing the yeast ras genes (Figure 5). Elevated amounts of an immunoprecipitable 42 kd protein are seen in yeast cells harboring high copy numbers of the plasmid YEPRAS2-1 containing RAS2 (Figure 2). A lower molecular weight (30 kd) protein is also seen. Since the molecular weight of yeast RAS2, calculated from DNA sequence analysis, is 35 kd, these immunoprecipitable proteins may have undergone post-translational modifications. These results suggest that we have cloned a complete copy of the RAS2 gene and that it is expressed. We have not observed detectable amounts of immunoprecipitable protein in yeast cells with high copy numbers of the cloned yeast RAS1, either because RAS1 protein does not cross-react with monoclonal Y13-259 or because YEPRAS1-2 (see Figure 2) does not contain the complete RAS1 gene, or because the RAS1 gene is not expressed. See Figure 5 for more experimental details.

Discussion
We have demonstrated the presence in yeast of two genes with striking homology to the mammalian ras genes. Our results are in agreement with those of Defeo-Jones et al. (1983), who also found two genes in yeast closely homologous to mammalian ras genes. Our predicted amino acid sequence for RAS1 is in accord with the partial amino acid sequence predicted for the gene they called c-ras<sup>nc</sup>-1. Our data indicates that the RAS1 and RAS2 proteins are equidistant from the three known mammalian ras proteins but are more closely related to each other than to their mammalian homologs. A third gene, YP2, has been described in Saccharomyces cerevisiae by Gallwitz et al. (1983), which can encode a protein with significant but much weaker homology to the mammalian ras proteins.
Our data indicate that the protein encoded by YP2 is also quite diverged from both yeast ras proteins. Little is known about the functional domains of the ras proteins. Nevertheless, we can clearly speak of domains of divergence. The first domain comprises the first 80 or so N-terminal amino acids. This domain is the most highly conserved. The three human ras genes, H-, K-, and N-ras, encode proteins with identical amino acids sequence in this region. It is within this region that homology between mammalian and yeast ras proteins is highest. The second domain comprises the next 80 amino acid positions. In this domain the three human ras proteins have begun to diverge from each other, with about 85% homology between any pair (Taparowsky et al., 1983). The yeast ras proteins are also highly homologous to each other in this second domain, but are not as homologous as in the previous domain. We observe only patchy homology between yeast and mammalian ras proteins here. The third domain we call the variable domain. In mammalian ras it is short, comprising 15 amino acids near the C terminus. In mammals, this domain varies radically among different members of the ras gene family, but it is highly conserved for the same member in different species and is therefore under evolutionary constraint (Shimizu et al., 1983a; Taparowsky et al., 1983). In yeast ras proteins we may also speak of a C-terminal variable domain since we can observe no clear homology between yeast RAS1 and yeast RAS2 proteins in this region. However, the variable domain in yeast ras proteins is much larger than the corresponding domain in mammalian ras proteins. Finally, at the very carboxy terminus of ras proteins there is again a small conserved domain. As we noted previously, all the mammalian ras proteins terminate with the sequence cySLAX, where A is an aliphatic amino acid and X is the terminal amino acid (Shimizu et al., 1983a; Taparowsky et al., 1983). Interestingly, both the yeast ras proteins have this terminal peptide. However, the yeast RAS1 and RAS2 proteins contain a larger common terminus: seven of the last eight amino acids are identical.

We think these domains of divergence correspond to functional domains of the ras proteins as well. We suggest that the N-terminal domain is the effector region of the ras proteins, involved in interactions of a catalytic or regulatory nature that have been conserved in evolution. It is in this region where certain amino acid substitutions can activate the transforming potential of the mammalian ras proteins. On the other hand, we believe the C-terminal variable domain contains the determinants of physiological specificity. Through this region the ras proteins may receive their normal physiologic signals, which are then transduced or mediated to the N-terminal domain by way of the intervening, semiconserved domain. In this view of things, the yeast RAS1 and RAS2 proteins would have the same effector functions, but would carry out these functions in response to different stimuli. We have demonstrated that the cloned yeast RAS2 gene expresses a protein that is immunoprecipitable with a monoclonal antibody raised against the mammalian H-ras protein. We are currently investigating whether overexpression or disruption of either of the yeast ras genes results in a detectable phenotype. Study of the yeast ras genes may greatly accelerate our understanding of the normal and transforming mammalian ras genes at three levels. First of all, the high degree of conservation at the N terminus suggests to us that the biochemical effector function of the ras genes may have been conserved in evolution. (For example, yeast and mammalian ras genes may be regulatory components of a homologous catalytic system.) This hypothesis can be put to a rigorous test by examining the function of mammalian/yeast ras chimeric genes. Second, the yeast ras proteins are homologous to the mammalian proteins about amino acid positions where amino acid substitutions lead to activation of the transforming potential of the mammalian protein. By examining the consequences of similar amino acid substitutions on the function of the yeast ras proteins, we may gain valuable insights into the molecular mechanism of ras activation. Third, what we learn about the physiologic function of the yeast ras proteins may provide tantalizing clues to the physiologic role of the ras proteins in mammalian cells.

**Experimental Procedures**

**Yeast Strains, Media, and Transformation**

General procedures for genetic manipulation of yeast were performed as described by Mortimer and Hawthorne (1969). DNA from strain DCS (MATa his3Δ1 leu2Δ2) was used for Southern blot analysis. This strain was also used for transformations with the high copy number plasmid clones of RAS1 and RAS2 (see below). Yeast transformations were done according to the method of Beggs (1978). Cells were grown either in rich medium (2% Bacto-peptone, 1% yeast extract, and 2% glucose) or synthetic medium (0.7% yeast nitrogen base without amino acids [Difco], supplemented with appropriate amino acids and nucleic acid bases, and 2% glucose). Synthetic media was used for both 6-methionine labeling and for selection and maintenance of transformants.

**Nomenclature**

Consonant with standard yeast nomenclature we have designated the ras homologous genes in yeast as RAS1 and RAS2. These correspond respectively to c-ras+1 and c-ras+2 of Diefenboke et al. (1983).

**Southern Analysis**

Yeast DNA was prepared essentially as described (Struhl et al., 1979). DNAs were digested with restriction endonucleases (suppliers New England Biolabs or Bethesda Research), and 5 μg was loaded onto agarose gels for electrophoresis and blotted as described (Shimizu et al., 1983a). Low-stringency hybridizations were in aqueous 6x SCC at 50°C with a final blot wash in 2X SCC at 50°C. The ras probes used were the 2.2 kb Bam HI-Eco RI fragment of the viral K-ras plasmid clone pHK11 (Ellis et al., 1981), the 1.0 kb Hinc II fragment of the viral K-ras plasmid clone pHK2-2 (Ellis et al., 1981), and the 0.5 kb Eco RV fragment of the N-ras cDNA plasmid clone pKBE1 (Taparowsky et al., 1983).

**Screening Plasmid Libraries**

The Grunstein and Hogness procedure (1975) was used to screen a genomic library that had been constructed in the plasmid vector YEPl3 from yeast DNA partially digested with Sau 3A restriction endonuclease (Broach et al., 1979; Nasmyth and Reed, 1980). Low-stringency hybridizations were performed as described above. For our initial screening we used a 0.6 kb Pvu II-Sma I fragment from the H-ras cDNA plasmid clone RSE (Perez et al., 1983). 5x cDNAs that gave reactive strong signals on duplicate filters were purified and further analyzed by restriction endonuclease-
ase digestion and Southern blotting. Three of these plasmid clones were distinct and contained in common a 3.6 kb Eco RI fragment that hybridized to mammalian ras. Next, the library was screened with the 0.5 kb Nco I-Sal I fragment of the N-ras cDNA plasmid clone pDPl. We obtained three candidates, two of which contained a 1.7 kb Hind III fragment that hybridized to mammalian ras. See Figure 2 for more experimental details.

**Immunoprecipitations**

The plasmid library clone YEpRAS1-2 and YEpRAS2-1 (see Figure 2) were used to transform DC 5 to obtain yeast with high copy number of either RAS1 or RAS2 genes, respectively. Extracts of these transformants, along with DC 5 transformed with YEpD1 (parental plasmid without any insert; Breach et al., 1979) were prepared from genetically growing cultures (10^7 total cells per extract) labeled with 250 μCi 35S-methionine (Amersham) for 90 min. The cells were lysed in 200 μl PBS containing 1% Triton X-100, 0.6% deoxycholate, 1 mM PMSF, and 0.1 mg/ml aprotinin (Sigma) by vortexing with glass beads on ice. This crude extract was clarified, immunoprecipitated with monoclonal antibody Y13-259 (Furth et al., 1982), and analyzed by SDS-PAGE electrophoresis as described (Shimizu et al., 1983b).

**Sequencing**

Restriction endonuclease fragments were separated by gel electrophoresis and sequenced by the method of Maxam and Gilbert (1980) after 3’-OH end-labelling with E. coli polynucleotide large fragment (Beresnes Research Labs) or 5’-OH-labelling with T4 polynucleotide kinase (Miles Labs). Confirmatory sequence data was obtained by the dideoxy method of Sanger et al. (1977) as modified by Biggin et al. (1983). Restriction endonuclease fragments were cloned into either M13 mp8 or M13 mp9 (Messing and Vieira, 1982). See Figure 2 for a detailed description.

**Acknowledgments**

This work was supported by grants from the National Institutes of Health and American Business for Cancer Research Foundation. S. P. is a postdoctoral fellow of the Leukemia Society of America; T. K. is on leave from the Department of Genetics, Osaka University Medical School; J. B. is an Established Investigator of the American Heart Association; O. F. is on leave from Istituto di Chimica Biologica, II Facolta di Medicina, Universita di Napoli, Italy.

The costs of publication of this article were defrayed in part by the advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 9, 1984

**References**


