Peas-Mea1-Ppp2r5d Overlapping Gene Complex: a Transposon Mediated-gene Formation in Mammals

Yasuhide OHINATA,1,2 Shizuyo SUTOU,1 and Youji MITSUI1,2,*

National Institute of Advanced Industrial Science and Technology (AIST), Central 6, Higashi 1-1-1, Tsukuba, Ibaraki 305-8566, Japan1 and Institute of Applied Biochemistry and TARA center, University of Tsukuba, Tsukuba, Ibaraki 305-0006, Japan2

(Received 21 February 2003; revised 20 March 2003)

Abstract

Human and mouse MEA1/Mea1 is flanked by two overlapping genes, a novel PEAS/Peas in a head-to-head orientation and PPP2R5D/Ppp2r5d in a tail-to-tail orientation making a Peas-Mea1-Ppp2r5d overlapping gene complex (PMP-complex). Genomic zoo blot analyses and database searching revealed that Mea1 exists only in mammals, while Peas and Ppp2r5d are conserved in eukaryotes. Mea1 and Peas are transcribed from a testis-expressed bidirectional promoter. Mea1-Ppp2r5d overlapping gene complex (MPOS) contains polyadenylation signals for both genes and shows marked conservation throughout mammals. Furthermore, the MPOS occupies 3′-region of transcripts of both genes is expected to form a clover-like intramolecular structure. Mouse genomic library Screening and database searches identified two MPOS-derived sequences in Odf2 gene and RP23-86H7 cosmid clone, respectively, in which MPOS might be a core segment for the retropositions. Thus, a key role of MPOS, a short transposable element containing polyadenylation signals on both strands, in the formation of the Mea1 during mammalian evolution is suggested.

Key words: overlapping genes; Peas; Mea1; Ppp2r5d; transposable element

1. Introduction

The male-enhanced antigen-1 gene (Mea1) was initially isolated with a polyclonal anti-H-Y antigen antibody as a candidate of serologically detectable H-Y antigen (SDM). It is expressed in the germ cells of mammalian testes, and is of interest from the viewpoints of male-germ cell development and autoimmune disease antigens. Our recent study identified two genes overlapping with Mea1 which occupied the central position. One of the flanking genes is a novel gene named Peas and the other is Ppp2r5d, forming the Peas-Mea1-Ppp2r5d overlapping gene complex (PMP-complex). Ppp2r5d, a mouse homolog of human PPP2R5D encodes the 3′ isoform of the PP2A regulatory B56 family.

PMP-complex is located at human 6p21.1 and mouse chromosome 17 24.30-cM. HLA/H-2 complex loci contain many overlapping genes. To list a few, CYP21 encoding steroid 21-hydroxylase (P450c21) overlaps with TNX, which encodes for an extracellular matrix protein, tenascin-X, in tail-to-tail orientation. CYP21 and TNX share complementary segments of 299-bases at the 3′-ends. On the other hand, ACAT2, encoding the acetyl-CoA acetyltransferase-2 overlaps with TCP1, encoding the t-complex protein-1, a molecular chaperone of the chaperonin family in tail-to-tail orientation. ACAT2 and TCP1 apparently arose during the transition from therapsid reptiles to mammals and have retained their identity for more than 200 million years.

The reverse flow of genetic information from RNA back into DNA is known as retroposition, and each transposed informational element is known as a retroposon. Highly repetitive elements in eukaryotic genomes, such as short interspersed repetitive elements (SINEs, typically < 500 bp), long interspersed repetitive elements (LINEs) with full-length copies ranging from 6 to 8 kb, and processed retropseudogenes belong to this category. Each mammalian order has characteristic families of SINEs that exist as multiple copies (e.g., human Alu, rabbit C, and mouse B1 and B2). However, sequence comparisons show that these SINEs have different origins. Human Alu and mouse B1 repeats probably arose from a 7SL RNA gene, while many SINEs, including rabbit C and mouse B2 repeats, appear to be derived in part from tRNA genes. Although the number of copies is smaller than that of SINEs, many transposable elements are considered to originate via various RNA in-
termediate molecules (e.g., tRNA, rRNA, mRNA, SRP, and hnRNA). In the present study, we have characterized Peas and Mea1 bidirectional transcription and the 110 bp Mea1-Ppp2r5d overlapping segment (hereafter called, MPOS) in the PMP-complex and discuss a possible role of MPOS in transposon-mediated gene creation.

2. Results and Discussion

2.1. Sequence analysis of Ppp2r5d-Mea1-Peas gene complex

By screening a mouse genomic library with a Mea1 cDNA probe, 21 clones were obtained. By restriction mapping, these 21 clones converged to three clones, and the sequence of 17941 bases was determined (DDBJ accession No.: AB074009). Two genes were found to overlap with Mea1 (Fig. 1A). One of them, Ppp2r5d (DDBJ accession No. AB074009) in a tail-to-tail orientation, encoded a mouse homolog of the protein serine/threonine phosphatase 2A (PP2A) regulatory subunit B56 isoform. Ppp2r5d and Mea1 share an overlapping region MPOS (Fig. 1B, box) with poly(A)+ signals (Fig. 1B, shadow) for both Mea1 and Ppp2r5d. Human, bovine, and mouse MPOS flanking sequences are shown in Fig. 1B. Even though the 110-bp MPOS fall in 3′-untranslated regions, the sequences were markedly conserved (identity: bovine/human, 98% and mouse/human, 96%). In contrast, both 3′-untranslated regions flanking MPOS were not as well conserved.

The other gene that overlapped with Mea1 in a head-to-head orientation, was a novel gene named Peas (DDBJ accession No.: human, AB055925; mouse, AB053465) (Fig. 1A). Peas and Mea1 share a bidirectional promoter (Fig. 2A), a GC-rich region roughly spanning 1 kb which contains many AP-2 binding motifs on both strands. The GC-content was over 75% in a 1-kb window (Fig. 1A).

2.2. Mea1 and Peas are inducible by addition of retinoic acid

Northern blot analyses revealed that Peas and Mea1 are dominantly expressed in the testis (Fig. 2B,C). In situ hybridization analyses using anti-sense probes revealed that Peas and Mea1 are dominantly expressed in testicular germ cells, spermatocytes and spermatidetes (Fig. 2D,E). Peas and Mea1 sense probes were also tested as negative controls, but these did not detect any signals (data not shown). The bidirectional promoter for Peas and Mea1 contains many AP-2 binding motifs. AP-2 is known to be inducible by retinoic acid (RA) in teratocarcinoma F9 cells. Furthermore RA signaling is an essential feature of spermatogenesis. In order to determine whether the bidirectional promoter controls transcription of Peas and Mea1 simultaneously, transcription levels of both genes in F9 cells were examined by adding RA or Bt2cAMP (Fig. 2F,G). The transcription levels of Peas and Mea1 increased by 4.7 and 2.1 times, respectively, after 48 hr of RA exposure (Fig. 2G). The addition of Bt2cAMP marginally increased transcription levels of both genes (Fig. 2F).

2.3. Sequence analyses of MPOS and MPOS-like segments

By screening a mouse genomic library with a 32P-labeled Mea1 cDNA probe including MPOS, we isolated several clones and determined their sequences. We found a clone carrying a part of outer dense fiber-2 gene (Odf2) intron that had a segment similar to the 3′-region of Mea1 cDNA with a MPOS-like sequence (Fig. 1C, box) (DDBJ accession No. AB097861). Odf2 is known to be specifically expressed in testis and its products localize in sperm outer dense fibers. A poly(A)+-like sequence was found close to the MPOS-like segment (Fig. 1C, underline). There are direct repeats consisting of 9-bases (GAGATGAGG, Fig. 1C, shadow), which flank the sequence similar to the 3′-region of Mea1 cDNA. These findings suggest that a Mea1 transcript via RNA polymerase II was retropositioned to this position. MPOS might be the core element in the retrotransposition.

Database searches identified another MPOS-like sequence in a cosmId clone, RP33-86H7 (AC068561, nucleic acid nos. 15671–15784) on mouse chromosome 6 (Fig. 1C, box). A poly(A)+-like sequence (Fig. 1C, underline) was found close to the MPOS-like segment (Fig. 1C, underline). There are direct repeats consisting of 9-base (TAT-ACTATA) tandem duplicates (Fig. 1C, shadow, nucleic acid nos. 15615–15658) which flank the sequence homologous to the 3′-region of Ppp2r5d cDNA. These findings suggest that a Ppp2r5d transcript via RNA polymerase II was retropositioned to this site.

We performed Southern blot analyses using probe A (mouse Mea1 exon 2, not containing MPOS) and probe B (mouse Mea1 exon 4, containing MPOS) (Fig. 1A). Probe A detected two bands (9.5 and 8.5 kb) on EcoRI blots (Fig. 3A). Probe B detected four bands (9.5, 8.5, 4.7, and 4.2 kb). The 4.7- and 4.2-kb bands are derived from MPOS-transposed segments. The faint 9.5-kb band that hybridized to probes A and B remains to be identified (Fig. 3A). Genomic zoo blot analysis with full-length mouse Mea1 cDNA probe containing MPOS showed that signals were observed only in mammals (Fig. 3A). These findings indicate that Mea1 and MPOS were probably emerged during mammalian evolution (see also Fig. 1B). The number of copies of Mea1 or MPOS varied in mammalian species, possibly reflecting the retroposition of MPOS. In mouse, an additional unknown 23-kb band was detected.
Figure 1. Structure of the PMP complex. A (top), GC-contents of the PMP complex. The window size is 1 kb. A (bottom), positions and orientation of three overlapping genes in mouse, *Pna*, *Mea1*, and *Ppp2r5d*. The positions of probes (A and B) used for genomic Southern analyses are indicated. B, alignment of the human, bovine, and mouse MPOs is shown as enclosed in a box. *Ppp2r5d* and *Mea1* poly(A)⁺ signals are shadowed. C, MPOs-derived sequence (box) in mouse *Od2* and mouse chromosome 6 RP23-86H7 cosmid. The other repeat (TATACAAT) appears ~500 bp downstream (not shown). The poly(A)⁺-like sequences are underlined. Approximately 1.5 × 10⁶ recombinants from a mouse (129SvJ) AFIX⁺II library (Stratagene, La Jolla, CA) were screened using a 3²P-labeled ~0.9-kb full-length mouse *Mea1* cDNA fragment as a probe. Positive plaques were purified through two rounds of re-screening. The genomic DNA from several phage clones were mapped, restricted, and subcloned into a pBluescript SK(−) vector (Stratagene).
Figure 2. Gene expression of Peas and Mea1. A, a schematic illustration of mouse Peas and Mea1 genomic structure; Mea1 and Peas are transcribed from multi-points in the less than 2.5 kb GC-rich region (mostly from those within the 1-kb region). B, C, Northern blot analyses of Peas (B) and Mea1 (C) showing their expression enriched in testis. The size markers for B and C are 0.5, 1, 1.5, 2, 2.5, 3, 4, 5, 6, and 9 kb from the bottom to top. A Northern membrane loaded with 2 µg of poly(A)^+ RNA from a variety of mouse tissues was purchased from Ambion (Austin, TX). The membrane was hybridized with a ^32P-labeled full-length coding fragment of mouse Mea1 or Peas for 16 hr at 45°C in a hybridization solution containing 50% (v/v) formamide. The membrane was washed under stringent conditions. D, In situ hybridization analysis of Peas showing its expression in testicular germ cells (Magnification, ×200). E, In situ hybridization analysis of Mea1 showing its expression in testicular germ cells (Magnification, ×200). Paraffin-sections (3 µm thick) of the mouse (BALB/c; age 8 weeks) testis were hybridized with anti-sense and sense digoxigenin (DIG)-11-UTP-labeled RNA probes (∼500 bases). Immunohistochemical detection of DIG-labeled RNA was carried out with alkaline phosphatase-conjugated anti-DIG antibody (Roche) and an alkaline phosphatase substrate kit VI (BCIP/NBT) (Vector, Burlingame, California). Counterstaining was performed with 0.1% safranin O. F, Northern blot analyses of Peas and Mea1 transcripts in F9 cells treated with RA or Bt_2cAMP. Mouse F9 cells were maintained in DMEM (Sigma, St. Louis, MO) containing 10% (v/v) FBS. For a differentiation study, F9 cells were incubated with 1 µM all-trans-retinoic acid (Sigma) or 1 mM dibutyl cyclic AMP (Bt_2cAMP, Wako, Osaka, Japan) for up to 48 hr. Northern blot analyses were performed using ^32P-labeled GAPDH cDNA fragments for an internal control. Relative transcription levels were determined with the NIH Image software. G, time courses of Peas (squares) and Mea1 (circles) transcripts induced in F9 cells by RA (line) and Bt_2cAMP (broken line).
Figure 3. Conservation of Mea1-like sequences. A, EcoRI-digested mouse (ICR) genomic DNA was analyzed by Southern hybridization with probe A (not containing MPOS) and with probe B (containing MPOS) (Fig. 1A). B, Genomic zoo blot analyses of Mea1. EcoRI-digested genomic DNAs from various animals were analyzed with a full-length mouse Mea1 cDNA probe. A genomic zoo-blot membrane loaded with 10 µg of EcoRI-digested genomic DNA in each lane was purchased from Seegene (Seoul, Korea). The membrane was treated with 32P-labeled mouse Mea1 coding fragment for 16 hr at 55°C in a hybridization solution (0.5 M Na2HPO4-H3PO4 (pH 7.2), 1 mM EDTA, 7% (v/v) SDS). The membrane was washed under stringent conditions. Mouse (ICR) genomic Southern membranes loaded with 10 µg of EcoRI-digested DNA in each lane were treated with probe A and B (Fig. 1A), 32P-labeled DNA fragments of the mouse Mea1 exon 2 and 3′-noncoding region, respectively, for 16 hr at 45°C. The membranes were washed under stringent conditions.

2.4. Modeling of secondary structures of 3′-regions of Mea1 and Ppp2r5d transcripts

Short transposable elements derived from RNA species can generally form intramolecular structures.9–13 Since MPOS contains palindromic poly(A)+ signals and complementary sequences, it may form an intramolecular structure at each 3′-terminal of Mea1 and Ppp2r5d transcripts. Secondary structure modeling showed that portions of Mea1-Ppp2r5d and Ppp2r5d-Mea1 mRNA transcripts could form clover-like structures that are complementary mirror images (Fig. 4A,B). The same modeling was also applied to tail-to-tail overlapping genes in HLA/H-2 loci (ACAT2-TCP1 and CYP21-TNX) (Fig. 4C,D). Although we could not identify the original RNA species of MPOS, the ability to form clover-like structures suggests that MPOS was derived from an RNA species.

2.5. Formation mechanisms of PMP-complex

Mea1 is found only in mammalian genomes and Peas and Ppp2r5d are genetically conserved in eukaryotes. Rabbit C-repeats can provide genes with functional poly(A)+ signals.19 Similarly, MPOS supplies poly(A)+ signals for both Mea1 and Ppp2r5d. It is therefore considered that an original RNA molecule of MPOS had retropositioned to the present position, the far, opposite side of the potential-bidirectional Peas promoter, and Mea1 was created in a mammalian ancestor (Fig. 5). The sequences of MPOS are tightly conserved throughout mammals (Fig. 1C, box), suggesting conservation of the original sequence. The origin of MPOS was thought to be an RNA species, but its original molecule has not been identified. PP2A regulatory B56 family is encoded by Ppp2r5a-e genes. Ppp2r5c and Ppp2r5d are closely related.4 The genomic region of Ppp2r5c contains a transposon called intracisternal A-particle (IAP) (AB026817). A pseudogene of Ppp2r5c also exists.20 Ppp2r5d have been duplicated from Ppp2r5c and retropositioned to the present position (Fig. 5). Since short transposable elements derived from small RNA molecules do not encode reverse transcriptase themselves, they need to take advantage of other reverse transcriptases. MPOS is a short transposable element so that could not move alone but...
could have transposed together with \textit{Ppp2r5d}. Our study supports the idea that MPOS transposition created a new mammalian-specific gene, \textit{Mea1}, and may have contributed to mammalian evolution.

References

1. Lau, Y-F., Chan, C. K., and Sparkes, R. 1989, Male-enhanced antigen gene is phylogenetically conserved and expressed at late stages of spermatogenesis, \textit{Proc. Natl. Acad. Sci. U.S.A.}, 86, 8462–8466.

2. Sutou, S., Kondo, M., Matsuda, M. et al. 2001, H-Y antigens as Y chromosome-encoded gene products and serologically detectable male antigens (SDM) as testis- or spermatogenesis-linked autosomal gene products, \textit{Arch. Tierz.}, 44, 11–21.

3. McCright, B. and Virshup, D. M. 1995, Identification of a new family of protein phosphatase 2A regulatory subunits, \textit{J. Biol. Chem.}, 270, 26123–26129.

4. McCright, B., Rivers, A. M., Audlin, S., and Virshup, D. M. 1996, The B56 family of protein phosphatase 2A (PP2A) regulatory subunits encodes differentiation-induced phosphoproteins that target PP2A to both nucleus and cytoplasm, \textit{J. Biol. Chem.}, 271, 22081–22089.

5. Tee, M. K., Thomson, A. A., Bristow, J., and Miller, W. L. 1995, Sequences promoting the transcription of the human XA gene overlapping P450c21A correctly predict the presence of a novel, adrenal-specific, truncated form of tenasin-X, \textit{Genomics}, 28, 171–178.

6. Speek, M., Barry, F., and Miller, W. L. 1996, Alternate promoters and alternate splicing of human tenasin-X, a gene with 5' and 3' ends buried in other genes, \textit{Hum. Mol. Genet.}, 11, 1749–1758.

7. Shintani, S., O'hUigin, C., Toyosawa, S., Michalová, V., and Klein, J. 1999, Origin of gene overlap: the case of TCP1 and ACAT2, \textit{Genetics}, 152, 743–754.

8. Singer, M. F. 1982, SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes, \textit{Cell}, 28, 433–434.

9. Ullu, E. and Tschudi, C. 1984, Alu sequences are processed 7SL RNA genes, \textit{Nature}, 312, 171–172.

10. Deininger, P. L. and Daniels, G. R. 1986, The recent evolution of mammalian repetitive DNA elements, \textit{Trends Genet.}, 2, 76–80.

11. Daniels, G. R. and Deininger, P. L. 1985, Repeat sequence families derived from mammalian tRNA genes, \textit{Nature}, 317, 819–822.

12. Ohshima, K., Hamada, M., Terai, Y., and Okada, N. 1996, The 3' ends of tRNA-derived short interspersed repetitive elements are derived from the 3' ends of long interspersed repetitive elements, \textit{Mol. Cell. Biol.}, 16, 3756–3764.

13. Brosius, J. 1999, RNAs from all categories generate retrosequences that may be excerpted as novel genes or regulatory elements, \textit{Gene}, 238, 115–134.

14. Ohinata, Y., Sutou, S., Kondo, M., Takahashi, T., and Mitsui, Y. 2002, Male-enhanced antigen-1 gene flanked by two overlapping genes is expressed in late spermatogenic cells con- structed from overlapping genes is expressed in late spermatogenesis, \textit{Biol. Reprod.}, 67, 1824–1831.

15. Luscher, B., Mitchell, P. J., Williams, T., and Tjian, R. 1989, Regulation of transcription factor AP-2 by the morphogen retinoic acid and by second messengers, \textit{Genes Dev.}, 3, 1507–1517.

16. Lufkin, T., Lohnes, D., Mark, M. et al. 1993, High postnatal lethality and testis degeneration in retinoic acid receptor alpha mutant mice, \textit{Proc. Natl. Acad. Sci. U.S.A.}, 90, 7225–7229.

17. van Pelt, A. M. and de Rooij, D. G. 1991, Retinoic acid is able to reinitiate spermatogenesis in vitamin A-deficient rats and high replicate doses support the full development of spermatogenic cells, \textit{Endocrinol.}, 128, 697–704.

18. Hoyer-Fender, S., Petersen, C., Brohmann, H., Rhee, K., and Wolgemuth, D. J. 1998, Mouse \textit{Odf2} cDNAs consist of evolutionary conserved as well as highly variable sequences and encode outer dense fiber proteins of the sperm tail, \textit{Mol. Reprod. Dev.}, 51, 167–175.

19. Krane, D. E. and Hardison, R. C. 1990, Short interspersed repeats in rabbit DNA can provide functional polyadenylation signals, \textit{Mol. Biol. Evol.}, 7, 1–8.

20. Muneer, S., Ramalingam, V., Wyatt, R., Schultz, R. A., Minna, J. D., and Kamiba- yashi, C. 2002, Genomic organization and mapping of the gene encoding the PP2A B56gamma regulatory subunit, \textit{Genomics}, 79, 344–348.