Volume 16 Number 24 1988 Nucleic Acids Research

Mouse spleen derived cDNA clones containing *per* repeat sequence

Shin-ichiro Nishimatsu[1,2], Kazuo Murakami[2], Youji Mitsui[1] and Norio Ishida[1]*

[1]Cell Science and Technology Division, Fermentation Research Institute, Agency of Industrial Science and Technology, Higashi 1-1-3, Tsukuba Science City 305 and [2]Institute of Applied Biochemistry, University of Tsukuba, Tennohdai 1-1-1, Tsukuba Science City 305, Japan

Submitted November 12, 1988 Accession nos. X12806–X12809 (incl.)

The period (*per*) gene had a hexamer repeat DNA sequence(ACNGGN)n, which affected the courtshipsong of *Drosophila*, in its protein coding region(1). Recently, we cloned 10 *per* repetitive sequences from the mouse genomic DNA library(2)(3). Young et al. reported the presence of several transcripts encoding *per* repeat sequences in the mouse spleen(4). In order to determine whether multiple genes encoding *per* repeat are expressed in the mouse tissue, we have screened a mouse spleen cDNA library(λgt10) using *per* repeat (cp2.2) as a probe. Four cDNA recombinants have been cloned and sequenced. Interestingly, all of these 4 clones had a *per* repeat sequence(underline) in the 3' untranslated region. These results suggest that the *per* repeat may be important for gene regulation. Using FASTP computer program and the algorithm of Lipman and Person(5), the putative ORF of these 4 cDNA clones are found to belong to the unknown peptides, which deserve further attention. Sp34 had entirely 43.3% homologous to *Drosophila per* gene at the nucleotide level. RNA polymerase III recognition site(6) was observed in the clone Sp28(double underline).

SP 34

```
CTCCCCTGTCAGAGTCTCCCTCGAGCCTCGATGCCGATTGTCTGCTTTGACCCTCCGTATGAGTTATACTTACATTGGCCCCTGGCTGGGGAAGTGCAAT   100
                                 M  P  I  V  C  F  D  P  P  Y  E  L  Y  L  H  W  P  L  A  G  E  V  Q  S
CCCAGTACTGGGGGACTGGAGACAGGGGGATTTCTGGTGCTCACTGGCCAGCACCAGGTTCACTGAGAGAGGCTGTCTCAGAGAAGAAAAAGATCATCAA   200
  Q  Y  W  G  T  G  D  R  G  I  S  G  A  H  W  P  A  P  G  S  L  R  E  A  V  S  E  K  K  K  I  I  K
GAAAGACACCTGACATGAACCACTGGCTTCCATAGCACAGGCACACGCACAGGCACACGCACAGGCACACGTACAACCCAGGAGTTCTAGCTAGTTCTGC   300
 K  D  T  *
TGCTTTAGCTCAGTGGCAGAGTTCTTAACATGCCCATGCGTGTCACCACACACATATTCATGCAAATAGAAAAAGTAGAGAAAATTAAAGTTATCAAAAA   400
TCTCTTCAAACCACATCTGAACATTAGTCCCCTTTCACCTCCTATATAGTCAAGATTTTTACTAGGCTGTCAATGAAAGTCAACCAATAAAAATCTCTAA   500
ACACCTCAAAAAAAAAAAAAAA   522
```

SP 28

```
CCCTTGGGTGGTGGCAGAATCAAGTTGCTAGCAGCAGGCAATCACTCGTGAGCTTTCTTCCTCTGTAGGTGCCTGGCCCATTTAGTTACACACTGCAAAT   100
CAAAACCTGCCACAACAGAATCAAGAGAAATGCTGACACTCCAATCCTCTGGTACAGGACAGAGGAAAGTGGCAGTACATATGGGATTATTAGAAAATAT   200
                              M  L  T  L  Q  S  S  G  T  G  Q  R  K  V  A  V  H  M  G  L  L  E  N  I
AAGTCTAAAAATATCTAAAGAAGTTGGAGGTCTATATGATACAAGGATGGCTCAACAAGGGCAGCTATTAAAGTTGTTGGTGGGTACAGTGAGAAAGTGG   300
 S  L  K  I  S  K  E  V  G  G  L  Y  D  T  R  M  A  Q  Q  G  Q  L  L  K  L  L  V  G  T  V  R  K  W
GACCATTCATTTTTTATGGATCTGTGAAATTTTATATACTTAGAATTTAAAACTTAAAAGATATCCTAACTACAAAACTTTTATAATAGACAGGCTATCA   400
D  H  S  F  F  M  D  L  *
GCTCATGTCTGAACAGAAGATCTATTATCAAATAATATAACTTCAGCAAATGCATTAATATAGTATATATATATATGTATGTGTATATATATATATATAT   500
ATATTAATATATATGAATGTCTTTTCAAGACAGGGTTTCTCTGTGTAGCCCTGGCTGTCCTGGAACTTACTCTATAGACCAGGCTGGTCTTGAACTCAGA   600
ACTGTGTCTTTCTTCTCAGTGTTGAGATTACAGGTATATGTCACCATACTCAACAATGATTTATGTATTTGTCGGAAATGTTTTAAACAAGAAGATTCTA   700
GAGCTCCTAAATGAATGATAACAATCAAAGGATGCTCCAGGCACAGGGGTGTCAGGCACAGGGTTGTCAGGCATAGGGGTGTCAGGCACAGGGGTGTCAG   800
CAGGGTCTCTGGGGCCCTTAAAAACCCTGGCTGGGCAGGGACTGTGGGGCTAATAATGAACTTCTGATATACTCCACTGTGTTTCTTTCGGATTTACATA   900
GATATTTCTAACCATCTCCCAAAGCTTCTGACATATTGTGCATATCTCCAGGTTACAAATGATTGTTCAAAAAGAGGGCATAATTTTGAGTTGAATTTTG  1000
AGAAACTTGTGTGAAATAGTTTGTCTTTAAAACCAAGGCAGTGGGGTTGGGGATTTAGTTCAGTGGTAAAGCACTTGCCTAGCAAGCAGAAGGCCCTGGG  1100
TTCTATCCTCAGCTATGGGAACAAACAAACAAACAAACAAACCTAAATACATAAATAAAAAAAAAAAAAAAA   1172
```

Sp 6

```
TTTTTTTTTTTTTTGATCATGGCTTAACTTAAYTGGTGCTCTCCTCCCACTTGTCTGCATTTGGCTCCGATTAAGTGAGGAAATATCAGTTTGAAATTTG   100
TGATTTTCTTTCTGTTTCGTCTGACCACCAAATCCATTTTCCTTTGAATTATCTGGTTTATAAATGATAGGCTTTCCACATCCCATGTTTTCTTCACCCG   200
ATCTGTAATGTGACCAGTGGACAACCTGCCTGAGTGCATCCTCAGTAGAGACCGCTGTACCATCCGGGTTTGCATAAAACAGCAAAAGGGGCTGTAAGTG   300
GCATCGTATGCACTTGGAGACCACATCTTTCCATCTGGTTCCCATCTTTAAAAGAAAAATCATAATAAATAAATATAAACCAATTCCTGAAATGTGTAAT   400
ATTAATAGGCAGAAACTGGATTCATTTACATTTCAAAGTACTTCAAAGACTTTCAGGGGCTAGCAGTCTAATTATTATGAATAATTTTAATAACTAGGCC   500
AAATACAAATTTTAGATTCTCATTTAAAAAAACAAACTGACAGGAAAGGTCTGATGAGATAGTTCAGTGGATAAAGCCTCTTGCTGCTAACCTGATGGCC   600
TGAGCTTGATTCCCAGGACCCACGGTGGGAGACCTGACATCACTGACTATGCTCTGACCTCACACATGAGCTTTAGATGTGTATTCCCTCTTCCCACATA   700
CACACAATACACAAATGTAAAAAAAAAAAAAAAAAATCAACAGGTAAAATGATACTTTCTGTACTAAGACTTAATTATTAATTTTATTTGTATACACATAT   800
CAACTATTGAACTCATGGAATCCCCCTATTTCAACTTGTCAAATATAGATGTTAAGCCACCATGCCTGGAATAATCACTAAATTTTAAACAGCTTTTCAT   900
ATACATGGTTGTGTGACTTTTGACAAAGCGTCCAAGACAATTTGATAGAGGAATATAAAGGAATTTGTTCAACAAATGGTGCTGGGAAGCTGGACAGCCA  1000
                                                                                M  V  L  G  S  W  T  A  K
```

```
AAGAGAAAACAACTCCTGCCCATGGCCCCCAAGGAAGCATCATCAAAGCATCTTTGCAGAAAATGCAGAGAAAATCTTCACGAGACAAGCTCAGAGAAGC  1100
 E  K  T  T  P  A  H  G  P  Q  G  S  I  I  K  A  S  L  Q  K  M  Q  R  K  S  S  R  D  K  L  R  E  A
AACAAATTATAAATCTGACTTCATCAAGTTAAGTAATAACATTTGCTCTTGGGAAGCTGTTACTAAGATATTGAAAGTGGCCAGCCAGGTGTGGTGGCAC  1200
 T  N  Y  K  S  D  F  I  K  L  S  N  N  I  C  S  W  E  A  V  T  K  I  L  K  V  A  S  Q  V  W  W  H
ACACCTTTAATCCCAGCACTCAGGAGGCCAACACAGGGGCTCTCTCGGAGCTTAAGGCCAGCATGATCTACATTGTAAGTTCCGGGCCAGTCAGTGCTAC  1300
T  P  L  I  P  A  L  R  R  P  T  Q  G  L  S  R  S  L  R  P  A  *
ACAACAAGACTCTGTCTCAAGAAGCAACTGAACGTAAACAACATATAAATAATAAAGGACTTGTGTGCAGAATAAAAAAGTCCTAGAACTTAAAAAGATG  1400
TTCCAATTTGAAAAATTGGTAGGACATTTGAATAGACCTCTCCCCCTCATATTTAAATAGATACCATGCACTCACAGACACAGACACAGACACAGACACA  1500
GACACACACAGACACACACACACACACACACACACACACACACACACACACAGGAAGCTGGCTGAAATACAAGGAACAGAAAGTCACCGCCACTCCCG  1600
TTGCTTATCTTTCGGCGGTTTCCTTTCTTCTATTGACCACAAGAAATGGGAGGATTTGGAATGTGTGCAGCTGAGTCGGCCTGTGCCTCTGGGCATTTGG  1700
ACACTTCTTGCTGTTTGAGAATTGGCTCTAGAACTTGTGGGCCACCTTGAGGTGTTTTTTTTTTTGTCCCTAATTGATTAGAGAGCCACCACATGGAGCA  1800
AGATACAAAGACACATACTGTGTCCAATACTGATTTGAAAGAGAGCCTATGGGTTTTTTACATTTTGTGGCAGTAGCCTTGCGTATTCTGCTGAGCTTCA  1900
TCCACACAGTTAAAACAGTCTCTCTCTTGACTTAACACCTGATGGAGCTCTGTGTGTACCTATAGGTAGCTCTGTGTTTCTGTTTGTTTAGGGATCTCCA  2000
TTTTCTCTCTACTCTGTCTGCCATCTAGATATTGGGGTGACTAGCCAAACCAGACCAGCACTTTCTCTCCAGCATGCTCTGTGACACATAGCAGCTCCCA  2100
AAGCATTCTACGATACGCCAGTGATGAGCATCCGTGACTCTGACCCTCATTTTGAGATCTGTACAGAGCTGTCCGTTCCTTTCACATCTACAAAAAAAAA  2200
AAA   2203
```

Sp 5

```
CCCCTTCTCCAGCATCATGTCTGCCCACGTGCTACCTACCATGTCAACGGACTAAACCTCTGAAAATGTAAACCAGCCCTAATGATAAGCTTTCCTTTAT   100
AAGAGTTGCTGTGGTCATGGTGTCTCTTCACAACAATAAAAACCCTAAATAAGATGCTATGTATATATGTTTGTGGTGTGTGTTTGTGTGTGTGTTTCTC   200
                                                                M  L  C  I  Y  V  C  G  V  C  L  C  V  C  F  S
TGTGTGTATGTGTGTACATGTATTGTGTGTATATGTACATGTGTGCACATATGCACACATATGGACAACAGCCCTCCACCTCAGGTGTCTTCTTCCCAAA   300
 V  C  M  C  V  H  V  L  C  V  Y  V  H  V  C  T  Y  A  H  I  W  T  T  A  L  H  L  R  C  L  L  P  K
TCCTTTTCTACCTTGTTATTTAAAGCAGGTTTTGGTATAGATACTTGTGTCATACCCAGTTTTTCTATGGGCATGCTAGTTGTCTTCATGGCAAGTATGC   400
S  F  S  T  L  L  F  K  A  G  F  G  I  D  T  C  V  I  P  S  F  S  M  G  M  L  V  V  F  M  A  S  M  L
TACCCACTGAGCCACCTCCCAGCCCCTGGGTTGCTCATTAACACACCCTGCAGTGCCAGTGCCAGTGCCAGTGCCAGTGCCAGTGCCAGTGCCAGTGCCA   500
  P  T  E  P  P  P  S  P  W  V  A  H  *
TCAGCCTCTCCCCGGTGCATTTAGCAGGGCATTTCTGGCATCGCCTCCAGGCCCACTACTTGGATCCAGCCTCTCACAGCCCACTTCTGACATCAAAGCC   600
AAAACAAGAGAGAAAGGGCATTAAGAAGCCAGGGAACAGAAGTAGGTTGATGTGGATATCATTTTCTAATTGTTTTGGGCTTCGCCAGACTTTCTTATAA   700
CCTCGCAAGCTGCCTGCACCCCCATGCCATCCATGAAATGTGCAGTTGCAGCTCCCTTGCTTTGGTGACCTTTGTCCTAGCTATCATGTTTGTGTCCCTT   800
CACACTTTCCCATCTTGCCTAGGCCTTTCTTCTTCCCAACTTTGACCATCTTCAGCCCAGATAAACACTGCCAAAAGAAATCCTTACAACATTCAACCTA   900
CGATCTCCTCGTCCACAGTCATAGCATCACCTCAAGGTGCCCAGTGCTCAAAGGCCCAGTCCGATATTCTAGTATTTCATCTCTCTCTGCTTTTGTCTGT  1000
CTGTCCATCTTTCTTCCTTCCTTCCTTTCTTCCTTTTTTGTTTCTTTTCTTTGTTGTGATTTTGAACTAGGGTCTCACTGTGTTCTCCAGAAGCCCCTGT  1100
TCCTCCATTATAGACATCACCCATCTTAACTTCATGTAATTTTTCATTCTTCCCTTCTCTTCTTAGCAGTAAATCCCGTTTTCCCCAGACCTGAAAACTA  1200
GCAGAGGAATTGCCTTTCCTTGCACTTTCCCTTGCACTTTTCTGAACCAGCGTCAGCTTGTGTTCTGGTGAGCAATGAAGTGTCAGGCCAGTGGAATCCT  1300
CTAAAAGTTGCACAGAAATAAGGCACATGGCTAAAGACACAGAGATTAGACTAGCAGGACAAATGGGTATTATCTGAACACACCAAATATAAAAAAAAAA  1400
AAA   1403
```

*To whom correspondence should be addressed

ACKNOWLEDGEMENT

We thank Mr. K. Saida for technical assistance. This work was supported by the Ministry of International Trade and Industry(Japan) for project of Basic Technology for Future Industry.

REFERENCES

1. Yu, Q., et al. (1987) Nature 326 , 765-769
2, Ishida, N., et al. (1988) Nucleic Acid Res. 16, 3581
3, Ishida, N., et al. ,(1988) J. Mol. Biol. Submitted
4. Shin, H-S., et al., (1985) Nature 317, 445-448
5. Lipman, D. J., et al. (1985) Science 227, 1435-1441
6. Crabtree,G. R. et al. (1985) J. Mol. Biol. 185, 1-19