Genomic structure of the rat major AP endonuclease gene (Apex) with an adjacent putative O-sialoglycoprotease gene (Prsmg1/Gcpl1) and a processed Apex pseudogene (Apexp1).

Ming Yao* Kosuke Akiyama† Yunshan Tan‡
Altaf Hossain Sarker** Shogo Ikeda†† Shahjalal Shafiul Alam‡‡
Ken Tsutsui§ Michihiro C. Yoshida¶ Shuji Seki∥

*Okayama University, †Okayama University, ‡Okayama University,
**Okayama University, ††Okayama University of Science, ‡‡Okayama University,
§Okayama University, ¶Hokkaido University, ∥Okayama University,

Copyright ©1999 OKAYAMA UNIVERSITY MEDICAL SCHOOL. All rights reserved.
Genomic structure of the rat major AP endonuclease gene (Apex) with an adjacent putative O-sialoglycoprotease gene (Prsml/Gcpl1) and a processed Apex pseudogene (Apexp1).

Ming Yao, Kosuke Akiyama, Yunshan Tan, Altaf Hossain Sarker, Shogo Ikeda, Shahjalal Shafiul Alam, Ken Tsutsui, Michihiro C. Yoshida, and Shuji Seki

Abstract

Genomic sequencing and chromosomal assignment of the gene encoding rat APEX nuclease, a multifunctional DNA repair enzyme, were performed. An active Apex gene and a processed pseudogene were isolated from a rat genomic library. The active Apex gene consists of 5 exons and 4 introns spanning 2.1 kb. The putative promoter region of the Apex gene lacks the typical TATA box, but contains CAAT boxes and a CpG island having putative binding sites for several transcription factors, such as Sp1, AP-2, GATA-1 and ATF. A putative O-sialoglycoprotease (a homologue of Pasteurella haemolytica glycoprotease, gcp; abbreviated as Prsml/Gcpl1) gene consisting of 11 exons and 10 introns spanning 7.3 kb lies immediately adjacent to the Apex gene in a 5’-to-5’ orientation. The Apex gene locus was mapped to rat chromosome 15p12 using in situ hybridization. The processed pseudogene (designated as rat Apexp1) has a nucleotide sequence 87.1% identical to that of the rat Apex cDNA, although several stop codons interrupting the coding sequences and multiple nucleotide deletions were observed. The Apexp1 is located in an inactive LINE sequence. Calculation of nucleotide substitution rates suggests that the immediate, active progenitor of Apexp1 arose 23 million years ago and that the non-functionalization occurred 15 million years ago.

KEYWORDS: apurinic, apyrimidinic endonuclease, glycoprotease, Aprx pseudogene, genomic sequencing, chromosomal mapping

*PMID: 10631378 [PubMed - indexed for MEDLINE]

Copyright (C) OKAYAMA UNIVERSITY MEDICAL SCHOOL
Genomic Structure of the Rat Major AP Endonuclease Gene (Apex) with an Adjacent Putative O-Sialoglycoprotease Gene (Prsmg1/Gcp11) and a Processed Apex Pseudogene (Apexp1)

Ming Yao$^a$, Kosuke Akiyama$^a$, Yunshan Tan$^a$, Altaf Hossain Sarker$^a$, Shogo Ikeda$^b$, Shahjalal Shafiul Alam$^b$, Ken Tsutsui$^a$, Michihiro C Yoshida$^c$ and Shuji Seki$^{a,d,e}$

$^a$Department of Molecular Biology, Institute of Cellular and Molecular Biology, Okayama University Medical School, Okayama 700-8558, Japan, $^b$Department of Biological Chemistry, Okayama University of Science, Okayama 700-0005, Japan, $^c$Chromosome Research Unit, Faculty of Science, Hokkaido University, Sapporo 060-0180, Japan, $^d$Department of Human Nutrition, Chugoku Junior College, Okayama 701-0197, Japan

Genomic sequencing and chromosomal assignment of the gene encoding rat APEX nuclease, a multifunctional DNA repair enzyme, were performed. An active Apex gene and a processed pseudogene were isolated from a rat genomic library. The active Apex gene consists of 5 exons and 4 introns spanning 2.1 kb. The putative promoter region of the Apex gene lacks the typical TATA box, but contains CAAT boxes and a CpG island having putative binding sites for several transcription factors, such as Sp1, AP-2, GATA-1 and ATF. A putative O-sialoglycoprotease (a homologue of Pasteurella haemolytica glycoprotease, gcp; abbreviated as Prsmg1/Gcp11) gene consisting of 11 exons and 10 introns spanning 7.3 kb lies immediately adjacent to the Apex gene in a 5'-to-5' orientation. The Apex gene locus was mapped to rat chromosome 15p12 using in situ hybridization. The processed pseudogene (designated as rat Apexp1) has a nucleotide sequence 87.1% identical to that of the rat Apex cDNA, although several stop codons interrupting the coding sequences and multiple nucleotide deletions were observed. The Apexp1 is located in an inactive LINE sequence. Calculation of nucleotide substitution rates suggests that the immediate, active progenitor of Apexp1 arose 23 million years ago and that the non-functionalization occurred 15 million years ago.

Key words: apurinic/apyrimidinic endonuclease, glycoprotease, Apex pseudogene, genomic sequencing, chromosomal mapping

APEX nuclease (also called BAP1, HAP1, APE, Ref-1 and APE1) is a multifunctional DNA repair enzyme having apurinic/apyrimidinic (AP) endonuclease, 3'-5' exonuclease, DNA 3' repair diesterase, DNA 3'-phosphatase and RNase H activities (1-3). It is also known to be a redox factor (Ref-1), stimulating DNA binding activity of AP-1 binding proteins such as Fos and Jun (4). cDNAs for APEX nuclease have been cloned from bovine, human, mouse and rat cDNA libraries (reviewed in Refs. 2 and 3). Genomic cloning and sequencing have also been reported for human APEX and mouse Apex genes (5-8). The Apex/Ref-1 disruption in germ cells leads to lethality during embryogenesis (9), and the establishment of Apex/Ref-1 (-/-) cell lines has not yet been reported (10). These results suggest that APEX nuclease performs multiple functions in living cells and is essential for cell survival. To analyze the biological function of APEX nuclease, we studied developmental changes and tissue variations of Apex gene expression in mice and rats by in situ and northern blot hybridizations (11, 12). In order to further clarify the genetic regulation and biological function of APEX nuclease, we cloned the Apex gene of the rat, an important model organism in biomedical research. During this study, we noticed the presence of genomic DNA restriction fragments that were hybridizable with labeled rat Apex cDNA by Southern blot analysis but were not derived from the active Apex...
gene. To investigate the origin of the unexpected restriction fragments, we screened intensively Apex-related genes in the rat genomic library, and cloned an Apex pseudogene in addition to an active Apex gene.

This paper describes the genomic structure, sequence and chromosomal localization of the rat Apex gene along with those of an adjacent Prsngl/Gepl gene [a gene for a putative metalloglycoprotease or for gcp (Pasteurella haemolytica glycoprotein)-like 1 (EC3.4.24.57)]. This paper also demonstrates the presence of an inactive, processed Apex pseudogene in the rat genomic DNA.

**Materials and Methods**

**Materials.** The reagents used in these experiments were obtained from the following sources: restriction enzymes from Toyobo Biochem. (Osaka, Japan); T4 DNA ligase from Promega Corp. (WI, USA); rat genomic library constructed using Sprague-Dawley/C adult rat DNA and cloning vector EMBL-3 SP6/T7 from Clontech Laboratories, Inc. (CA, USA); Mega-prime™ DNA Labeling System, [α-32P] dCTP (3,000 Ci/mmol), and positively charged nylon membrane (Hybond-N+) from Amersham Japan (Tokyo, Japan); PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit from PE Applied Biosystems (CA, USA); Blood and Cell Culture DNA Maxi Kit, QIAGEN Lambda Mini Kit, QIAprep Spin Miniprep Kit and QIAquick Gel Extraction Kit from QIAGEN K.K. (Tokyo, Japan). The other materials used were obtained as described previously (13).

**Cloning of the rat Apex gene and its pseudogenes.** Recombinant bacteriophage clones in a rat genomic library in bacteriophage vector EMBL-3 SP6/T7 (Clontech) were plated with Escherichia coli K802. Approximately 1 × 10^8 bacteriophage clones were screened using the 32P-labeled rat Apex cDNA probe by the plaque hybridization technique (14). A 32P-labeled probe was prepared using the EcoRI-SmaI fragment from the rat Apex cDNA clone pUC18-RAPEX80 (12) as a template and [α-32P]dCTP (3,000 Ci/mmol) by the random priming labeling method. Prehybridization was carried out at 68°C in 6 × SSC, 150 mM NaCl and 15 mM sodium citrate, pH 7.0, 5 × Denhardt’s reagent (1 × Denhardt’s reagent: 0.02% Ficol, 0.02% polyvinylpyrrolidone and 0.02% bovine serum albumin), 0.5% sodium dodecylsulfate (SDS) and 100 μg/ml denatured salmon testes DNA for 4 h. Hybridization was carried out at 68°C in the same solution containing the 32P-labeled cDNA probe for 18 h. The membrane was washed at room temperature twice in 2 × SSC and 0.1% SDS for 15 min, and then twice at 68°C in 1 × SSC and 0.1% SDS for 30 min. After the washing, the membrane was air dried and autoradiographed at −80°C using Fuji X-ray film with an intensifying screen. Screening was repeated 3 times to isolate positive isolated clones. The recombinant phage DNA was isolated from the positive clones by the plate lysate method. The insert DNAs of the positive clones were subcloned into plasmid vector pUC18 and subjected to DNA sequencing.

**DNA sequencing and analysis of the rat Apex gene.** DNA sequences were determined on both strands by the dideoxy chain termination method using double-stranded templates and modified AmpliTaq polymerase (ABI PRISM Cycle Sequencing Kit from PE Applied Biosystems) and by an ABI 373S DNA Sequencer (Perkin-Elmer Japan Co., Ltd., Chiba, Japan). Sequencing primers used were the M13 universal primer and specific oligonucleotide primers synthesized according to the sequences determined. The CpG islands (CpG-rich regions) were defined according to the criteria reported by Gardiner-Garden and Frommer (15).

**Southern blot analysis of rat genomic DNA.** High molecular weight DNA was isolated from rat liver by using a Blood and Cell Culture DNA Maxi Kit (QIAGEN). Ten micrograms of DNA digested with a restriction enzyme were electrophoresed on a 0.7% agarose gel in the standard electrophoresis buffer containing 89 mM Tris-HCl, 89 mM boric acid and 2 mM EDTA (pH 8.0), and transferred to a nylon membrane (Hybond-N+) by the capillary transfer technique. The blotted membrane was hybridized with 32P-labeled rat Apex cDNA probe prepared as described above. Prehybridization was carried out at 68°C in 6 × SSC, 5 × Denhardt’s reagent, 0.5% SDS and 100 μg/ml denatured salmon testes DNA for 4 h, and then hybridization was carried out at 68°C in the same solution containing 32P-labeled cDNA probes (2 mg/ml) for 20 h. The membrane was washed twice at room temperature in 2 × SSC and 0.1% SDS for 15 min, twice at 68°C in 1 × SSC and 0.1% SDS for 30 min, and then twice at 68°C in 0.1 × SSC and 0.1% SDS for 5 min. After the washing, the membrane was air dried and autoradiographed at −80°C using Fuji X-ray film with an intensifying screen for 4 weeks.

**Fluorescence in situ hybridization (FISH).**
Air-dried chromosome preparations were made from fibroblasts derived from primary skin cultures of a WKAH strain rat by standard methods. Chromosomal DNA was denatured by treatment in 70% formamide, 2× SSC for 3 min at 75°C. In situ hybridization with a biotinylated rat probe was performed according to a modification of the method of Kondoh et al. (16). Briefly, the DNA probe (clone A; 13.4 kb) labeled with biotin-16-dUTP (Boehringer Mannheim, Australia) by nick translation was coprecipitated with sonicated salmon sperm DNA and E. coli tRNA, dissolved in 100% formamide, denatured for 10 min at 75°C, and then resuspended in a hybridization mixture containing 4× SSC, 0.2% bovine albumin, 10% dextran sulfate, and 50% formamide.

After hybridization to metaphase spreads and incubation with anti-biotin Fab'-alkaline phosphatase (ALP) conjugate (1:50) for 45 min at 37°C, biotinylated probes were detected with the HNPP-azo dye solution by the method previously described (16). Chromosomes were counterstained with Hoechst-quinacrine (17) and mounted in MacIlvaine's buffer (pH 4.5) for simultaneous observations of fluorescent signals and G-bands with a Nikon fluorescence microscope equipped with a BV-2A filter set (excitation wavelength, 400–440 nm).

We have used ISCN terminology (18) in describing the chromosomes of the RNO (Rattus norvegicus) for rat chromosomes and numbered the rat G-bands according to the system of Satoh et al. (19).

Results and Discussion

Isolation of Apex-related genomic clones. A rat genomic library was screened to isolate Apex genomic clones as described in Materials and Methods. Fourteen recombinant clones hybridizable to the rat Apex cDNA probe were isolated. We selected a clone containing the whole region of the Apex gene with a long 5’ flanking sequence for further analysis of the active gene. The insert DNA was subcloned into plasmid vector pUC18, and the clone was designated as clone A. Three independent clones having an identical processed pseudogene (Apexp1) were also subcloned into plasmid vector pUC18 (clones 2, 21 and 25). Among them, the clone 21 was selected for further analysis, because it contained an insert (8.2 kb) long enough to analyze the pseudogene with the 5’- and 3’-flanking regions.

Sequencing and analysis of the rat Apex gene. The sequence of the clone A insert (13.4 kb) covered the entire Apex gene with 9.9 kb, 5’ flanking and 1.4 kb, 3’ flanking regions. By comparing this sequence with the published Apex cDNA sequence (12), we defined the exon/intron structure of the rat Apex gene (Fig. 1; accession no. AB023065). The gene consists of 5 exons and 4 introns, with a total length of 2.1 kb. The nucleotide sequence of the exons corresponds exactly to the rat cDNA sequence (12). All exon-intron junctions follow the GT/AG rule (20). The translation initiation and termination codons (ATG and TGA) are located in exon II and exon V, respectively. In order to tentatively assign a base position within the gene and in the flanking regions, the adenine residue in the translation initiation codon of the Apex gene was designated as position +1 and the 5’ neighboring residue guanine as position −1. With this as a starting point, positive numbers were given to positions downstream from the starting point and negative numbers to those upstream. The splicing sites in the amino acid codons are completely conserved as in the mouse and human genes (7, 8). The putative promoter region of the Apex gene lacks the typical TATA box, but contains CAAT boxes and a CpG island having putative binding sites for several transcription factors, such as Sp1, AP-2, GATA-1 and ATF.

Genes adjacent to the Apex gene. The sequence analysis of the 9.9 kb, 5’ flanking region of the Apex gene revealed the presence of a gene coding for a putative glycoprotein homologous to Pasteurella haemolytica glycoprotein (gcp) (accession nos.: U15958, M62364) (21). The known substrates for the Pasteurella haemolytica glycoprotein are glycoproteins with extensive clusters of negatively charged sugar residues such as O-sialylglycan or sulfoglycan conjugates, including human cell surface glycoproteins (glycophorin A, CD24, CD34, CD43, CD44 and CD45), ligands for P- and L-selectins, the tumor antigen epitec, the vascular adhesion protein VAP-1, platelet glycoprotein Ib, and cranin, a brain O-sialylglycoprotein (22). Based on the information reported for the prokaryotic gcp enzyme and the information found in EST Data Bases, the gene encoding the putative mammalian homologue of the gcp enzyme is tentatively designated as a metalloproteinase (EC3.4.24.57) for glycoprotein (Prsmgl) or gcp (Pasteurella haemolytica glycoprotein)-like 1 (Gcpl1). By comparing the 5’ flanking sequence of the rat Apex gene with mammalian EST-sequences (accession nos.: AA051631, AA105521, AA401938, AA572107, AA822991, AF052137, H31957) homologous to the Pasteurella haemolytica
glycoprotease, we defined the exon/intron structure of the rat Prsmg1/Gcpl1 gene (Fig. 1; accession no. AB023065). The gene consists of 11 exons and 10 introns spanning 7.3 kb. All the exon-intron junctions follow the GT/AG rule (20). The Prsmg1/Gcpl1 gene lies immediately adjacent to the Apex gene in a 5'-to-5' orientation. The closest exons for the Apex and Prsmg1/Gcpl1 genes, the exon 1 of the Apex gene and the exon 1 of the Prsmg1/Gcpl1 gene, lie within a 366 bp sequence. The CpG island is located between the introns 1 of the Apex gene and of the putative Prsmg1/Gcpl1 gene (Fig. 1). The rat Prsmg1/Gcpl1-coding sequence (accession no. AB023065) deduced from the exons appears to code an enzyme consisting of 335 amino acids with a calculated molecular weight of 36,356. The predicted amino acid sequence of the rat Prsmg1/Gcpl1 protein conserves a putative zinc-binding site (His-109 and His-113).

We also found putative human PRSWM1/GCPL1 and mouse Prsmg1/Gcpl1 genes in homologous, 5' flanking regions of the human APEX and mouse Apex genes on chromosomes 14q11.2-q12 and 14C2-D1, respectively. The putative rat Prsmg1/Gcpl1 protein shows 31.8% identity in 321 amino acid residues to the Pasteurella haemolytica glycoprotease (accession nos.: U15958, M62364), 94.9% identity in 335 residues to the putative human PRSWM1/GCPL1 protein, and 97.6% identity in 335 residues to the putative mouse Prsmg1/Gcpl1 protein (unpublished results).

To investigate consensus sequences present commonly in the probable promoter regions of the rat, mouse and human major AP endonuclease genes, the genomic sequences of these APEX (or Apex) genes were compared in 5' terminal and 5' flanking regions (Fig. 2). A CAAT box and a GC box (Sp1 binding site) are equally located in the intervening sequences between these APEX (or Apex) and PRSWM1/GCPL1 (or Prsmg1/Gcpl1) genes. The intervening sequences are short and quite similar between the rat and mouse. The intervening human sequence is relatively long, fairly different from those of the rat and mouse, and contains Alu-like sequences as described previously (8). The promoters of the Apex and Prsmg1/Gcpl1 genes are probably overlapped, based on the fact that the intervening sequence is short and that CAAT and GC boxes having characteristics that allow them to function in either orientation are equally present.
The sequence analysis using EST data bases (accession no.: H31707) of the 3' flanking region (1.4 kb) of the Apex gene in the clone A indicated the presence of the last exon of an unidentified gene laid in a 3'-to-3' orientation (Fig. 1). The polyadenylation signal (AATAAA) of the 3'-adjacent gene is located at position 1908/1903 (accession no.: AB023065).

**Chromosome mapping of the Apex gene.**
A total of 10 metaphases were examined. All the metaphases had twin spots on one homologue at region p12 of chromosome 15, indicating that the rat Apex gene is located on RNO15p12. A representative example of the FISH mapping is shown in Figure 3. Although no obvious morphological similarity is detectable between RNO15 and MMU14, on which the mouse Apex gene is located, the present mapping demonstrates that the Apex gene is a member of a conserved syntenic group between rat 15 and mouse 14, such as nucleoside phosphorylase on RNO15 and MMU14 (19.50 map units).

After finishing the present experiments, we noticed a report of Johansson et al. (23) on cytogenetic localization of rat cancer-related genes including two Apex-related genes which were designated Apex1 and Apex2. These Apex-related genes have not yet been cloned or sequenced. However, Johansson et al. reported that there appear to be two loci for the Apex gene in the rat, one located on RNO4q12 and the other on RNO15p14. Based on previous reports that the chromosomal localizations of mouse Apex and human APEX genes are MMU14C3 and HSA14q11.2-q12, respectively, and that MMU14 shows homology to part of RNO15, they described that the gene located on RNO15 appears to be the Apex1 gene, homologous to the active human and mouse genes. Further, they suggested without any sequence information that the Apex2 gene on RNO4q12 might represent a pseudogene. As described above, we assigned the rat Apex gene to RNO15p12 after cloning and sequencing it. We could not detect any significant hybridization signal on RNO4. The difference between our results and those of Johansson et al. is likely to be due to the difference in the probes used for in situ hybridization. Johansson et al. (23) used the rat Apex cDNA cloned by us (12), whereas we used the rat Apex genomic clone A as the probe. Our present results agree with theirs in that the active Apex gene is located on rat chromosome 15. It is not clear, however, why their location (RNO15p14) for Apex1, the active Apex gene described in their study, was somewhat different from the location (RNO15p12) of the active Apex gene in the present paper. We suggest that their Apex2 gene on RNO4q12 is not an active Apex gene, as will be discussed later.

**Characterization of the rat Apex pseudogene 1 (Apexp1).** The clone 21 having an 8.2 kb long fragment encompassing the Apex pseudogene (Apexp1) was sequenced and analyzed precisely (the sequence is deposited in Genome Data Bases with the accession no. AB023066). In order to assign tentatively a base position within the insert of the clone 21, the guanine residue at the 5' end of the insert sequence on which the Apexp1 lies was designated as position +1. With this as a starting point, positive numbers were given to positions downstream.

Apexp1 is a processed pseudogene comprising almost the entire sequence of rat Apex cDNA, from its exon 1 to the polyA signal (AATAAA), with the exception of the polyA tail. Although the Apexp1 sequence is 87.1% identical to the Apex cDNA, in-frame termination codons and deletions or insertions that change the reading frame preclude the possibility that this pseudogene encodes a functional APEX nuclease. The nucleotide sequence encompassing Apexp1, which is located at base positions from 1837 to 3026 in the insert of the clone 21, is related to long interspersed repetitive elements (LINEs) (about 74% identical nucleotides/4350 bp to a LINE3 element—accession no. M13100), with the exception of the sequences from 1 to 780 nt and from 6231 to 8236 nt in the insert, which have no significant sequence homologies with LINEs (Fig. 4; accession no. AB023066). Apexp1 is located in a LINE element in the opposite direction and
flanked by short direct repeats (such as CATTTC, TTTTTC and TCTTTCAG) which might be involved in the retrotransposition of the gene. The LINE element is thought to be functionally inactive due to the presence of in-frame termination codons and deletions or insertions that change the reading frame.

Processed pseudogenes have been described for numerous housekeeping genes in mammals (24). In terms of Apex-related genes, however, the presently described rat Apex1 is the first pseudogene reported at the nucleotide level in vertebrates, although Johansson et al. (23) recently suggested the presence of a pseudogene in their cytogenetic study. Pseudogenes are formed during the process of evolution. The extent of nucleotide sequence divergence from a functional mRNA can be regarded as a quantifiable interval in an evolutionary clock.

To estimate the time elapsed since the divergence of Apex and Apex1, we aligned the rat Apex1 sequence with those from rat Apex cDNA (accession no.: D44495) (12) and human APEX cDNA (accession no.: D90373) (25) and calculated the rate of nucleotide substitutions for each position of the codons, according to the method reported previously (26, 27) (Table 1). According to the computation from Li et al. (26), which assumes different substitution rates for inactive pseudogenes and functional genes, it can be estimated that the immediate, active progenitor of the Apex1 gene diverged from the Apex gene about 23 million years ago and that non-functionalization occurred 15 million years ago. Therefore, it appears that Apex1 is a late pseudogene that must be limited to a few species of rodents.

We have attempted to localize the chromosome of the Apex1 gene using the 8.2 kb Apex1 genomic fragment, but so far have failed to localize it probably due to the presence of the repetitive LINE sequence in the genomic fragment. The Southern blot analysis indicated the presence of a single active Apex gene and a single Apex pseudogene, as described below (Fig. 4). Combining our present results and the results of Johansson et al. (23), we suggest that the Apex2 of Johansson et al. on RNO4q12 may be Apex1.

**Southern blot analysis.** To analyze the existence of Apex-related genes in the rat genome, Southern blot analysis of rat genomic DNA was performed using the 32P-labeled, rat Apex cDNA probe and several restriction enzymes (EcoRI, PstI, HindIII, XbaI, NcoI, BamHI and ScaI) (Fig. 4). Restriction fragment lengths of genomic DNA containing the true Apex gene or a part

---

**Fig. 4** Southern blot analysis and restriction maps of the rat Apex gene and its pseudogene. (A) Southern blot analysis of genomic DNA from the rat liver. Rat genomic DNA was digested with the restriction enzyme indicated on the top. Ten micrograms of DNA per lane were subjected to electrophoresis on 0.7% agarose gel and then transferred to a nylon membrane. The blots were hybridized with the 32P-labeled rat Apex cDNA. Numbers in the left column indicate size markers in kilobase pairs. Bands corresponding to the Apex gene and its pseudogene (Apex1) are indicated by stars and circles, respectively. (B, C) Organization and restriction maps of the isolated genomic clone A (B) containing the Apex gene and clone 21 (C) containing the Apex1 gene. Recognition sites for restriction enzymes BamHI (B), EcoRI (E), HindIII (H), NcoI (N), PstI (P), ScaI (S) and XbaI (X) are shown over the genes. (B) Five exons of the Apex gene in clone A are indicated by solid boxes. The Arabic numeral under each exon indicates the exon number. (C) The Apex1 sequence in clone 21 is indicated by the shaded box. Arrows under the diagram of the clone 21 indicate the coding directions of the progenitor genes (Apex and LINE).
of the gene, and those of genomic DNA containing Apex1 (a pseudogene) or a part of the gene, were deduced from the 13.4 kb sequence of the clone A and the 8.2 kb sequence of the clone 21, respectively (Fig. 4). The presence of both a single Apex gene and a single Apex1 gene in the rat genome was confirmed by the Southern blot analysis. Almost all the hybridizing fragments detected by the Southern blot experiment corresponded to the restriction enzyme fragments deduced from the sequences of the Apex and Apex1 genes, although minor unidentified fragments were also observed (Fig. 4). The origins of the unidentified fragments are not known. They may indicate the presence of a gene(s) distantly related to the Apex gene, because the unidentified fragments show only weak hybridization signals.

In summary, we performed genomic cloning, sequencing and chromosomal assignment of Apex-related genes. A single active Apex gene is located on RNO15p12. A putative glycoprotease gene is located immediately adjacent to the Apex gene in a head-to-head orientation. An Apex pseudogene (Apex1) located in a LINE element is also shown to be present in rat genomic DNA. The present results should provide useful information for the further genetic study of APEX nuclease.

Acknowledgments. We thank Mr. M. Uji of Hokkaido University and Mr. T. Nakamura and Ms. K. Tabuchi of Okayama University for their technical assistance. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture and in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan (Repair, Recombination and Mutagenesis, 08280101). Footnotes: The nucleotide sequence data reported in this paper have been deposited in the GSDB, DDBJ, EMBL and NCBI nucleotide sequence data bases under the Accession Nos. AB023065 and AB023066.

References
9. Xanthoudakis S, Smeyne RJ, Wallace JD and Curran T: The redox/DNA repair protein, Ref-1, is essential for early embryonic develop-


Received July 19, 1999; accepted September 24, 1999.