Volume 17 Number 24 1989 Nucleic Acids Research

Nucleotide sequence of the REC1 gene of *Ustilago maydis*

David W.Holden, Ad Spanos and Geoffrey R.Banks

Genetics Division, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Submitted November 20, 1989 EMBL accession no. X17663

Conventional genetic studies have demonstrated that the REC1 gene of *Ustilago maydis* plays a central role in regulating the genetic system of this fungus. *rec1* mutants are very sensitive to UV light and produce a significant proportion of non-viable cells. In diploid cells, the *rec1* mutation causes variable colony morphology, increases the spontaneous level of mitotic recombination but reduces radiation induced recombination (1). We have begun a molecular analysis of this gene by isolating a subclone from a cosmid library of *Ustilago* DNA (2) that complements all of the above phenotypes when transformed into *rec1* strains. Sequencing of the complementing region revealed an open reading frame (ORF) of 1566 bp that could encode a protein of 522 amino acids and 56,800 daltons. We have confirmed that this ORF represents the REC1 gene by gene disruption and complementation experiments (3). No significant homologies were found between this gene and sequences in the EMBL/GENBANK and SWISSPROT data banks.

```
   1  ATTAACCAACAGAAATTAAGAAGAAATGGTGAAGGATGCAGGGTCCAGTCGCGTTGAATTTTAAGTCGCGTTTGGTGGGTGCCAAAGACCAACCCGCAGCTT
 103  AGCTTAGCTGAGTCGCGTCATGGATGTCACAGCGACCTTGACTAACTTTAACCAACTTCTTTACTCTTGCTCTTGTCGGATGGATGACAGCGACAACTCAAG
 205  CTCATGACTTGAGACCATCACATGCTTTCTCGAAAACTGCTCCTTGGCAACACTATTTTACCTCCACGCGCTGACGTTGCTGAATCCAGTAAACTAGCCCAT
 307  CTTCCTCGATCCGCATCGGCCACTTGGTCATTTCACTCCACTCTCGCCAGTTGACGCCACGTCAGCATGCCGGCCGAGGGAGCTTGCGATGCTGCATCTCTC
                                                                                M  P  A  E  G  A  C  D  A  A  S  L     12
 409  ATGACGCTGACAGCTACTTTGTCCGATGTCACGGGTCTTGCCAACTTGCTCAAGTCGGTTGCCATTCAGACTCATGCCGTAGTCATCGCCTCAAGTTCGGGC
       M  T  L  T  A  T  L  S  D  V  T  G  L  A  N  L  L  K  S  V  A  I  Q  T  H  A  V  V  I  A  S  S  S  G     46
 511  CTCGAGATCATCACTGAGCTCAATCGCACGCTGCAAGCTCACGCTTACCTCTATTCGCACATGTTTGACTCCTACCGCTTTGAAAACGCTCAAGACGATGTC
       L  E  I  I  T  E  L  N  R  T  L  Q  A  H  A  Y  L  Y  S  H  M  F  D  S  Y  R  F  E  N  A  Q  D  D  V     80
 613  CGAGGCTCAACATCATTGCAAGCCCGTTCAAGACCAAAGAAGCGCTCCAAACTGACATCCAAACACGCCGAGACAGCCGACTCTCAATCTTCTGCCGCCTCG
       R  G  S  T  S  L  Q  A  R  S  R  P  K  K  R  S  K  L  T  S  K  H  A  E  T  A  D  S  Q  S  S  A  A  S    114
 715  TCCGATCACGAATCTGGTCAATCACAGGCTCACACGAAAAAGCGTTTCAGAGCAAACTCTCATTCATACGCTCAAGAGGCCGACAGAGTTCACGACGAGCCA
       S  D  H  E  S  G  Q  S  Q  A  H  T  K  K  R  F  R  A  N  S  H  S  Y  A  Q  E  A  D  R  V  H  D  E  P    148
 817  GACAGTGTCTCGTTCGAAGTCAACCTCCAAACTTGGATCTCTTGCCTCAATATCTTCGGCGGCGTAGGACCCTCTCGTCCGCACAGCTCCTCCAGCGGCCTT
       D  S  V  S  F  E  V  N  L  Q  T  W  I  S  C  L  N  I  F  G  G  V  G  P  S  R  P  H  S  S  S  S  G  L    182
 919  CCTGGCTTCCGACCAGAACAGGGCAGTGCTGAAGCTCCACCAGGTGGGCGAGGATATCAACGCACACGTTACGGTGTTGCTGACGCTTACGGTGCTGAACGT
       P  G  F  R  P  E  Q  G  S  A  E  A  P  P  G  G  R  G  Y  Q  R  T  R  Y  G  V  A  D  A  Y  G  A  E  R    216
1021  GGTACCTCCGTCGAACGTGGATTTGATCGTAATCCCTTCTCGTCGTCTGCAAAGGCCACGCGGATGAAACTCAGCTATCAAGGTCATGGTAATCCACTCGTT
       G  T  S  V  E  R  G  F  D  R  N  P  F  S  S  S  A  K  A  T  R  M  K  L  S  Y  Q  G  H  G  N  P  L  V    250
1123  CTCGAGCTTGAGCAAGACGCCAACGTATTGACGCGTGTATCTATGTCGACATACGAACCCAGCTTCCTCACAGATATGGTCTTTGAACCGCAGAATATGGTG
       L  E  L  E  Q  D  A  N  V  L  T  R  V  S  M  S  T  Y  E  P  S  F  L  T  D  M  V  F  E  P  Q  N  M  V    284
1225  GCTCAAGTCATTGTTGCATCCGAGCTCATGCAATCTGCCTTTACTGAGATTGATGCCAGCTGCAAGAAACTCTCGATCCTCATCACATCGCCCCACTCATTA
       A  Q  V  I  V  A  S  E  L  M  Q  S  A  F  T  E  I  D  A  S  C  K  K  L  S  I  L  I  T  S  P  H  S  L    318
1327  TCGACGTATGACGGTGATCAGCGCACAGAGGCTCCGGCACCCACCAAGAGGAACACTTCAGCCAGCATGCTCAAGTTCAGGGCCATCTCGGATACCGGATCG
       S  T  Y  D  G  D  Q  R  T  E  A  P  A  P  T  K  R  N  T  S  A  S  M  L  K  F  R  A  I  S  D  T  G  S    352
1429  TCAGAAATGGAGTTCCCAGCCAGTCTCACCTCATCTGACCCTACGGGCGTTATAGAGAAATTCGTAGCCCTGCCTGGCAGCAGCGAACAATGGTACGACTTC
       S  E  M  E  F  P  A  S  L  T  S  S  D  P  T  G  V  I  E  K  F  V  A  L  P  G  S  S  E  Q  W  Y  D  F    386
1531  ACATTGCTCTCACGAACCATGTCTGTCTTGCGATCGTCGATAAAGACCTCTCTCCGAATGGACGAGGCTGGTCTCATTAGCTTTCAGTTTATGATGCCAAAA
       T  L  L  S  R  T  M  S  V  L  R  S  S  I  K  T  S  L  R  M  D  E  A  G  L  I  S  F  Q  F  M  M  P  K    420
1633  TATCGACGTGCCGCCGCAGCGGGTGCGCCACTAACAAACGCTGCTGCTGGCCAGGCTGCACATGAAGATGAGCAAGACGCCTTCTGCGAGTTCCTGGTGAGT
       Y  R  R  A  A  A  G  A  P  L  T  N  A  A  A  G  Q  A  A  H  E  D  E  Q  D  A  F  C  E  F  L  V  S    454
1735  ACTTTTACATGCATCGCTGACTCCAGCTTGGCGCTTAAACCACCGCGATTCGGCAAACTCTGTTCGTACTCGGATTGCCGGATTCCGCGGATTCGGACTATG
       T  F  T  C  I  A  D  S  S  L  A  L  K  P  P  R  F  G  K  L  C  S  Y  S  D  C  R  I  P  R  I  R  T  M    488
1837  ATTAGTGCAGTGCATACTGACCTTGTGGAACGTTACTTTCATGTCTCTTTTTTTCGTGGTCTGTCAACGGTGTTCAGTGCTGTCCGCTCGATACAAGTACCT
       I  S  A  V  H  T  D  L  V  E  R  Y  F  H  V  S  F  F  R  G  L  S  T  V  F  S  A  V  R  S  I  Q  V  P    522
1939  TAATCGTGTGAAGCCTTGCGCAATCGCCGCTGAAGTTGATCTTTGCGGCTGATATCCTCTTTTCAAAAATGTGAATCACGAACATGCTCGCCATCTGTTAGG
       *
2041  TTTATCATGTATTGAAGAAGGTGTGTGTCTGAGTAACGCAACAAACGCCGGAGGATCAACCAGTTCTTGGTATGCTGACATGGTCAGCCGCCTCACGATTCT
2143  GCAACAGCATGGTGAAGCTCATTGTGGACAGTGCGCATCAAAAATGGGCTCAGAATTTGCGCACGTTTACATGTGCTCGTGTTTGAAACGCCGGCAAGCGAG
2245  GTTCGGATGCAAACGTGTGCCAAGCTCAAGCAAGCGCTCGCTACCGACTGGAGCAAGCTT
```

Acknowledgement: We thank N.Kanuga for technical help during the course of this work.

References: (1) Holliday,R., Halliwell,R.E., Evans,M.W. and Rowell,V. (1976) Genet. Res. 27, 413-453. (2) Wang,J., Budde,A.D., and Leong,S.A. (1989) J. Bacteriol. 171, 2811-2818. (3) Holden,D.W., Spanos,A. and Banks,G.R. (manuscript in preparation).