Molecular cloning of retrovirus-like genes present in multiple copies in the Syrian hamster genome

Atsushi Suzuki¹, Hidero Kitasato², Masaya Kawakami¹ and Masao Ono¹*

¹Department of Molecular Biology, School of Medicine, and ²Department of Microbiology, School of Hygenic Science, Kitasato University, Sagamihara-shi, Kanagawa 228, Japan

Received 31 August 1982; Accepted 17 September 1982

ABSTRACT
Endogenous retrovirus-like sequences homologous to intracisternal type-A particle (IAP) genes, which are present in the inbred mouse (Mus musculus) genome, were cloned from a Syrian hamster gene library. A typical hamster IAP gene was 7 kb long and segments homologous to long terminal repeat (IAP) sequences present in Mus musculus IAP genes were located at both ends of the gene. Contrary to the pattern found in the Mus musculus IAP genes, the organization of the cloned hamster IAP genes was not markedly polymorphic and deletion was not observed among these cloned genes. A sequence about 0.8 kb long and located close to the 3' end of the hamster IAP gene was well conserved in both IAP gene families, although they showed less overall homology with one another. The reiteration frequency of the hamster IAP genes was calculated to be 950 copies per haploid genome. Since such IAP genes with the above properties were not found in the genome of the Chinese hamster, whose progenitors diverged from those of the Syrian hamster about 7.5 Myr ago, the integration of a huge number of Syrian hamster IAP genes must have occurred subsequent to such divergence.

INTRODUCTION
Intracisternal type-A particles (IAPs) are retrovirus-like particles found in most tumor and early embryonic cells of inbred mice (1-6) and of other mammals (7-13). As shown in other retrovirus particles, IAPs have particle-associated single-stranded poly(A)-containing RNAs (IAP RNAs) of 4 to 7 kb long (14-16). IAP RNA, however, has no apparent sequence homology with either type-B or type-C murine retroviral RNA (17) although a certain class of retroviruses endogenous to the Asian murine species, Mus caroli and Mus cervicicolor, has been reported to share partial sequence homology with IAP RNA (18,19).

DNA sequences complementary to mouse IAP RNA (IAP genes) are reiterated at about 1,000 copies per haploid genome of both normal and transformed cells in the inbred mouse (15,20,21) and localized dispersedly on many chromosomes (20). In contrast, the numbers of endogenous genes homologous to murine type-B and type-C retroviruses have been reported to be less than 1/20 of that of IAP genes present in inbred mouse genome. A typical IAP gene isolated...
from the Mus musculus gene libraries was 7 kb long (16,22,23) and at both ends of the gene had long terminal repeat (LTR) sequences 0.4 kb long (16,23) with U3-U5 structures (16,24) known as other retrovirus LTRs (25). The structures of cloned IAP genes were fairly polymorphic and a portion of the IAP gene was missing in some clones (15,16).

Unlike murine type-B and type-C endogenous retrovirus genes (26-29), similar numbers and organizations of these IAP genes were widely distributed among subspecies of Mus musculus (16,21) which are thought to have diverged from one another about one to two Myr ago (30). Furthermore, sequences homologous to Mus musculus IAP genes were detected in the DNAs prepared not only from Mus caroli and Mus cervicolor, but also from other rodent animals, and even from cats and monkeys (31). Among the rodent animals, repetitive sequences homologous to Mus musculus IAP genes were detected in Syrian hamster genome (31). On the other hand, such distinctive sequences were not found in the DNA prepared from rats (Rattus norvegicus) or Chinese hamsters (Cricetulus griseus) which are generally thought to be phylogenetically much closer to inbred mice and Syrian hamsters, respectively (32).

We were much interested in these IAP gene-like sequences present in the Syrian hamster genome because of their unusual distribution in the order Rodentia. In order to prove the structure, reiteration frequency, and origin of these genes, hamster IAP genes were cloned from a gene library. We found that highly repetitive and mutually homologous hamster IAP genes were 7 kb in length and contained putative LTR sequences, although the sequence homology between Mus musculus and Syrian hamster IAP gene families was limited.

**MATERIALS AND METHODS**

**Preparation of DNAs and Filter Hybridization**

High molecular weight DNAs were prepared from livers of Syrian hamsters and BALB/c mice, as described previously (15). Restriction endonucleases and T4 ligase were obtained from Takara Shuzo Co., and digesions were carried out according to the instructions of the supplier. DNA fragments were electrophoresed in an agarose gel (15) and transferred to a nitrocellulose membrane (Schleicher and Schüll) following the method of Southern (33). The sizes of the digested fragments (in kb) were determined from the Hind III digest of λc1857 DNA on the same gel. Recombinant DNA clones containing IAP genes from the Swiss/Webster gene library have been described previously (15,23). DNA fragments for probes were prepared (15) and nick-translated.
triphosphate (2,000 Ci/m mole, Amersham). Filter hybridization was performed in 50 mM Tris-HCl (pH 7.5), 1 M NaCl, 10 mM EDTA, 10 x Denhardt's solution, 0.1% SDS, 100 µg/ml of sonicated, denatured E. coli DNA, and 1-2 x 10^6 cpm/ml of 32P-labeled probe, at 65°C for 16 to 18 hrs. The filter was washed 4 times (10 min per wash) in 0.3 x SSC, 0.1% SDS at 65°C.

Construction and Screening of a Syrian Hamster Gene Library

The Syrian hamster gene library was prepared as described previously (15). To prepare the library, partial Eco RI digested DNA from Syrian hamster livers was size fractionated on a 10 to 40% sucrose gradient and the 13- to 17-kb region was used for in vitro packaging with the Charon 4A arms. Using 1.1 µg of hamster DNA and 3.0 µg of Charon 4A arms, in vitro packaging yielded 3.2 x 10^6 independent recombinants, which were amplified, as described in a conventional manner (15). Partial Eco RI digested gene library of BALB/c liver DNA was also prepared as described above.

The gene libraries were screened (35) and individual clones were plaque purified two times. Lambda clones containing Syrian hamster IAP genes were digested with Eco RI and subcloned into Eco RI-digested pBR 322, as described previously (15). Lambda and plasmid DNA were prepared using conventional techniques (15).

RESULTS

Cloning of Syrian Hamster IAP Genes

Gene libraries were screened with probes prepared from a Mus musculus IAP gene (Fig. 1). Using probe I (1.4 kb fragment, see Fig. 3), which was generated from the middle portion of the Mus musculus IAP gene (15,23), we could not find any significant difference between the number of spots per filter from BALB/c mouse and that from Syrian hamster, although this probe hybridized more intensely with BALB/c clones than with hamster ones. As suggested previously (31), this result strongly indicated the presence of Syrian hamster IAP genes which are highly repetitive but less homologous to Mus musculus ones.

As judged from the average number of three experiments in which 1.5 x 10^4 recombinant plaques were screened, the frequencies of positive spots were determined to be 0.37% on BALB/c and 0.50% on hamster libraries, respectively. Since more than 95% of the screened phages were determined to be recombinants, the number of IAP genes possessing a certain part of the probe I sequence was calculated to be 700 copies per BALB/c and 950 copies per
Fig. 1. Plaque hybridization of the gene libraries with a probe prepared from the cloned Mus musculus IAP gene. 1.5 x 10^4 recombinant phages from the gene libraries were grown on 9 cm plastic dishes and screened as described (35) using 32P-labeled probe I DNA prepared from pS81A (23) DNA (see Fig. 3). (a) BALB/c liver library, (b) Syrian hamster liver library, (c) Charon 4A.

Syrian hamster haploid genome. These values were obtained considering the haploid genome size of the mouse and the Syrian hamster as 3.0 x 10^6 kb (36) and the average length of the insert DNA of 9 BALB/c and 9 hamster IAP clones as 15 kb. The number of IAP genes in the BALB/c genome calculated by this procedure was in good agreement with those obtained by the method of reassociation kinetics (15,20) and filter hybridization (21).

Probes II (6.8 kb fragment), which contained most of the Mus musculus IAP gene and a 0.9 kb flanking sequence, hybridized almost as frequently (0.53%) with recombinant plaques prepared from the hamster library as probe I did. In contrast, probe II hybridized a little more frequently (0.51%) with those from the BALB/c library than probe I did, possibly due to the presence of the clones lacking a sequence corresponding to probe I (15,16).

Organization of Cloned Hamster IAP Genes

From the hamster gene library, recombinant phages which possessed sequences hybridizable with probe II were selected as clones of Syrian hamster IAP genes. DNAs prepared from plaque purified clones of the hamster IAP genes were digested with Eco RI and analyzed by the Southern blotting technique (Fig. 2). In eight clones analyzed, all had at least two Eco RI fragments hybridizable with probe II and one of them was almost identical in length in all 8 clones. The size of this common Eco RI fragment was determined to be 1.7 kb long. All clones but λH15 had an Eco RI fragment strongly hybridizable with probe II.
Fig. 2. Southern blotting analysis of the recombinant lambda DNAs containing Syrian hamster IAP genes. Plaques hybridizable with probe II were purified as the clones of Syrian hamster IAP genes. The recombinant DNAs prepared from these plaques were digested with Eco RI and analyzed by the Southern blotting technique. (a) Eco RI digested DNAs (0.75µg) electrophoresed on a 0.75% agarose gel and stained with ethidium bromide. Hind III digested λcycl2857 DNA was used for the estimation of molecular weight indicated as kilo-bases (kb) in this and following figures except in Fig. 4. (b) Southern blotting and hybridization analysis with probe II. ▶️, Common 1.7 kb Eco RI fragments; ●, Eco RI fragments of 5' side IAP genes; □, Eco RI fragments of 3' side IAP genes.

For more detailed analysis, Eco RI fragments hybridizable with probe II were subcloned from lambda phage Charon 4A into plasmid pBR 322 and common restriction enzyme sites of each clone were determined to give a preliminary organization map of the hamster IAP gene (Fig. 3). A typical IAP gene cloned from *Mus musculus* genome was 7 kb long with a 0.4 kb LTR sequence located at both ends of the gene; a Pst I site was always positioned almost in the middle of this LTR sequence (16, 23). In the hamster IAP genes, we found common Pst I sites located 7 kb apart from each other. Furthermore Bam HI, Hind III, and Sst I sites were localized on the sequences within a distance of 0.1 kb from these Pst I sites. Therefore, the clustered sites appeared to be located in the middle of the putative LTR sequences present in the hamster IAP gene.

In order to prove this possibility we first tried to make a more detailed restriction enzyme map on these four clustered restriction enzyme sites. Orders and distances among these restriction enzyme sites were completely the
Fig. 3. Structure of cloned Syrian hamster IAP genes. The relative positions of the restriction enzyme sites were determined by single and double digests of subcloned DNAs with Eco RI ( ), Bam HI ( ), Hind III ( ), Pst I (○) and Sst I ( ). Linkage of each Eco RI fragment was determined by either Bam HI, Hind III or Pst I digestion of the lambda DNA and hybridization with probe II. These cloned sequences were aligned according to the common restriction enzyme sites on each sequence. The orientation of the hamster IAP gene was defined from that of the Mus musculus IAP gene (16,23). The 5' end of the gene is positioned at the left. (a) A typical hamster IAP gene is shown with the putative LTR sequences which are assumed to be 0.4 kb in length according to those observed in Mus musculus genes. (b) A typical Mus musculus IAP gene (23). Probes used in this study are also indicated.

As we have done previously (23), cross-hybridization experiments were...
Fig. 4. Terminally redundant sequences present in hamster IAP genes analyzed on polyacrylamide gel. Subcloned DNAs (5μg) were digested with Sst I plus Bam HI, or Hind III, or Pst I and analyzed on a native 8% polyacrylamide gel with Hinf I digested pBR 322 DNA fragments (indicated as base-pairs) as standards (37). 1, Hinf I digested pBR 322 (3μg); 2, pH18A digested with Sst I plus Bam HI; 3, pH18B digested Sst I plus Bam HI; 4, pH10C digested with Sst I plus Bam HI; 5, pH18B digested with Sst I plus Pst I; 6, pH10C digested with Sst I plus Pst I; 7, pH10C digested with Sst I plus Hind III.

carried out for further confirmation of the presence of LTR sequences. For this purpose, probe III and IV DNAs were prepared from subcloned pH2B DNA digested with Eco RI and Hind III. Then, subcloned DNA presumably possessing 5' LTR sequences (5.0 kb pH10B, 5.8 kb pH18A) and 3' LTR sequences (3.0 kb pH2B, 2.1 kb pH10C, 3.9 kb pH18B) were digested with Eco RI and Hind III, and analyzed by the Southern blotting technique using $^{32}$P-labeled probe III and IV. As shown in Fig. 5, probe III hybridized with DNA fragments located upstream from the Hind III site in the putative LTR sequence, and probe IV hybridized with those positioned downstream from this Hind III site. These results again strongly supported the presumption concerning the presence of LTR sequences. We next investigated whether these putative LTR sequences of the hamster IAP genes were homologous to those of the Mus musculus IAP genes. For this analysis the 1.5 kb fragment (probe V) containing 5' LTR segment and the 0.9 kb flanking region were prepared from subclone pS19D of the Mus musculus IAP gene (23). This probe hybridized only with DNA fragments present upstream from the Hind III site in the putative LTR sequence although the intensities of these bands were much weaker than those formed by the
Fig. 5. Terminally redundant sequences present in hamster IAP genes analyzed by blot hybridization. Eco RI plus Hind III digested subcloned DNAs (0.5 μg) were resolved on a 1.2% agarose gel and analyzed by the Southern blotting method. (a) A Southern blot of the gel hybridized with probe III. (b) An identical blot hybridized with probe IV. (c) Another blot hybridized with probe V. 1, pH18A; 2, pH10B; 3, pH18B; 4, pH2B; 5, pH10C.

hybridization among the contaminated vector pBR 322 DNAs. This result indicated that the putative LTR sequence of the hamster IAP gene had some but weaker homology with the LTR segment of the *Mus musculus* IAP gene and that homologous region must have been located at the 5’ half of the LTR sequence, namely the U3 region.

The Pst I sites were observed in the middle of almost all putative LTR sequences not only of the hamster and *Mus musculus* (16,23), but also of *Mus caroli* and *Mus cervicolor* IAP genes (manuscript in preparation). These Pst I sites were evolutionarily conserved suggesting that the sequences around these sites might have a significant function for the IAP genes. Hind III sites located in the vicinity of the Pst I sites in the putative LTR sequence...
Nucleic Acids Research

were found in some clones of *Mus musculus*, and in most clones of *Mus caroli* (manuscript in preparation) IAP genes. On the other hand, in the IAP gene families no LTR sequence with either Bam HI or Sst I site in it has been reported other than the putative hamster LTR segment.

As shown in the *Mus musculus* IAP genes (15,16,22), sequences flanking the putative hamster LTR segments were heterogeneous, indicating that the hamster IAP genes must have been positioned at many sites in the genome. Contrary to the *Mus musculus* IAP genes, the structures of the hamster IAP genes, as determined by the locations of several restriction enzyme sites, resembled each other and no deletion was observed among these cloned hamster IAP genes.

**Organization of IAP Genes in Syrian Hamster Genome**

In order to examine the organization of the hamster IAP genes in the normal hamster genome, Syrian hamster liver DNA was digested with several restriction enzymes and analyzed by the Southern blotting method using three probes (I,II, and VI) prepared from the *Mus musculus* IAP gene. The results which had been predicted from the organization map of the cloned hamster IAP genes were obtained upon digestions of the hamster DNA either with Bam HI, Eco RI, Hind III, or Sst I (Fig. 6). Probe II hybridized with Bam HI digested hamster DNA to give 1.4, 1.8, and 4.0 kb fragments. The size of these three bands well coincided with those expected from the organization map. Among these bands, the 1.8 kb fragment with which probe II hybridized most intensely could be positioned at the 3' most end on the cloned gene. In contrast, probe I hybridized exclusively with the 1.4 kb fragment. On the cloned IAP genes prepared from *Mus musculus* (15,16), *Mus caroli*, *Mus cervicolor* (manuscript in preparation), and Syrian hamster gene libraries, we could find these 1.4 kb fragments interposed by two Bam HI sites and positioned at some distance downstream from the middle of the gene. These fragments were also identified in cellular DNAs prepared from these animals upon digestions with Bam HI.

Digestion of the hamster DNA with Hind III gave 1.7, 2.4, and 4.1 kb bands hybridizable with probe II. As expected, probe I hybridized with the 2.4 kb band and probe VI, which was positioned on the 0.8 kb fragment of the *Mus musculus* IAP gene just downstream from the probe I DNA, hybridized with the 1.7 kb segment. A 4.1 kb band hybridizable with all three probes must have been derived from the hamster IAP genes lacking a Hind III site located between the 2.4 and 1.7 kb Hind III fragments, although no such clone could be identified among our 8 clones. A 1.7 kb band which showed the most
Fig. 6. Identification of IAP genes in hamster genome. Syrian hamster liver DNA (2µg) was digested with restriction enzymes and analyzed by the Southern blotting method with probe prepared from the Mus musculus IAP gene. (a) Bam HI digest. (b) Hind III digest. (c) Sst I digest. (d) Eco RI digest. 1, Hybridized with probe II; 2, Hybridized with probe I; 3, Hybridized with probe VI. Experiment B, C, and D were preformed on the same agarose gel.

intense hybridization with probe II was again positioned at the 3' most end on the cloned gene map. Probe II did not hybridize significantly with the 1.25 or the 1.85 kb band which could be generated from the 5' half of the cloned hamster IAP gene by Hind III digestion, although longer exposure or less stringent conditions for filter hybridization and washings gave detectable bands in these regions.

Sst I digested hamster DNA possessed the predicted 1.4, 1.8, and 4.5 kb fragments hybridizable with probe II, and probe I hybridized exclusively with the 4.5 kb band. The Sst I site located closer to the 3' end and identified on 1.7 kb common Eco RI fragment was so polymorphic among the
cloned hamster genes that we could easily understand the reason why probe VI hybridized with both 1.4 and 1.8 kb fragments. In addition to this Sst I site, some polymorphic nucleotide sequences were identified on the cloned hamster IAP genes by the presence or absence of certain restriction enzyme sites. For example, the internal Pst I or Sst I site was missing in pH14B or pH15A, respectively, while an additional Pst I site was observed in pH18A.

Since probe VI DNA was positioned just downstream from probe I DNA on the *Mus musculus* IAP gene and the bands hybridizable with probe VI were always identified as fragments located downstream from those identified by probe I on the cloned hamster IAP gene map, the orientation shown in Fig. 3 must be correct if we define the orientation of the hamster IAP gene according to that of the *Mus musculus* one. In spite of the similar size of the *Mus musculus* and hamster IAP genes, the sequences of these two gene families lacked significant homology, as expected by both solution (31) and filter hybridizations. Most homologous regions could be mapped on a 0.8 kb region located between the right side common Eco RI site and the putative 3' LTR segment, in light of the fact that probe II hybridized most intensely with the fragments containing this region in Fig. 2, and with the corresponding fragments in Fig. 6. This region did not include the putative LTR sequence because probe V, which contained the *Mus musculus* LTR sequence, could only weakly hybridize with the putative hamster LTR sequence (Fig. 5). Less homologous sequences must have been located on the 5' half of the hamster IAP gene because probe II did not hybridize intensely either with these regions of the cloned hamster IAP genes (Fig. 2) or with the fragments predicted to be generated from this portion of the hamster IAP gene in the genome by the digestion of certain restriction enzymes. Except for the highly homologous region, the 3' half of the hamster IAP gene was homologous to some extent between these gene families. The differences in the varieties and locations of the restriction enzyme sites in the two cloned IAP gene families also suggested the presence of considerable sequence heterogeneity.

**DISCUSSION**

When certain genes seem to be considerably repetitive, we can easily count the gene number from the frequency of certain clones obtained from a well-characterized gene library, even if we have to use a less homologous probe for the detection. According to this procedure, we could estimate the number of Syrian hamster IAP genes to be 950 copies per haploid genome. This number was similar to that of *Mus musculus* IAP genes (15,20,21) but was at...
least 20 times larger than those of type-B or type-c endogenous retrovirus genes present in the murine genome. Using the cloned IAP gene segments prepared from different species of Mus, such as Mus caroli and Mus cervicolor, we have demonstrated the number of IAP genes to be at least 200 in Mus caroli, and 400 in Mus cervicolor per haploid genome (manuscript in preparation). These numbers are about 10 times larger than the values reported previously (18). If our figure is correct, then IAP genes were present as considerably repetitive genes in both rodent genera.

A typical IAP gene cloned from a Syrian hamster gene library was 7 kb in length, and has an endogenous retrovirus gene-like structure with putative LTR segments at both ends of the gene. Furthermore, 0.8 kb regions located close to the 3' end of these IAP genes were highly homologous to the Mus musculus IAP genes. This region corresponds to the c region in a typical retrovirus genome which has gag-pol-env-c structure and conservation of the sequence corresponding to this region has been reported in other retrovirus families (38,39), although the function of this region is still unknown.

Using various parts of a cloned BALB/c IAP gene as probes on a Southern blotting analysis of Syrian hamster DNA, a restriction enzyme map concerned with Hind III sites on the Syrian hamster IAP gene has been reported (31). It is significant that our organization map obtained using molecular cloning techniques agrees fundamentally with their gene map.

IAP-like particles have been reported to exist not only in Mus (1-6) and Syrian hamster cells (7,8), but also in many rodent (9-11), and bovine (13) cells. Furthermore, sequences hybridizable with Mus musculus IAP gene have been detected in monkey, cat, and even bat DNAs in addition to DNAs from rodent animals (31). However the number and structure of these sequences present in these animals remain to be determined.

The origin of the hamster IAP genes is still unclear. About 7.5 Myr ago the progenitors of Syrian hamsters are thought to have diverged from those of Chinese hamsters (32). Therefore, integrations of a huge number of Syrian hamster IAP genes must have occurred after the divergence of these two hamster genera since IAP genes, such as those found in the Syrian hamster genome, could not be detected in the Chinese hamster (31, unpublished observation).

Intracisternal retrovirus-like particles have been observed in tissue culture (such as BHK cells) and transformed cells derived from Syrian hamsters (7,8). Since both ³²P-labeled probe I and II hybridized with RNA prepared from BHK cells but no significant hybridization between these probes and normal Syrian hamster liver RNA was detected (unpublished obser-
vation), these particles present in BHK cells must have RNAs transcribed from the Syrian hamster IAP genes. The transcription of IAP genes thus seems to correlate with the expression of IAPs. Furthermore, since IAPs have always been observed in transformed or established cells, except those in early stages of development, the function of IAP genes might be involved in the proliferation of abnormal cells. Recently, we have obtained an interesting report suggesting that one type of IAP gene, which shared only 2% of the IAP gene family present in Mus musculus, was amplified by two to four fold in myeloma cells (21). In addition, it is known that in different categories of tumor cells, such as myeloma and neuroblastoma, different types of IAP RNAs are transcribed (14-16). Since the oncogenesis of cells is currently thought to be induced by the genetic activation of an oncogene of cellular origin, the expression and/or rearrangement of IAP genes might play an important role in such activation in some malignant cells.

ACKNOWLEDGMENTS

We would like to thank Dr. Masuo Obinata for the useful technical informations and Dr. Norman D. Cook for the critical reading of this manuscript. This work was supported in part by a research grant from the Japanese Ministry of Education, Science and Culture.

*To whom correspondence should be addressed

REFERENCES