

DNA sequence of the transposase gene of the new category of class II transposon, Tn2501

A.K.Turner and J.Grinsted

Department of Microbiology and Unit of Molecular Genetics, University of Bristol, Bristol BS8 1TD, UK  
 Submitted October 26, 1987 Accession no. Y00502

The indicated gene for the Tn2501 transposase encodes a protein of 994 amino acids. This shows similarity to transposases of other class II transposable elements, although the sequences do not align exactly and "padding" is required to obtain maximum % positional identity. When compared with the Tn21 transposase (1), this is 55% and with the Tn3 transposase (2) 31%.

```

M P R R Q Q I A C T S E E Q E R L L V I P D D E I I L T R M C F L N E O D I A L I N
10
ATGCCACGGGACAGACTACTGCAGCTGAAAGAGCAGGACGGCTTGCCTGTATTCACAGATGATGAAATATTCGACCCGAAATGTTTCTGAAAGACAGGATATGCACTCATTAAT 120
K H R R P A N R L G F A V L L C Y L R G P G F I P D K S N V P H S S V I S R I A
240
AAACACGGGGCGGACAGAAATGGCTGGGPTTTCGTGTATTAAGTCTGTATTCTGGTGGACCGTGGATTTATCCGGACAAAAGCAATGCCCTCAGATAGGCTCATATCCAGAAATGCCCT
360
S R L K L Q F D L M P P E Y A S R E Q T R W E H L T E L Y R L K L S P F S R S L
TCCGCACTGAAGCTCACCCCTGATTTATGGCCGAAATATGCTTCAAGAGGCAAAACCGCTGGAAACATCTGACTGAGCTTTATGATACCTGAAATATTCCTCCGTCAGCGCATCCTG
480
Q K D C I P H L F P Y M R T D K G F M L A E F M L H L H S N W I F P S V D
CAAAAAGACTGTATCCGGATCTGCATCCCTATGGATGGCAAGCAGCAAGGGTTTATGCTGGGCGAAGAAATGCTCAATGGCTACATAGCAATAATTTCCCTCTGTGATG
560
V I E R T L L A E A V L A D R A V F S A L T A Q L E K Q H S L A L D S L L K S E
GTGATGCAOAGGAGCTTGGCGAAGCCCTGCAGCTGGCTATGATAGGCGGTATTTTCAGCACTTACCGCACAACTGGAAAACAGCAGATAAATCAGCAGCTGCAGACATGCTCTCAATCAGAG
640
G E Q A S R L A W L L O P P G K I N G K N V L Q H I D R L N S I A A L G L P D G
GGTAAACAGCTTCCCGCTGGCATGGTACTTACAGCTCCGGGTAATAACAGCGTAAGAAGCTGCGCAACATATCGACCGGCTTAATTCATCGCCGGCTGGGATGCCGTATGCTG
720
I T L S I H Q N R L L K L A R E G R K M S S R D L A K F T D V R R Y A T L V C V
ATTACGGCTTTCGATCCACCAAGCAAGCTGTACTAGCTGGCCGCTGAGGGCCGGAATAAGCAGCAGAGACTGGCAAAATCAGCGATTCAGACGTTACCGCTACGGCTGTTCGGCT
800
T G E A Q A T L P D E V I E L H E R I L G T L E P S R A K R T O A E R L Q L T G K
ATTACAGAACGCGACAGCTACTGATGAGTTTACGGCTGACGACGAGCTTATGGGACTCTATTAGTGGGCAAAAGCAGACAGCAAAAGCAGAGGCTTACAGCTGACGGGAA
960
L I O S K L K Q Y V T V G Q A L L H A R E S G E D P W A A I E D V L P W Q E F I
CTCATCAGAGCAAGTTAGCAAAATATGTTACTGTCCGTCAGGCACTGCATCGAGAGAAATCCGGTGAAGAATCCCTGGCGGCAAGAGAAAGATGCTCTCCCTCCGACGGAGTTCACT
1080
N S L E E T Q F L S R K G N F D P L H L I T E K Y S T L R K Y A P R M L S A L O
AACCGCTGGAAGAAAGCAGTGTTCGCTCCGTAAGGGCAATTCGACCGCTTCCAGCTGATCACCGAAAATACAGTAGCCGTGGTAAATACGGCCCGGCTATGCTTCCAGATTGAG
1200
F I A T P F A G F L S D A L D T I R D M Y R K O L R A K V P P A A P T G F I P E S
TTCATCCCGACACCTCCCGCAGACTGGCATGCGCTGGACACCATCAGGACATGACACCTTAAACACTTCCTAAAGVTTCCCGCCAGCTCCCAACAGGATTTATCCCTGAAAGC
1320
W R K L V L T P E S G T D R K Y E F C V N W E L K G A L R S G D I W V K G S R
TGGGAAAACCTGGTACGGCTTCAGGATGACACCGCAAGTACTAGAGTTTGGCGTAATGAATCAACTCAGGGGGCTTATAGATCCGCTATATGGTAAAGAGTATGGCCGCG
1440
Y R N F D D Y L I P A A E F E K S R H N D Q L Q L A V Q T D C R A Y L Q A R M T
TACAGAAATTTGATGATTAATCTACCCCGGCTGCTAGATTGAGAAATCCCGACATAATGACAGCTTACAACCTGGCGCTCAGACCGATTCGCGCGGATACCTCAGCCCGGATATGAGC
1560
L L A S R L E E V N A M L A G A D L P D V D I S D K G V I T P L E N S V P S G
CTTCTGCTACGGCTGGGAAGATTAAAGCGATGGCTTCCGCGGTATTGGCCGATGTTGGATATCTCAGATAAAGGGCTAAATAACTCTCCCGCTGGAGAACGTGTTCCGGGG
1680
A S P F A D L V Y G M L P H P K I T E I L E E V D N W T G F T R H F A H L K N N
CCTTCCCGCTTCGGATTTGGTATTATGGATCTCTCTCACTCCGAAATACGGAGATACGGAGAAAGTGGACAACTGGAGGGGTTTACGGCTACTCCGACACTCCAAAATAAT
1800
N V R P K I G R I N L G L I L D L G L I L R L D L H D K R L F I H K G A R E Y F G L Q S V S
AAGCTCAGACCAAAAGACGAGACTGTCTCTCAGGATTTCTGGCGAGGACATCAATCTGGGCTGACAAAATAATGGCGGAGTCTCSCCGGGGCAACAAAATGCTACTGAGG
1920
I Q A W Y I R D E T Y S A A L A E F L V N A O K A R P L A A F W G D G T T S S D
ATTCAGGATGTCATCAGGAAATGAAACCTATTCAGCGCAGCTGGCGAGCTGGTCAAGCTCAGAAAAGCGGCACTGGCTGGCTGATTCAGGACAGATCAGCATGCTGCAGAC
2040
G C N F R V G S H G R Y A G O V N L K Y G O E P G V Q I Y T H I S D Q Y S P F Y
GGCAAAACTTTCGGGTAGCCAGTACCGGGTATCGCGGTCAGGCTCAATCTAAATATGGCTCAGAGCCGCGTGCAGATTATCAGCGCATATCTCAGACCAATATAGCCCGTTCTAC
2160
A K V I S R V R D S T H V L D G L L Y H E S D L E I T E H Y T D T A G F T E H V
GCCAAGTATACCGGGTCCGGCTACACCCATGCTGTGATGGTCTGCTGTATCATGAAAGCACTGGAAATACCGAGCATACACAGATCAGCCAGGTTTCACTGAACATGTT
2280
F A L M H L L G F A F P R I R D L H D K R L F I H K G A R E Y F G L Q S V S
TTGCTGTGATGATGCTGGGATTCCTTTGGCACCGAGATCGGGATCTTCATGCAAGGGCGCTGTTTATTCAGAAAGCGGACTCATCGGGGCTTCAGTCTGCATCATCA
2400
T T S L N I A H W I E L R L A N A S I K Q G T V T A S L M I K K L A S Y
ACAAACCGCTGAATAAAGACACTTACGGCCACTGGGGAGATATTACGGCTAGCAGGCTGATTAACAGGGGACTGTACCGCATGCTGATGATATAAAAAGTTGGCCAGTTAC
2520
P K O N G L A K A L E R I E R T L F M L D W F R D P L R R R V Q A G L N
CCAAAACAGAAATGGACTGACCAAGCGCTGAGAGAGATTGGCCGCACTGACGGGAGCGCTGTTTATGCTGGACTGGTTTCGATGATCCCGCGGTACCGCCAGCTACAGACGAGGCTGAAT
2640
K G E A R N A L R A V F M H R L G E I R D R G L E N Q S Y R A S G L T L L T A
AAGGGGAGCCCGGTAATGCCCTGGCGAGCGGTTTTATGCGCCGGCTGGTAAATCAGGATGCTGGCGGAGAAATCAGAGCTACCGCCAGCTGAGCTACGTTACACTGCTG
2760
A T L W N T V Y I E R A I E S L K R K C I P I N W O L V S H L S P L G W E H I
GGGATCAGTTATGGACCGGATATATAGAAAGACGATGAGTCACTAAAAGCAAGAGTATTCGGATTAATCAACAGCTGTATCTCATTTCCCTCCGCTGGGCAATATC
2880
N L S G D Y V W R N N I K L G S G K Y R S L R T V D T E L Y K K Q S *
AATCTGAGTCGGATACCTGCGGGTAATAAATTAAGCTAGGGCTCGGAAAATACCGCTCATATTAGCCAGCTGCAGCCGAAATGCTAGCAAAAACAGTCTGACCGGGGATAATTTCC
3000
GTTTTCCAGGGCGGCC 3018
    
```

The complete nucleotide sequence of the *tnpA* gene of Tn2501. The 5'-3' strand of the terminal 3018bp of Tn2501 is shown, and the position of the *tnpA* gene and its encoded transposase indicated. The last 48bp at the 3' end are the inverted repeat of the element (3), and nucleotides 1-613 have already been published (4).

References

1. Ward, E. and Grinsted, J. (1987) Nucleic Acids Res. 15, 1799-1806.
2. Heffron, F. et al (1979) Cell 18, 1153-1163.
3. Michiels, T. and Cornelis, G. (1984) J. Bacteriol. 158, 866-871.
4. Michiels, T. et al (1984) J. Bacteriol. 169, 624-631.