Nucleic Acids Research, 1993, Vol. 21, No. 6 1345–1349

Evolutionary conservation of excision repair in *Schizosaccharomyces pombe*: evidence for a family of sequences related to the *Saccharomyces cerevisiae RAD2* gene

Antony M.Carr, Katherine S.Sheldrick, Johanne M.Murray[1], Rowyda Al-Harithy[1], Felicity Z.Watts[1] and Alan R.Lehmann

MRC Cell Mutation Unit, Sussex University, Falmer, Brighton BN1 9RR and [1]School of Biological Sciences, Sussex University, Falmer, Brighton BN1 9QG, UK

Received January 15, 1993; Revised and Accepted February 16, 1993 EMBL accession no. X66795

ABSTRACT

Cells mutated at the *rad13* locus in the fission yeast, *Schizosaccharomyces pombe* are deficient in excision-repair of UV damage. We have cloned the *S.pombe rad13* gene by its ability to complement the UV sensitivity of a *rad13* mutant. The gene is not essential for cell proliferation. Sequence analysis of the cloned gene revealed an open reading-frame of 1113 amino acids with structural homology to the *RAD2* gene of the distantly related *Saccharomyces cerevisiae*. The sequence similarity is confined to three domains, two close to the N-terminus of the encoded protein, the third being close to the C-terminus. The central region of about 500 amino acids shows little similarity between the two organisms. The first and third domains are also found in a related yet distinct pair of homologous *S.pombe/S.cerevisiae* DNA repair genes (*rad2/YKL510*), which have only a very short region between these two conserved domains. Using the polymerase chain reaction with degenerate primers, we have isolated fragments from a gene homologous to *rad13/RAD2* from *Aspergillus nidulans*. These findings define new functional domains involved in excision-repair, as well as identifying a conserved family of genes related to *RAD2*.

INTRODUCTION

The structural integrity of the DNA is fundamental to cell proliferation in all organisms (1). Intricate DNA repair pathways exist to restore this integrity after DNA damage. Deficiencies in DNA repair enzymes result in a number of human genetic disorders, the best known of which is xeroderma pigmentosum (2). Seven excision repair defective complementation groups have been described for xeroderma pigmentosum and a further two may be associated with the related disease Cockayne's Syndrome (3). Yeasts are becoming increasingly important in the study of DNA repair, as shown by the findings that the human excision repair genes *ERCC1*, *ERCC2*, *ERCC3* and *XPAC* are homologous to the *S.cerevisiae RAD10* (4), *RAD3* (5), *SSL2* (6) and *RAD14* (7) genes respectively. The degree of structural conservation between species is generally highest in domains which are thought to be important for the functions of the gene products.

The fission yeast, *Schizosaccharomyces pombe*, is finding increasing use as a model organism for studying basic metabolic processes in eukaryotic cells. In particular, cell cycle control has been intensively studied in *S.pombe* (8). We are using *S.pombe* to study DNA repair processes (9–12). Radiation-sensitive mutants of *S.pombe* have been assigned to approximately 23 complementation groups (13, 14), although recent work suggests that a few of these complementation groups may not in fact be distinct (15; our unpublished observations). The functions of the genes defective in these mutants have until recently been poorly characterised, but literature published in the 1970s (reviewed in 13) together with more recent work suggest that groups of these genes are involved in excision repair (eg *rad13*, *rad15* and *rad16*), in mitotic arrest following DNA damage (*rad1*, *rad3*, *rad9*, and *rad17*) (16, 17), and possibly in recombination repair (reviewed in 9, 18).

Rad13 mutants of *S.pombe* are sensitive to UV-but not to γ-irradiation, they retain caffeine sensitization to the lethal effects of UV-irradiation, and they are extremely hypermutable by UV-irradiation (13). These are the properties expected of *S.pombe* mutants deficient in nucleotide excision-repair. Fabre and Moustacchi demonstrated that *rad13* mutant cells were unable to excise thymine-containing dimers from their DNA (19), whereas Birnboim and Nasim detected only a reduced rate of excision (20). Further evidence for a role of the *rad13* gene in excision-repair came from the findings of McCready *et al.* (21) that the *RAD2* gene of *S.cerevisiae*, also known to be involved in excision-repair (see 22) was able to complement partially the UV-sensitivity of a *S.pombe rad13* mutant. In this paper we describe the cloning and characterisation of the *S.pombe rad13* gene and we show that it is the structural homolog of the *S.cerevisiae RAD2* gene. The sequence similarity is confined to very specific regions of the gene. We have also cloned the *rad2*

gene of *S.pombe* (note that the *rad* number designations of the two organisms are unrelated) and we show that it too has regions of homology to *rad13*. These findings define a new series of conserved domains involved in excision repair, as well as showing that *rad13/RAD2* is a member of a gene family.

MATERIALS AND METHODS

Procedures for culture of *S.pombe* strains, construction of *S.pombe* gene banks (23), transformation of *S.pombe*, UV selection (12), DNA and RNA extractions, and DNA sequencing have all been described in our earlier work (10–12). Amino acid sequences were compared using the AALIGN program of DNAStar.

Degenerate PCR with DNA from *Aspergillus nidulans*

Reactions contained in 100µl, 50mM KCl, 10mM Tris–HCl pH8.3, 1.5mM $MgCl_2$, 0.01% gelatin, 0.125mM dNTP's, 100pmol of each primer, 0.1mM tetramethylammonium chloride, 140ng genomic DNA from *A.nidulans* (kindly supplied by P.Strike, Liverpool University), and 1 Unit of 'Amplitaq' (Cetus). Incubation times were 30 seconds at 94°C; 60 seconds at 53°C; 60 seconds at 70°C. Sequences from domain A (see Results) were amplified using the following combinations of degenerate primers (see Figure 3): 13A (GAYGCNWSNATHTGGATHTAYCARTT) and 13 D (CCNCCRTCRAANACRAANACNGG) for 20 cycles; 5µl of the resulting mixture was used as template with primers 13A and 13C (ARYTTRCADATNCKNCKRAARAA) for a further 20 cycles: 5µl of the resulting mixture was used as template with primers 13B (TTYYTNAARGCNGTNMGNGAYMARGA) and 13C for a third round of 20 cycles. Sequences from domain C were amplified by using similar combinations of primers: 13W (AARMGNGAYDSNGAYGARGTNAC) and 13Z (GTRTARTCNSWNCCNARNARNKG) for 20 cycles; 5µl of the resulting mixture was used as template with primer 13X (ATGATHAARGARKKNCARGARYT) and 13Y (GTNCCNCCRAANARRAANACRTC) for 20 cycles: 5µl of the resulting mixture was used as template, again with primers 13X and 13Y, for 20 cycles. The resulting fragments (81 and 153 bp predicted length respectively) were visualised on 2% agarose gels.

Products were excised from 2% low melting point gels and purified using the 'Mermaid' kit (BIO 101). They were then cloned into the 'T-vector' described by Kovalic *et al.* (24) and sequenced using standard procedures.

RESULTS

Cloning of the *rad13* gene of *S.pombe*

The *rad13.A ura4.D18* mutant was transformed with a gene bank constructed from *S.pombe* genomic DNA in a *S.pombe* shuttle vector, pUR19 (23), which contains *ars-1* and *ura4* sequences. A total of 45000 transformants were subjected to repeated rounds of exposure to 100 Jm^{-2} of UV light. After three rounds of selection, colonies resistant to UV were tested individually for co-stability of the UV-resistant and *ura*$^+$ phenotypes. Plasmids were rescued into *E.coli* from two independently isolated complemented colonies. Retransformation of these plamids into the original *rad13* mutant resulted in restoration of UV-resistance almost to the level of that in wild-type cells (Figure 1A). Restriction maps of these plasmids are shown in Figure 1B. The region common to both insert DNAs was sequenced in both directions using a combination of exonuclease III deletions (25) and oligonucleotide priming. The sequence is deposited in the EMBL database, ref. X66795. The sequence analysis revealed a single open reading frame of 1113 amino acids, whose location is shown in Figure 1B, and the size is consistent with a 3.8kb mRNA species detected by Northern analysis (Figure 1C).

In order to determine whether the *rad13* gene is essential for viability of *S.pombe*, a deleted plasmid was created by replacement of the *HpaI–HindIII* fragment containing most of

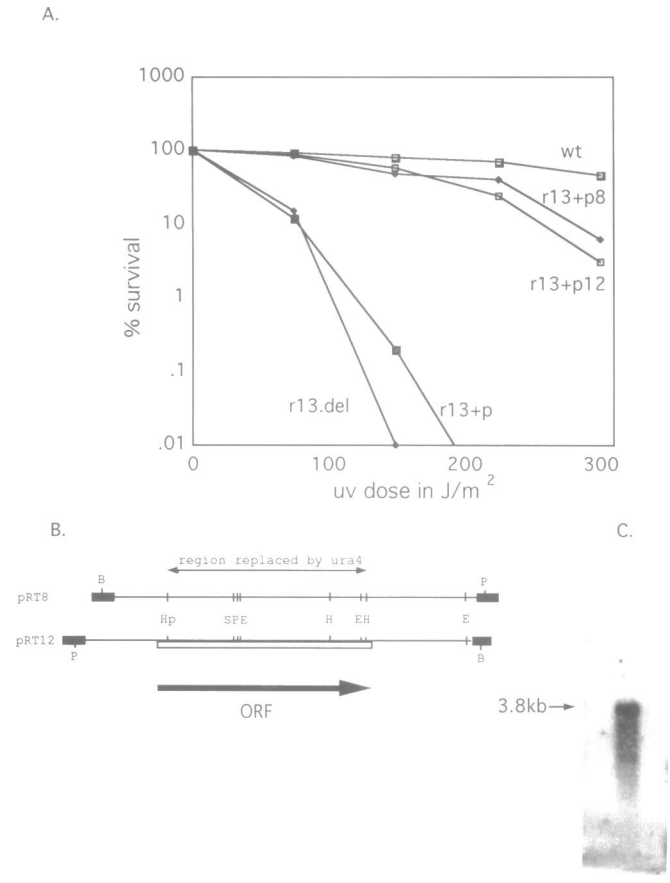
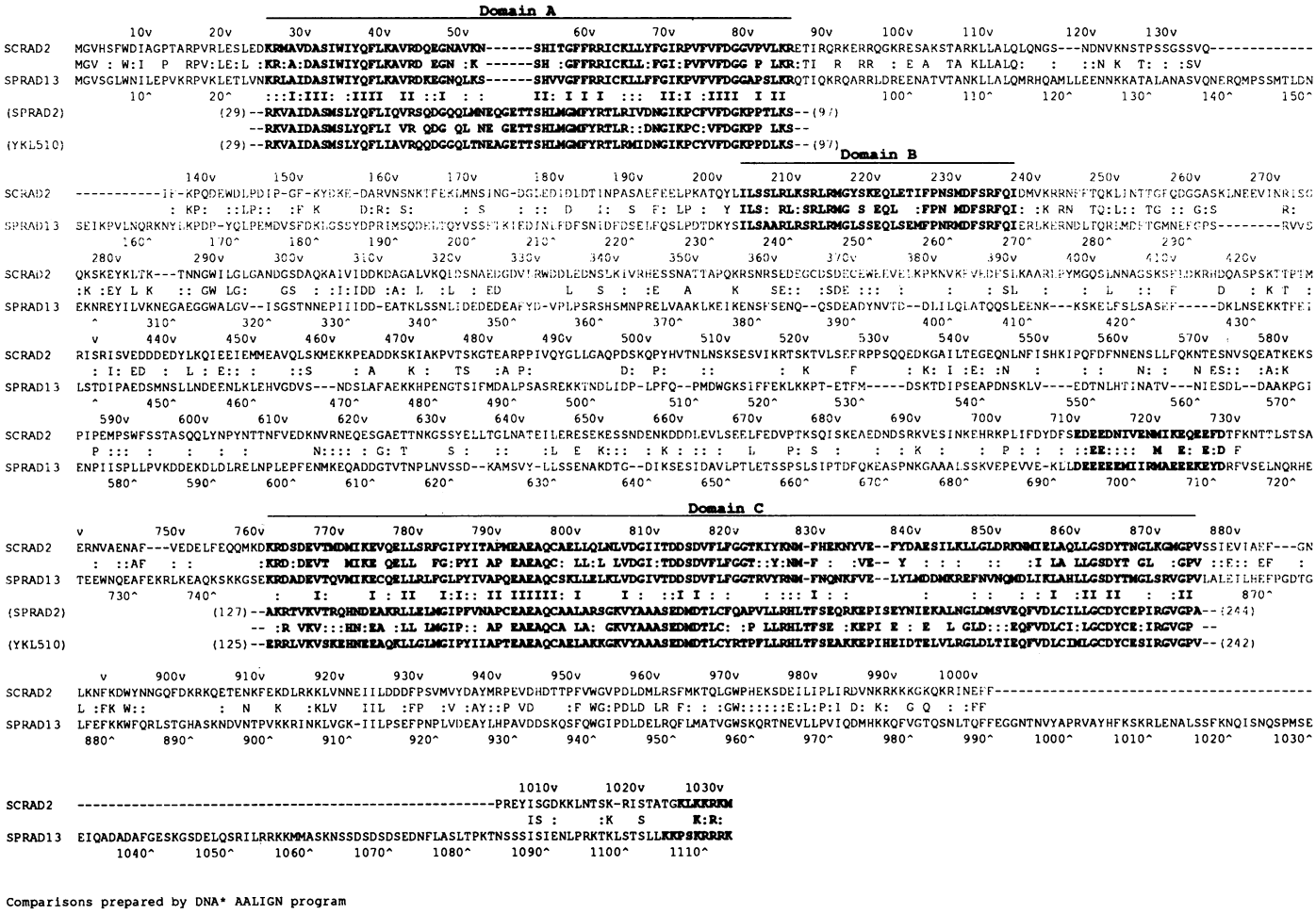


Figure 1. Characterisation of the *S.pombe rad13* gene. **A.** UV survival curves of *rad13* mutant cells and transformants derived from them. (r13+p) *rad13.A* cells with vector alone; (r13+p8 and r13+p12) *rad13.A* cells plus complementing plasmids; (r13.del) strain deleted for *rad13* sequences; (wt) wild type control. **B.** Restriction maps of the complementing plasmids, pRT8 and pRT12. The single ORF of 1113 amino acids is indicated (open box and arrow). B=*Bam*H1 H=*Hin*dIII, E=*Eco*RI, S=*Sal*I, P=*Pst*I, Hp=*Hpa*I. Flanking vector sequences are indicated by solid boxes. **C.** Northern blot analysis of 5µg of *S.pombe poly* A+ RNA hybridized with an ORF specific *rad13* probe.

Table 1. Comparison of properties of predicted products of *RAD2* and *rad13* genes

	RAD2	*rad13*
size (aa's)	1031	1113
Acidic (%)	17.5	17
Basic (%)	14	13
Hydrophilic (%)	27	30
Hydrophobic (%)	28	26
Basic Tail	+	+
pI	5.0	4.9
Charge at pH7	−35	−40

A



B

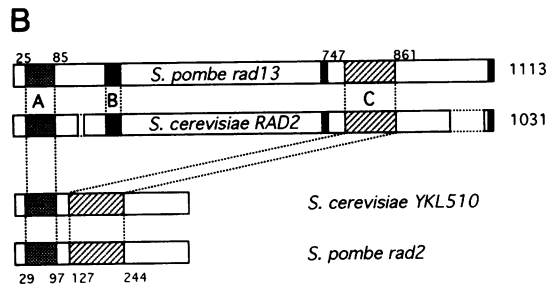


Figure 2. Conservation of a family of *RAD2* related genes. **A.** The *S.pombe rad13* (SPRAD13) sequence is compared to the *S.cerevisiae RAD2* (SCRAD2) sequence (27). Amino acid identities and similarities are shown in the central line, identity is indicated by the appropriate amino acid and conservative changes by a colon. In addition, the regions from the *S.pombe rad2* (SPRAD2) and *S.cerevisiae YKL510* sequences corresponding to domains A (position 25–85 of SPRAD13) and C (position 747–861 of SPRAD13) are shown. Domain B (position 232–268 of SPRAD13) is apparently not found in the *rad2/YKL510* gene pair. Residues common to all four proteins are indicated by a vertical line. Residues showing conservative substitutions between all four sequences are indicated with a colon on the equivalent line. The regions indicated by the dashed lines indicate gaps introduced for optimum alignment. Regions shown in bold are the three conserved domains, A, B and C, a highly acidic region and the basic tail. **B.** Diagrammatic representation of the locations of the important regions of *rad13* and *RAD2*. Dotted lines indicate gaps inserted in the *RAD2* gene product for optimum alignment with *rad13*. Also shown are the regions of conservation between the *rad13/RAD2* and the *rad2/YKL510* gene pairs.

the open reading frame (aa's 24–1096) (see Figure 1B) with a *ura4*$^+$ gene as described by Barbet *et al.* (23). After linearising this construct it was transformed into a *ura4.D18* diploid strain and integrants isolated. Following sporulation, 10 tetrads were dissected and all resulted in four progeny, two *ura*$^+$ and two *ura*$^-$. The UV sensitivity of the deletion strain was very similar to that of the *rad13.A* mutant (Figure 1A). After crossing the *rad13*::*ura4*$^+$ to the *rad13.A* mutant, no UV resistant segregants were observed upon random analysis of more than 500 spores. These results demonstrate that the *rad13* gene is not essential for cell proliferation and that the gene was effectively inactivated in the *rad13.A* mutant.

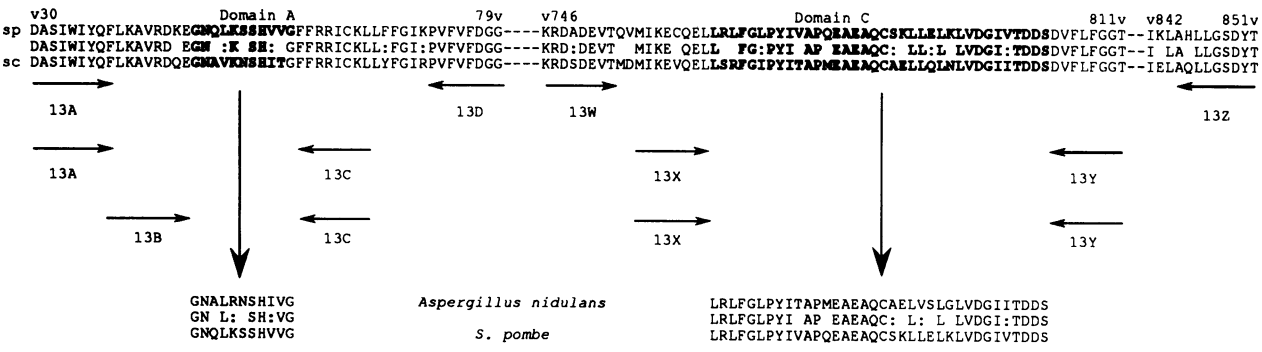


Figure 3. Demonstration of evolutionary conservation of the *rad13/RAD2* gene pair. Degenerate PCR was used to generate fragments of a *rad13/RAD2* homolog from *A.nidulans*. **Top:** Part of the homologous domains A and C from *S.cerevisiae* (sc) and *S.pombe* (sp). **Middle:** Horizontal arrows denote the positions of the degenerate primers used in successive rounds of PCR with *A.nidulans* DNA. The primers are described in full in Materials and Methods. **Bottom:** The PCR products were cloned and sequenced. Top line: amino acid sequences encoded by the PCR products amplified from *A.nidulans* DNA. Bottom line: corresponding amino acids from *S.pombe*. Middle line: amino acids identical (shown by letter code) or conserved (:) between the *A.nidulans* products and *S.pombe*.

Analysis of the *rad13* gene product

The predicted protein product of the *rad13* gene has a molecular weight of 126kD, an acidic pI of 4.9 with 13% basic, 17% acidic, 30% hydrophilic and 26% hydrophobic residues. Its most striking feature is the structural homology to the *RAD2* gene of *S.cerevisiae* (Figure 2A). There are three areas (domains A, B, C) with an unusually high degree of amino acid identity, namely 74%, 74% and 61% aa sequence identity over 66, 35, and 115 aa's in domains A, B and C respectively. If conservative amino acid substitutions are taken into account, the similarities in the three domains become 92%, 83% and 77%. The two main areas of similarity (domains A and C) are located close to the N and C termini of the *rad13/RAD2* open reading frames, and are separated by approximately 660 amino acids which, with the exception of domain B, have only limited sequence identity. Optimal alignment of the two sequences entails the insertion of a gap at position 133 of *RAD2* (Figure 2A). Sequencing of an *S.pombe rad13* cDNA clone indicates that there is no intron in *rad13* corresponding to this gap. Apart from these domains of high sequence identity, there is in addition an acidic region at aa's 693–711 of *rad13* and a corresponding acidic region at aa's 710–728 of *RAD2*. Although only 5 aa's in this region are identical, between the two species a further 7 have conservative substitutions. Finally the introduction of a gap between residues 1002 and 1003 of RAD2 permits the alignment of the C-Terminal 29 amino acids with a moderate degree of sequence conservation, both gene products having a basic C-terminus. The various important regions are aligned diagrammatically in Figure 2B. The sequence alignment enables us to define important regions of the *rad13/RAD2* protein that are likely to be involved directly in the function of the gene product. Apart from the specific regions of sequence similarity, the overall properties of the two proteins are very similar in terms of size, acidic pI, basic tail and hydrophilicity (see Table 1).

The *rad2* gene of *S.pombe*

We have also recently isolated and sequenced an *S.pombe* gene complementing the *S.pombe rad2.44* mutation. Detailed analysis of this gene will be presented elsewhere. Its predicted product shows 56% identity to an *S.cerevisiae* ORF, *YKL510*, identified in the chromosome X1 sequencing project (26). Figure 2A shows that in addition, the *S.pombe rad2* gene and *YKL510* share significant sequence identity with the *S.pombe rad13/S.cerevisiae RAD2* sequences in the A and C domains. Within these domains the sequence identity between the two members of a homologous pair is 74–86% over approximately 64 aas in domain A and 61–69% over approximately 116 aas in domain C, whereas the identity in the equivalent regions between the two gene pairs is 41% and 22% respectively (Figure 2A). Within gene pairs the relative spacing of the A and C domains is conserved (*rad13/RAD2*, approximately 660 aas; *rad2/YKL510*, 30 aas), but between gene pairs there is no such conservation (Figure 2B). These observations suggest that there is a family of *rad13/RAD2* related DNA repair proteins that may be derived from the same progenitor molecule.

Homologs from *Aspergillus nidulans*

As the evolutionary divergence between *S.pombe* and *S.cerevisiae* is approximately equivalent to the divergence between each yeast and mammalian species (28), it is highly probable that the sequences conserved between these two yeasts will also be found in other distantly related eukaryotes. In order to test this hypothesis we have carried out PCR with degenerate primers corresponding to sequences conserved between the *S.pombe rad13* and *S.cerevisiae RAD2* genes to generate novel DNA fragments from the distantly related fungus *Aspergillus nidulans*. The degenerate primer PCR was carried out as described in Materials and Methods. The primers corresponded to all possible codons coding for the indicated amino acids. The PCR products of the expected size were cloned into the 'T-vector' described by Kovalic *et al.* (24) and sequenced. As shown in Figure 3 the fragments generated from *A.nidulans* DNA are very similar to the corresponding sequences from the two yeasts. The aa sequence of the A domain is 64% identical to that from *S.pombe*, 73% identical to that from *S.cerevisiae*, and that of the C domain is 78% identical to *S.pombe*, 83% identical to *S.cerevisiae*.

DISCUSSION

The isolation and characterisation of homologous genes from different organisms provides valuable information on structurally conserved regions which are likely to be important for protein function or for protein/protein interactions. In the case of DNA repair genes, it has recently become clear that the excision repair

pathway is conserved in *S.cerevisiae*, *S.pombe* and man (reviewed in 29, 30). Thus the human *ERCC-2* gene, recently shown to be defective in xeroderma pigmentosum cells from complementation group D (Weber, cited in 30), has a high degree of sequence similarity to the *RAD3* gene of *S.cerevisiae* (5) and to the *rad15* gene of *S.pombe* (12, 31). The *RAD3* gene is known to encode an ATP-dependent DNA helicase and the regions of homology are highest in seven domains conserved in DNA helicases. Likewise the *ERCC3* gene (32) has homologs in both yeasts (6; cited in 30). We have now shown that the *S.pombe rad13* gene is homologous to the *S.cerevisiae RAD2* gene but this conservation is confined to three relatively short regions. Within these regions the degree of identity is extremely high, which strongly suggests that these domains are essential for protein function, although the nature of this function remains unknown. The *RAD2* protein is a weakly expressed gene which is essential for the incision step of excision-repair following UV damage (22). Unlike the other *RAD* genes involved in excision repair, the *RAD2* gene is inducible by DNA damaging agents (33). The domain-specific sequence similarity between *RAD2* and *rad13* is in direct contrast to our findings to be published elsewhere of a different pair of homologs, namely *RAD1* of *S.cerevisiae* and *rad16* of *S.pombe*. In the latter pair of excision repair genes, the sequence similarity (approximately 30% identity) extends over almost the entire length of the presumptive proteins.

Two of the conserved domains in *rad13/RAD2* have also been found in another homologous pair of genes, *rad2/YKL510* suggesting (1) that these genes are also involved in excision-repair (there is at present no direct evidence for this), (2) that they have a function related to that of *rad13/RAD2*, and (3) that this represents a gene family. The identification of conserved genes in the two highly diverged yeasts strongly suggests that these genes will be conserved throughout the eukaryotic kingdom. Using degenerate primer PCR we have isolated fragments from the *A.nidulans* genome that show homology to domains within the *rad13/RAD2* gene pair, and we expect that a similar approach will enable the homologous genes to be isolated from other organisms. Such experiments will assist in the isolation of the complete set of proteins involved in excision repair in mammalian cells and the reconstruction of this system *in vitro*.

While this work was being prepared for publication we learned (M.MacInnes, personal communication) that the human *ERCC5* gene (34) was homologous to *RAD2*. Comparison of the sequences of the *ERCC5* predicted protein with *rad13* (to be presented in full by MacInnes *et al.*) shows that sequence similarity is confined mainly to domains A and C. We are also aware that S.Clarkson (personal communication) has cloned a human homologue of *RAD2*.

ACKNOWLEDGEMENTS

This work was supported in part by EC Grant number B17-0026 to A.M.C. and A.R.L., MRC grant number G88/03213CB to J.M.M. and F.Z.W., a MRC HGMP studentship to K.S.S. and a Saudi Arabian Government grant to R.A.H. We thank Dr J.H.J.Hoeijmakers for helpful suggestions.

REFERENCES

1. Friedberg, E. C. (1985) *DNA Repair* (Freeman, San francisco).
2. Lehmann, A. R., & Dean, S. W. (1989) in *Handbook of Experimental Pharmacology*, eds. Cooper, C. S. and Grover, P. L., (Springer, Berlin), Vol. 94(II), pp. 71–101.
3. Venema, J., Mullenders, L. H. F., Natarajan, A. T., Van Zeeland, A. A., & Mayne, L. V. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4707–4711.
4. Van Duin, M., De Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M. H. M., Hoeijmakers, J. H. J., & Bootsma, D. (1986) *Cell* **44**, 913–923.
5. Weber, C. A., Salazar, E. P., Stewart, S. A., & Thompson, L. H. (1990) *EMBO J.* **9**, 1437–1448.
6. Gulyas, K. D., & Donahue, T. F. (1992) *Cell* **69**, 1031–1042.
7. Bankmann, M., Prakash, L., & Prakash, S. (1992) *Nature* **355**, 555–558.
8. Nurse, P. (1990) *Nature* **344**, 503–507.
9. Lehmann, A. R., Carr, A. M., Watts, F. Z., & Murray, J. M. (1991) *Mutation Res.* **250**, 205–210.
10. Murray, J. M., Carr, A. M., Lehmann, A. R., & Watts, F. Z. (1991) *Nuc. Acids Res.* **19**, 3525–3531.
11. Fenech, M., Carr, A. M., Murray, J. M., Watts, F. Z., & Lehmann, A. R. (1991) *Nuc. Acids Res.* **19**, 6737–6741.
12. Murray, J. M., Doe, C., Schenk, P., Carr, A. M., Lehmann, A. R., & Watts, F. Z. (1992) *Nuc. Acids Res.* **20**, 2673–2678.
13. Phipps, J., Nasim, A., & Miller, D. R. (1985) *Adv. Genet.* **23**, 1–72.
14. Lieberman, H. B., Riley, R., & Martel, M. (1989) *Mol. Gen. Genet.* **218**, 554–558.
15. Schmidt, H., Kapitza-Fecke, P., Stephen, E. R., & Gutz, H. (1989) *Curr. Genet.* **16**, 89–94.
16. Al-Khodairy, F., & Carr, A. M. (1992) *EMBO J.* **11**, 1343–1350.
17. Rowley, R., Subramani, S., & Young, P. G. (1992) *EMBO J.* **11**, 1335–1342.
18. Subramani, S. (1991) *Mol. Microbiol.* **5**, 2311–2314.
19. Fabre, F., & Moustacchi, E. (1973) *Biochim. Biophys. Acta* **312**, 617–625.
20. Birnboim, H. C., & Nasim, A. (1975) *Mol. Gen. Genet.* **136**, 1–8.
21. McCready, S. J., Burkill, H., Evans, S., & Cox, B. S. (1989) *Curr. Genet.* **15**, 27–30.
22. Friedberg, E. C. (1988) *Microbiological reviews* **52**, 70–102.
23. Barbet, N. C., Muriel, W. J., & Carr, A. M. (1992) *Gene* **114**, 59–66.
24. Kovalic, D., Kwak, J.-H., & Weisblum, B. (1991) *Nuc. Acids Res.* **19**, 4560.
25. Henikoff, S. (1984) *Gene* **28**, 351–359.
26. Jacquier, A., Legrain, P., & Dujon, B. (1992) *Yeast* **8**, 121–132.
27. Madura, K., & Prakash, S. (1986) *J. Bacteriol.* **166**, 914–923.
28. Sipiczki, M. (1989) in *Molecular Biology of fission yeast*, eds. Nasim, A., Young, P. Johnson, B.F., (Academic Press, New York.), pp. 431–448.
29. Hoeijmakers, J. H. J., & Bootsma, D. (1990) *Cancer Cells* **2**, 311–320.
30. Lehmann, A. R., Hoeijmakers, J. H. J., Van Zeeland, A. A., Backendorf, C. M. P., Bridges, B. A., Collins, A., Fuchs, R. P. D., Margison, G. P., Montesano, R., Moustacchi, E., Natarajan, A. T., Radman, M., Sarasin, A., Seeberg, E., Smith, C. A., Stefanini, M., Thompson, L. H., Van Der Schans, G. P., Weber, C. A., & Zdzienicka, M. Z. (1992) *Mutation Res.* **273**, 1–28.
31. Reynolds, P. R., Biggar, S., Prakash, L., & Prakash, S. (1992) *Nuc. Acids Res.* **20**, 2327–2334.
32. Weeda, G., Van Ham, R. C. A., Vermeulen, W., Bootsma, D., Van Der Eb, A. J., & Hoeijmakers, J. H. J. (1990) *Cell* **62**, 777–791.
33. Siede, W., & Friedberg, E. C. (1992) *Mol. Gen. Genet.* **232**, 247–256.
34. Mudgett, J. S., & MacInnes, M. A. (1990) *Genomics* **8**, 623–633.