Susceptibility of muridae cell lines to ecotropic murine leukemia virus and the cationic amino acid transporter 1 viral receptor sequences: implications for evolution of the viral receptor

Katsura Kakoki[1,3], Akio Shinohara[2], Mai Izumida[1], Yosuke Koizumi[1], Eri Honda[1], Goro Kato[2], Tsukasa Igawa[3], Hideki Sakai[3], Hideki Hayashi[1], Toshifumi Matsuyama[1], Tetsuo Morita[4], Chihiro Koshimoto[2], and Yoshinao Kubo[1,5]*

[1]Division of Cytokine Signaling, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8523, Japan

[2]Department of Bio-resources, Division of Biotechnology, Frontier Science Research Center, University of Miyazaki, Miyazaki 889-1692, Japan

[3]Department of Urology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8523, Japan

[4]Department of Plant and Animal Science, Faculty of Agriculture, University of Miyazaki, Miyazaki 889-2192, Japan

[5]Department of AIDS Research, Institute of Tropical Medicine, G-COE, Nagasaki University, Nagasaki 852-8523, Japan

Correpondence: Yoshinao Kubo, Division of Cytokine Signaling, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan

E-mail: yoshinao@nagasaki-u.ac.jp

Running title: Evolution of the ecotropic MLV receptor

Abstract

Ecotropic murine leukemia viruses (Eco-MLVs) infect mouse and rat, but not other mammalian cells, and gain access for infection through binding the cationic amino acid transporter 1 (CAT1). Glycosylation of the rat and hamster CAT1s inhibits Eco-MLV infection, and treatment of rat and hamster cells with a glycosylation inhibitor, tunicamycin, enhances Eco-MLV infection. Although the mouse CAT1 is also glycosylated, it does not inhibit Eco-MLV infection. Comparison of amino acid sequences between the rat and mouse CAT1s shows amino acid insertions in the rat protein near the Eco-MLV-binding motif. In addition to the insertion present in the rat CAT1, the hamster CAT1 has additional amino acid insertions. In contrast, tunicamycin treatment of mink and human cells does not elevate the infection, because their CAT1s do not have the Eco-MLV-binding motif. To define the evolutionary pathway of the Eco-MLV receptor, we analyzed CAT1 sequences and susceptibility to Eco-MLV infection of other several murinae animals, including the southern vole (*Microtus rossiaemeridionalis*), large Japanese field mouse (*Apodemus speciosus*), and Eurosian harvest mouse (*Micromys minutus*). Eco-MLV infection was enhanced by tunicamycin in these cells, and their CAT1 sequences have the insertions like the hamster CAT1. Phylogenetic analysis of mamalian CAT1s suggested that the ansestral CAT1 does not have the Eco-MLV binding motif, like the human CAT1, and the mouse CAT1 is thought to be generated by the amino acid deletions in the third extracellular loop of CAT1.

Introduction

A change in the cell surface receptor for a virus is one of a host defense mechanism against virus infection; for example, the C-terminally truncated-CCR5 variant is known to confer resistance to human immunodeficiency virus (HIV) infection [1]. We have analyzed the viral receptors for ecotropic murine leukemia viruses (Eco-MLVs) in mouse (*Mus musculus*), rat (*Ratus norvegicus*), and *Mus dunni* cells as a model of receptor evolution that confers resistance to a virus infection [2-4].

Eco-MLVs can infect mouse and rat cells, and recognize the multi-membrane spanning cationic amino acid transporter 1 (CAT1) as the receptor for infection [5]. Eco-MLV binds the YGE or HGE motif in the third extracellular loop of the CAT1 [6,7]; the CAT1 has two N-linked glycosylation sites near the Eco-MLV binding motif. Nucleotide sequences near the virus-binding motif are highly diversified between mouse, rat, hamster, mink, and human, suggesting that the region is under selective pressure.

Rat cells are much less susceptible to Eco-MLV infection than mouse cells, and hamster cells are completely resistant to infection. Treatment of rat and hamster cells with tunicamycin, an N-linked glycosylation inhibitor, enhaces susceptibility to Eco-MLV infection [8,9]. Furthermore, an amino acid substitution at the glycosylation site of the rat CAT1 increases susceptibility to Eco-MLV infection [2]. These results indicate that N-linked glycosylation of the rat and hamster CAT1 proteins inhibits Eco-MLV infection. Although the mouse CAT1 is also glycosylated at the same amino acid residues as the rat and hamster CAT1s, it does not affect Eco-MLV infection [2]; rat and hamster CAT1 proteins have three- and six-amino acid insertions near the viral binding domain of the protein compared to the mouse CAT1. We have previously reported that a deletion of the amino acid insertion in the rat CAT1 confers increased susceptibility, and abrogates the glycosylation-mediated inhibition of Eco-MLV infection, indicating that

the amino acid insertion in the rat CAT1 is the determinant for the glycosylation-dependent infection inhbition [3]. The longer insetion present in the hamster CAT1 compared to the rat protein may confer complete resistant of hamster cells to Eco-MLV infection. In addition, glycosylation of the *Mus dunni* CAT1 also inhibits Eco-MLV infection as a result of a one-amino acid insertion in the YGE virus-binding motif [4].

To confirm the pathway of evolution for mammalian CAT1s, we established immortalized cell lines from several murinae animals, and then determined their susceptibility to Eco-MLV infection and their CAT1 sequences. We showed that the CAT1 sequences of the southern vole, large Japanese field mouse, and Eurosian harvest mouse were shown to have amino acid insertions similar to the hamster CAT1. Phylogenetic analysis of mammalian CAT1 sequences revealed that the CAT1 ancestor is the human-type CAT1, and evolved to the mouse-type CAT1 by deletion rather than by insertion. This study reviews the evolutionary pathway of the ecotropic MLV receptor, CAT1, in relation to the viral infection.

Materials and Methods

Aminals

The southern vole (*Microtus rossiaemeridionalis*) was obtained from the closed colony maintained at the Frontier Science Research Center, University of Miyazaki. A wild large Japanese field mouse (*Apodemus speciosus*) and an Eurasian harvest mouse (*Micromys minutus*) were captured in Kiyotake, Miyazaki City, Miyazaki Prefecture, Japan, with the approval of the prefectural governor (No. 24940-2696). This animal study is approved by the Ethics Committee of Nagasaki University (No. 0812080723), and the Committee for the Ethics on Animal Experiments at the University of Miyazaki (No. 2008-505).

Cells

Mouse NIH3T3, rat F10, and human TELCeB6 [10] cells were cultured in Dulbecco's modified Eagle's medium (D-MEM) supplemented with 8% fetal bovine serum. Kidney grafts of *M. minutus, M. rossiaemeridionalis,* and muscle grafts of *A. speciosus* were isolated and were treated with trypsin to separate cells. The cells were cultured with D-MEM containing 20% FBS for more than one year. Cells from *M. rossiaemeridionalis,* and *A. speciosus* were passed by 1/6 dilution every 3 days. Cells from *M. minutus* were passed by 1/2 dilution every 6 days.

Expression plasmids

An expression plasmid of the ecotropic Friend MLV Env has been already described [11]. A VSV-G expression plasmid was obtained from Dr. L. Chang through the AIDS Research and Reference Reagent Program, NIAID, NIH, USA. [12].

Transduction assay

TELCeB6 cells [10] were transfected with an expression plasmid containing the Friend MLV envelope protein using the Fugene transfection reagent (Promega). Because the expression plasmid also encodes the neomycin resistance gene, the transfected cells were selected by genetisin (Invitrogen). Culture supernatnats of the genetisin-resistant cell pool were inoculated into target cells in the presence of polybrene (4 μg/ml) (Sigma-Aldrich). To construct a VSV-pseudotyped MLV vector, TELCeB6 cells were transiently transfected with a VSV-G expression plasmid and their culture supernatants were collected 2 days after transfection. The culture supernatants were then inoculated into target cells. To estimate transduction titer, the inoculated cells were stained with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (Wako) 2 days after the inoculation and then blue cells were counted.

Isolation of CAT1 sequences

Total RNA samples were prepared from cells using Isogen (Invitrogen). cDNAs were synthesized by reverse transcriptase (TaKaRa) and a fragment containing the third extracellular loop of CAT1 and its upstream region was amplified by PCR (TaKaRa) using the cDNA as template. Nucleotide sequences of the PCR primers are AAC CTG ATT CTC TCC TAC ATC and GTG GTG GCG ATG CAG TCA AAG. The PCR products were cloned into pTarget vector (Promega) and nucleotide sequences of the insert DNAs were determined (Applied Biosystems). The primers were synthesized by Genenet Co., LTD. Accession numbers of CAT1s from A. speciosus, *M. rossiaemeridionalis,* and *M. minutus* are AB839945, AB839946, and AB839947, respectively.

Alignments and phylogenetic analysis

Rodent CAT1 gene sequences were compared with other mammalian CAT1 gene sequences obtained from the DNA database: laboratory rat (*Rattus norbegicus*; AB066224) [2], laboratory mouse *(Mus musculus*; M26687) [5], human (*Homo sapiens*; X59155) [7], pig (*Sus scrofa*; AY371320) [13], domestic dog (*Canis lupus familiaris*; XM_854224) [14], American mink (*Neovison vison*; U49796) [15], domestic cat (*Felis catus*; XM_003980275) [16], chimpanzee (*Pan troglodytes*; XM_001139004), horse (*Equus cabllus*; XM_001492839) [17], cattle (*Bos taurus*; NM_001135792) [18], giant panda (*Ailuropoda melanoleuca*; XM_002914759) [19], African elephant (*Loxodonta africana*; XM_003414018), bonobo (*Pan paniscus*; XM_003818324) [20], Syrian hamster (*Mesocricetus auratus*; U26454), Chinese hamster (*Cricetulus griseus*; U49797), orangutan (*Pongo abelii*; XM_002824135) [21], small-eared galago (*Otolemur garnettii*; XM_003797609), Northern white-cheeked gibbon (*Nomascus leucogenys*; XM_003270277), rabbit (*Oryctolagus cuniculuc*; XM_002721425), lesser Egyptian jerboa (*Jaculus jaculus*; XM_004659984), naked mole rat (*heterocephalus glaber*; XM_004854792), and degu (*Octodon degus*; XM_004631056).

Obtained three rodent CAT1 gene sequences were compared with other mammalian CAT1 gene sequences obtained from the DNA database. All sequences were onece tranlated into amino acids and then aligned using MUSCLE [22] implemented in MEGA ver 5.1 [23]. These aligned amino acids sequeces were reversely translated into nucleotide sequences, and use them for the analyses.

For phylogenetic analyses, we employed the CAT2 sequences of human (D29990) [24] and mouse (L03290) [25] for the out-group. Therefore, we re-aligned all of the sequences with theese CAT2 sequences following the same procedure described above. A phylogenetic tree was constructed by

Bayesian method. The dataset was divided into three partitions with codon position (1st, 2nd, and 3rd), and optimum substitution models for each partition were selected by Kakusan 4 [26] based on the Bayesian information criterion; general time reversible (GTR) [27] with gamma distribution (+G), GTR+G, and HKY85 [28] +G models were selected for the 1st, 2nd, and 3rd position, respectively. The Bayesian analysis was conducted using MrBayes v3.2.1 [29] with 3 million generations of two independent runs of four Markov chains. We sampled one tree every 100 generations and calculated a consensus topology with discarding the first 25% of trees. Final average standard deviation of split frequencies of the Bayesian analysis was 0.017942, and all average effective sample sizes were more than 200.

Results

Susceptibility of rodent cells to Eco-MLV infection

Immortalized cell lines were established from the inbred southern vole (*Microtus rossiaemeridionalis*), wild large Japanese field mouse (*Apodemus speciosus*), and wild Eurosian harvest mouse (*Micromys minutus*) to assess their susceptibility to Eco-MLV infection. Cells were inoculated with the Friend MLV Env protein-carrying MLV vector, and transduction titers were measured. Infected cells were detected in the *A. speciosus* and *M. rossiaemeridionalis* cells; we observed the transduction titers for these cells to be 1/100 to 1/1000 times lower than those of mouse NIH3T3 cells (*M. musculus*), similar to that found on rat F10 cells (*R. norvegicus*) (Fig. 1A). In contrast, infected cells were not detected from *M. minutus*, showing that these cells are resistant to Eco-MLV infection.

To determine whether N-linked glycosylation inhibits the Eco-MLV vector infection in these rodent cells, the cells were pretreated with tunicamycin (100 μg/ml) for 24 h, and then were inoculated with the Eco-MLV vector. In the *A. speciosus, M. minutus,* and *M. rossiaemeridionalis* cells, tunicamycin treatment enhanced Eco-MLV vector infection, as in the rat F10 cells (Fig. 1B), showing that glycosylation inhibits the Eco-MLV infection in these cells. In contrast, we have already reported that tunicamycin treatment of NIH3T3 cells does not affect the Eco-MLV vector infection [2,3]. However, infection by VSV-G-pseudotyped MLV vector was not affected by the tunicamycin treatment of these cells (data not shown), suggesting that the glycosilation-mediated inhibition of MLV vector infection is dependent on the viral Env protein. These data indicate that glycosylation of CAT1 proteins of *A. speciosus, M. minutus,* and *M. rossiaemeridionalis* inhibits Eco-MLV infection, suggesting that these CAT1 sequences have the amino acid insertion similar to the rat and hamster CAT1.

Nucleotide sequences of the CAT1 genes of rodent cells

DNA fragments containing the third extracellular loop and its upstream region of the CAT1 gene were isolated by RT-PCR from *M. rossiaemeridionalis, A. speciosus, and M. minutus* cells, and their nucleotide sequences were determined. Amino acid sequences of the third extracellular loop and its upstream region from the various mammals are shown in Figs 2A and B, respectively. The CAT1 sequences of *M. rossiaemeridionalis, A. speciosus,* and *M. minutus* contain the amino acid insertions and the YGE or HGE Eco-MLV-binding motif in the third extracellular loop, as found in the hamster CAT1. Interestingly, the CAT1 sequences from rabbit (*Oryctolagua cuniculus*) and lesser Egyptian jerboa (*Jaculuc Jaculus*) have divergent types of deletions in the third extracellular loop.

When nucleotide sequences were compared between rodents including *M. musculus, R. norvegicus, M. minumus, A. speciosus, M. rossiaemeridionalis*, Chinese hamster (*C. griseus*), Syrian hamster (*M. auratus*), naked mole rat (*H. glaber*), degu (O. degus), rabbit (*O. cuniculus*), and lesser Egyptian jerboa (*J. jaculus*), the third extracellular loops (Fig. 3A) (76.4 +/- 6.8 %) were less homologous than their upstream regions (Fig. 3B) (89.1 +/- 4.8 %) ($P=9.49 \times 10^{-23}$). Furthermore, the changes in the first and second nucleotides of codons in the CAT1 third extracellular loops (ECL3s) of rodents were more dominant than those in the upstream regions (Table 1), and nonsynominous mutations in the ECL3s were more abundant than those in the upstream regions, compared to the mouse CAT1 ($P=8.12 \times 10^{-11}$). In contrast, the amino acid sequences of the extracellular loops of the CAT1s between cat, dog, pig, horse, cattle, giant panda, elephant, small-eared galago, northern white-cheeked gibbon, Sumatran orangutan, bonobo, chimpanzee, and human are more similar than those between the rodents (Fig. 3C) (86.8 +/- 7.2 %) ($P=2.93 \times 10^{-14}$). These results suggest that the third extracellular region is under stronger selective pressure in rodents than in higher mammals.

Additionally, immortalized cells were established from inbred steppe lemming (*Lagurus lagurus*) and Mongolian gerbil (*Meriones unguiculatus*). CAT1 sequences could not be amplified from these cells by RT-PCR (data not shown). Consistently, when these cells were inoculated with the Eco-MLV vector, LacZ-expressing cells were not detected even in the presence of tunicamycin, although VSV-G-pseudotyped vector could transduce the cells. The. *L. lagurus* and *M. unguiculatus* cells might be resistant due to the lack of susceptible CAT1 expression.

Phylogenetic analysis of the ecotropic receptors

The phylogenetic tree of the CAT1s from the various mammals was constructed and is shown in Fig. 4. Rodents including mouse, rat, hamster, *A. speciosus*, *M. mimutus*, and *M. rossiaemeridionalis* and belong to one group. Human and primates belong to another group. The phylogenetic tree is consistent with the standard classification of mammals. By the phylogenetic tree, the Eco-MLV binding motif could be obtained before mouse, rat, and hamster are apearred (open arrow head in Fig. 4), and the amino acid deletions in the third extracellular loops of mouse, rat, rabbit, and lesser Egyptian jerboa CAT1s could occur independently (closed arrow heads in Fig. 4). This phylogenetic analysis also suggests that the ansestor CAT1 contain the amino acid insertion as present in the human CAT1s, and thereafter the mouse- and rat-type CAT1s could be generated from the hamster-type CAT1 by amino acid deletion (Fig. 5).

Discussion

This study provides the evolution pathway of the mammalian CAT1s as follows (Fig. 5): The ansestor of the Eco-MLV receptor in mammals does not have the Eco-MLV binding motif, like the human CAT1. Since Eco-MLVs cannot infect the human CAT1, Eco-MLVs should not circulate at that time. The ansestral human-type CAT1 might be converted to the hamster-type CAT1, aquiring the virus-binding motif (open arrow head in Fig. 4). Since the hamster-type CAT1 is much less susceptible to Eco-MLV infection than the mouse CAT1, Eco-MLV should not spread at that time. The mouse type CAT1s might be generated from the hamster-type CAT1 through amino acid deletion (closed arrow in Fig. 4). Since the mouse CAT1 is fully susceptible to Eco-MLV infection, Eco-MLV could circulate after the deletion occurred. Consistent with this speculation, it has been reported that many wild mice contain endogenous polytropic and xenotropic MLVs in their genomes, but wild mice carrying the endogenous ecotropic MLV are less numerous [30], suggesting that the ecotropic MLV appeared later than the polytropic and xenotropic viruses. Therefore, it is unlikely that Eco-MLV was derived from a virus which is related to Eco-MLV and efficiently interacts with the hamster-type CAT1 in the ancestor species of the *Mus* gunus. Since the third extracellular loops of *Mus* subgenus aminals, including M. *dunni*, *M. spicilegus*, and *M. minutoides*, have amino acid deletions as seen in the *M. musculus* CAT1 [31,32], the deletion is thought to have occured before the *Mus* subgenus appeared.

The homology of the CAT1 third extracellular loop among the rodents is relatively lower than among the higher mammals, suggesting that the regions are under selective presure in rodents. Since phylogenetic analysis of the mammalian CAT1s provided a possibility that Eco-MLV appeared after the *Mus* subgenus was generated, the selective pressure might not be the Eco-MLV infection itself. The CAT1 containing the amino acid deletion might be advantageous in the *Mus* subgenus and, eventually the

deletion might permit the appearance of EcoMLVs. The observation that mouse, rat, rabbit, and lesser Egiptian jerboa CAT1s indipendently have amino acid deletions in the third extracellular loop suggests that the third extracellular loop of CAT1 is disadvantegious for rodens.

Since the one amino acid insertion in the *Mus dunni* CAT1 is not found in the CAT1s of other *Mus* animals, it is thought that the one-amino acid insetion occurred to inhibit infection by Eco-MLV or other CAT1-recognizing viruses after the *Mus* subgenus appeared. CAT1-recognizing virus(s) might be widely spreaded among *M. dunni* population. To resolve this issue, futher study is required.

Acknowledgements

The VSV-G expression plasmid was kindly obtained from Dr. L. Chang through the AIDS Research and Reference Reagent Program, NIAID, NIH, USA. TELCeB6 cells were kindly provided by Dr. F. Cossett. We thank Ms. F. Tsujita, Ms. Y. Kobayashi, and Ms. M. Haraguchi for assistance. This study was partially supported by the Japan Society for the Promotion of Science.

References

1. Y. Huang, W.A. Paxton, S.M. Wolinsky, A.U. Neumann, L. Zhang, T. He, S. Kang, D. Ceradini, Z. Jin, K. Yazdanbakhsh, K. Kunstman, D. Erickson, E. Dragon, N.R. Landau, J. Phair, D.D. Ho, R.A. Koup, Nature Med. 2, 1240-1243 (1996)
2. Y. Kubo, T. Ono, M. Ogura, A. Ishimoto, H. Amanuma, Virology 303, 338-344 (2002)
3. Y. Kubo, A. Ishimoto, T. Ono, H. Yoshii, C. Tominaga, C. Mitani, H. Amanuma, N. Yamamoto, Virology 330, 82-91 (2004)
4. H. Yoshii, H. Kamiyama, H. Amanuma, K. Oishi, N. Yamamoto, Y. Kubo, J.Gen.Virol. 89, 297-305 (2008)
5. L.M. Albritton, L. Tseng, D. Scadden, J.M. Cunningham, Cell 57, 659-666 (1989)
6. L.M. Albritton, J.W. Kim, L. Tseng, J.M. Cunningham, J.Virol. 67, 2091-2096 (1993)
7. T. Yoshimoto, E. Yoshimoto, D. Meruelo, J.Virol. 67, 1310-1314 (1993)
8. M.V. Eiden, K. Farrell, C.A. Wilson, J.Virol. 67, 4056-4061 (1994)
9. N. Tavoloni, A. Rudenholz, Virology 229, 49-56 (1997)
10. F.L. Cosset, Y. Takeuchi, J.L. Battini, R.A. Weiss, M.K. Collins, J.Virol. 69, 7430-7436 (1995)
11. Y. Kubo, A. Ishimoto, H. Amanuma, J.Virol. 77, 7510-7516 (2003)
12. L.J. Chang, V. Urlacher, T. Iwakura, Y. Cui, J. Zucali, Gene Ther. 6, 715-728 (1999)
13. Z. Cui, S. Zharikov, S.L. Via, S.I. Anderson, A.S. Law, A.L. Archibald, E.R. Block, Genomics 85, 352-359 (2005)
14. K. Lindblad-Toh, C.M. Wade, T.S. Mikkelsen, E.K. Karlsson, D.B. Jaffe, M. Kamal, M. Clamp, J.L. Chang et al., Nature 438, 803-819 (2005)
15. H. Wang, E. Klamo, S.E. Kuhmann, S.L. Kozak, M.P. Kavanaugh, D. Kabat, J.Virol. 70, 6884-6891

(1996)

16. J.U. Pontius, J.C. Mullikin, D.R. Smith, Agencourt Sequencing team et al., Genome Res 17, 1675-1689 (2007)

17. C.M. Wade, E. Giulotto, S. Sigurdsson, M. Zoli, S. Gnerre, F. Imsland, T.L. Lear, D.L. Adelson, E. Bailey et al., Science 326: 865-867 (2009)

18. A.V. Zimin, A.L. Delcher, L. Florea, D.R. Kelley, M.C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C.P. van Tassell, T.S. Sonstegard, G. Marcais, M. Roberts, P. Subramanian, J.A. Yorke, S.L. Salzberg, Genome Biol. 10, R42 (2009)

19. R. Li, W. Fan, G. Tian, H. Zhu, L. He, J. Cai, Q. Huang, Q. Cai, B. Li, Y. Bai, Z. Zhang, Y. Zhang, W. Wang et al., Nature 463, 311-317 (2010)

20. K. Prufer, K. Munch, I. Hellmann, K. Akagi, J.R. Miller, B. Walenz, S. Koren, G. Sutton, C. Kodira, R. Winer et al., Nature 486, 527-531 (2012)

21. D.P. Locke, L.W. Hillier, W.C. Warren, K.C. Worley, L.V. Narareth, D.M. Muzny, S.P. Yang, Z. Wang, A.T. Chinwalla et al., Nature 469, 529-533 (2011)

22. R.C. Edgar, Nucleic Acid Res. 32, 1792-1797 (2004)

23. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, Mol.Biol.Evol. 28, 2731-2739 (2011)

24. R. Hoshide, Y. Ikeda, S. Karashima, T. Matsuura, S. Komaki, T. Kishino, N. Niikawa, F. Endo, I. Matsuda, Genomics 38, 174-178 (1996)

25. E.I. Closs, L.M. Albritton, J.W. Kim, J.M. Cunningham, J.Biol.Chem. 268, 7538-7544 (1993)

26. A.S. Tanabe, Molecular Ecology Resources 11, 914–921 (2011)

27. S. Tavaré, Some probabilistic and statistical problems in the analysis of DNA sequences. Pp. 57–86

in Some mathematical questions in biology - DNA sequence analysis (R. M. Miura, ed.). American Mathematical Society, Providence, RI. (1986)

28. M. Hasegawa, H. Kishino, T. Yano, J.Mol. Evol. 22, 160–174 (1985)

29. F. Ronquist et al., Systematic Biology 61, 539–542 (2012)

30. C.A. Kozak, R.R. O'Neill, J.Virol. 61, 3082-3088 (1987)

31. Y.T. Jung, C.A. Kozak, J.Virol. 77, 5065-5072 (2003)

32. Y. Yan, C.A. Kozak, J.Virol. 82, 6120-6129 (2008)

Figure legends

Fig. 1 Susceptibility of rodent cells to Eco-MLV infection. Rodent cells were inoculated with the Eco-MLV (A and B). Cells were treated or untreated with tunicamycin for 24 h and then were inoculated with the viral vector. Transduction titers are indicated. These experiments were repeated three times and representative results are shown.

Fig. 2 Amino acid sequences of CAT1s. The amino acid sequences of the third extracellular loops (A) and the upstream regions (B) of CAT1s are indicated. Bars and dots indicate identical amino acids and deletions, respectively. The Eco-MLV binding motif is underlined.

Fig. 3 Comparison of nucleotide sequences of CAT1s. Nucleotide sequences of the third extracellular loops (A) and the upstream regions (B) of CAT1s from indicated animals. Nucleotide sequences of the third extracellular loops (A) and their upstream regions (B) of rodent CAT1s are compared. Nucleotide sequences of the third extracellular loops of CAT1s from the higher animals are compared (C).

Fig. 4 Phylogenetic tree of the mammalian CAT1s. Closed arrow heads indicate the time points when amino acid deletions in the third extracellular loops of CAT1s occurred. An open arrow head shows the time point when the Eco-MLV binding motif obtained.

Fig. 5 Evolutionary pathway of CAT1 predicted by this study. Block letters show glycosylation sites. Dots show amino acid deletions. The Eco-MLV binding motifs are underlined.

Fig. 1A

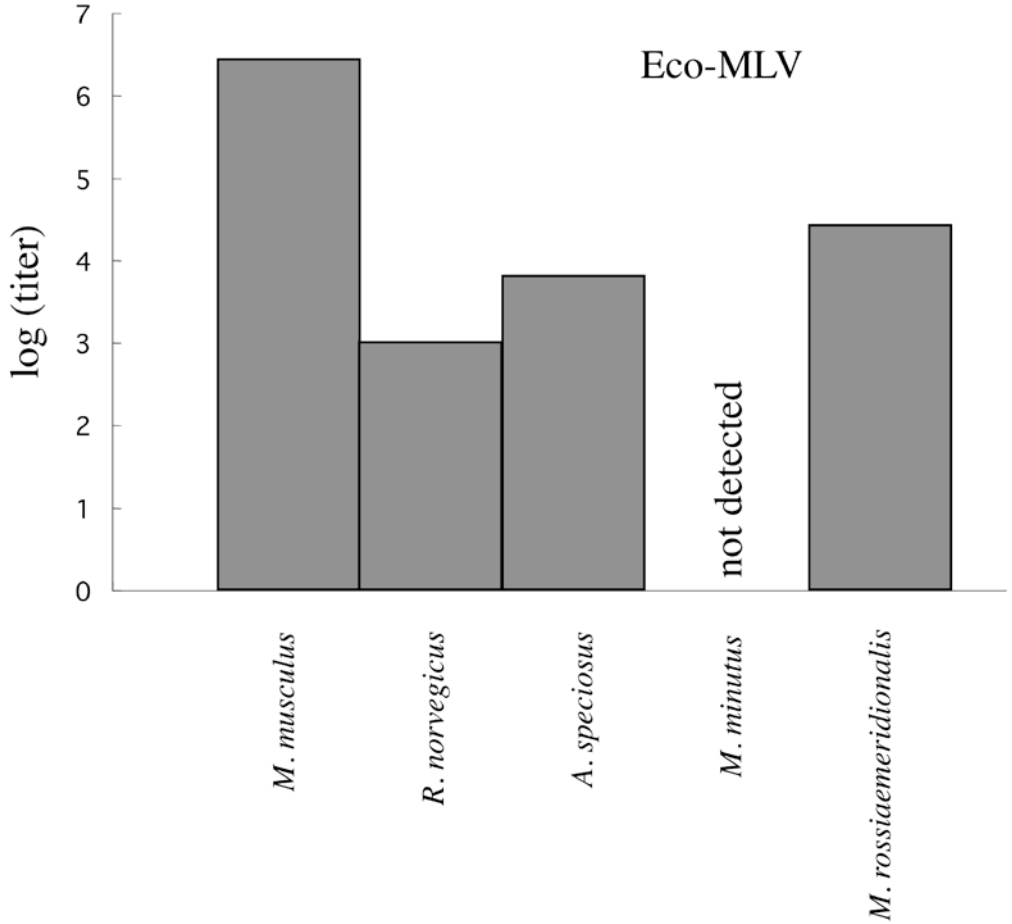
Eco-MLV
log (titer)
0
1
2
3
4
5
6
7
not detected
M. musculus
R. norvegicus
A. speciosus
M. minutus
M. rossiaemeridionalis

Fig. 1B

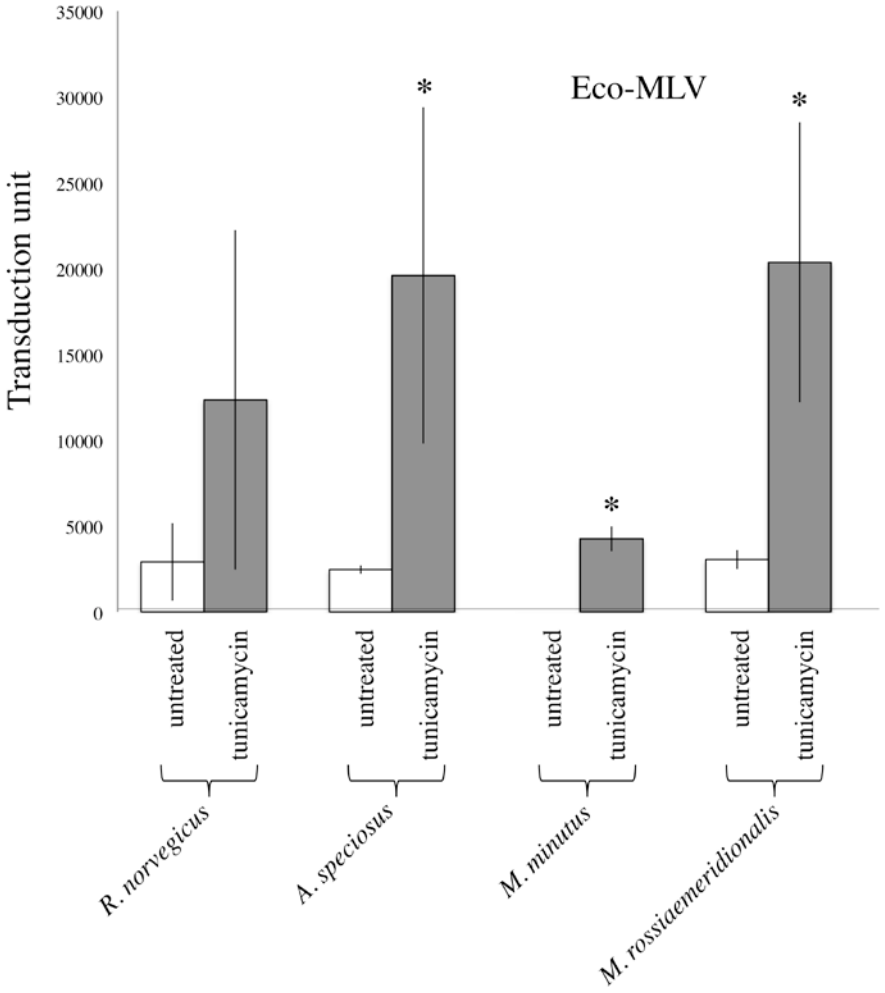
Eco-MLV
Transduction unit
35000
30000
25000
20000
15000
10000
5000
0
*
*
*
untreated
tunicamycin
untreated
tunicamycin
untreated
tunicamycin
untreated
tunicamycin
R. norvegicus
A. speciosus
M. minutus
M. rossiaemeridionalis

Fig. 2A

A

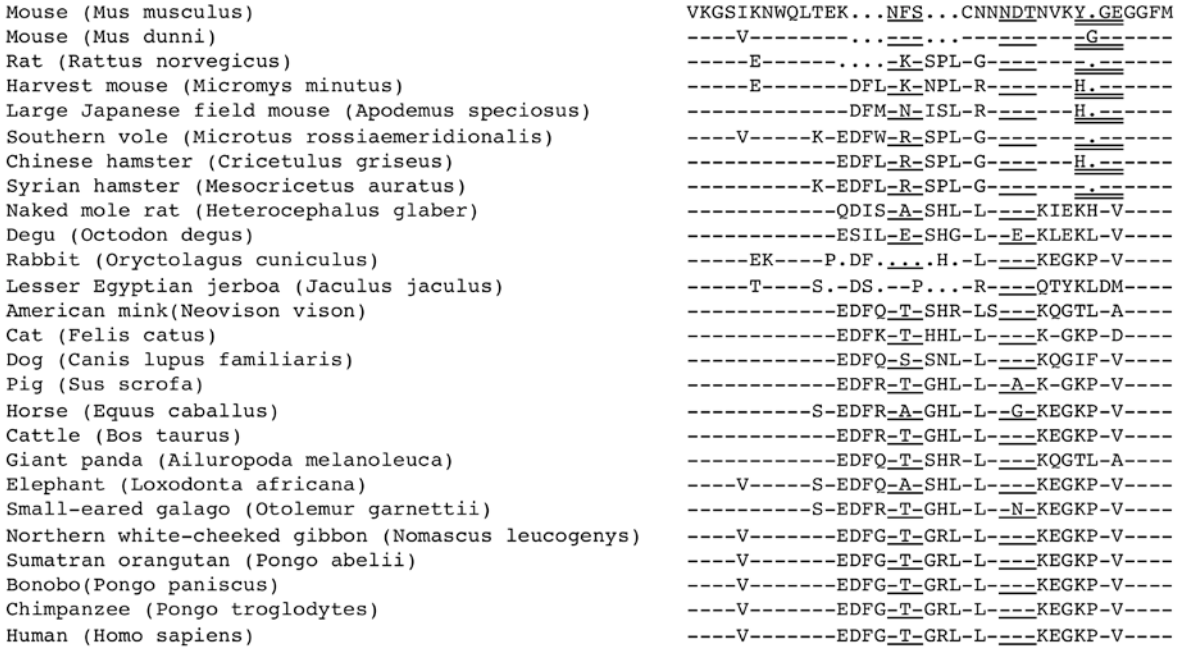
Mouse (Mus musculus) VKGSIKNWQLTEK...NFS...CNNNDTNVKY.GEGGFM
Mouse (Mus dunni) ----V--------...---...----------G------
Rat (Rattus norvegicus) -----E------....-K-SPL-G-------=.------
Harvest mouse (Micromys minutus) -----E-------DFL-K-NPL-R-------H.------
Large Japanese field mouse (Apodemus speciosus) ------------DFM-N-ISL-R-------H.------
Southern vole (Microtus rossiaemeridionalis) ----V-----K-EDFW-R-SPL-G-------=.------
Chinese hamster (Cricetulus griseus) ------------EDFL-R-SPL-G-------H.------
Syrian hamster (Mesocricetus auratus) ----------K-EDFL-R-SPL-G-------=.------
Naked mole rat (Heterocephalus glaber) ------------QDIS-A-SHL-L---KIEKH-V----
Degu (Octodon degus) ------------ESIL-E-SHG-L--E-KLEKL-V----
Rabbit (Oryctolagus cuniculus) -----EK----P.DF.....H.-L---KEGKP-V----
Lesser Egyptian jerboa (Jaculus jaculus) -----T----S.-DS.--P...-R---QTYKLDM----
American mink(Neovison vison) ------------EDFQ-T-SHR-LS---KQGTL-A----
Cat (Felis catus) ------------EDFK-T-HHL-L---K-GKP-D----
Dog (Canis lupus familiaris) ------------EDFQ-S-SNL-L---KQGIF-V----
Pig (Sus scrofa) ------------EDFR-T-GHL-L--A-K-GKP-V----
Horse (Equus caballus) ----------S-EDFR-A-GHL-L--G-KEGKP-V----
Cattle (Bos taurus) ------------EDFR-T-GHL-L---KEGKP-V----
Giant panda (Ailuropoda melanoleuca) ------------EDFQ-T-SHR-L---KQGTL-A----
Elephant (Loxodonta africana) ----V-----S-EDFQ-A-SHL-L---KEGKP-V----
Small-eared galago (Otolemur garnettii) ----------S-EDFR-T-GHL-L--N-KEGKP-V----
Northern white-cheeked gibbon (Nomascus leucogenys) ----V-------EDFG-T-GRL-L---KEGKP-V----
Sumatran orangutan (Pongo abelii) ----V-------EDFG-T-GRL-L---KEGKP-V----
Bonobo(Pongo paniscus) ----V-------EDFG-T-GRL-L---KEGKP-V----
Chimpanzee (Pongo troglodytes) ----V-------EDFG-T-GRL-L---KEGKP-V----
Human (Homo sapiens) ----V-------EDFG-T-GRL-L---KEGKP-V----

Fig. 2B

B

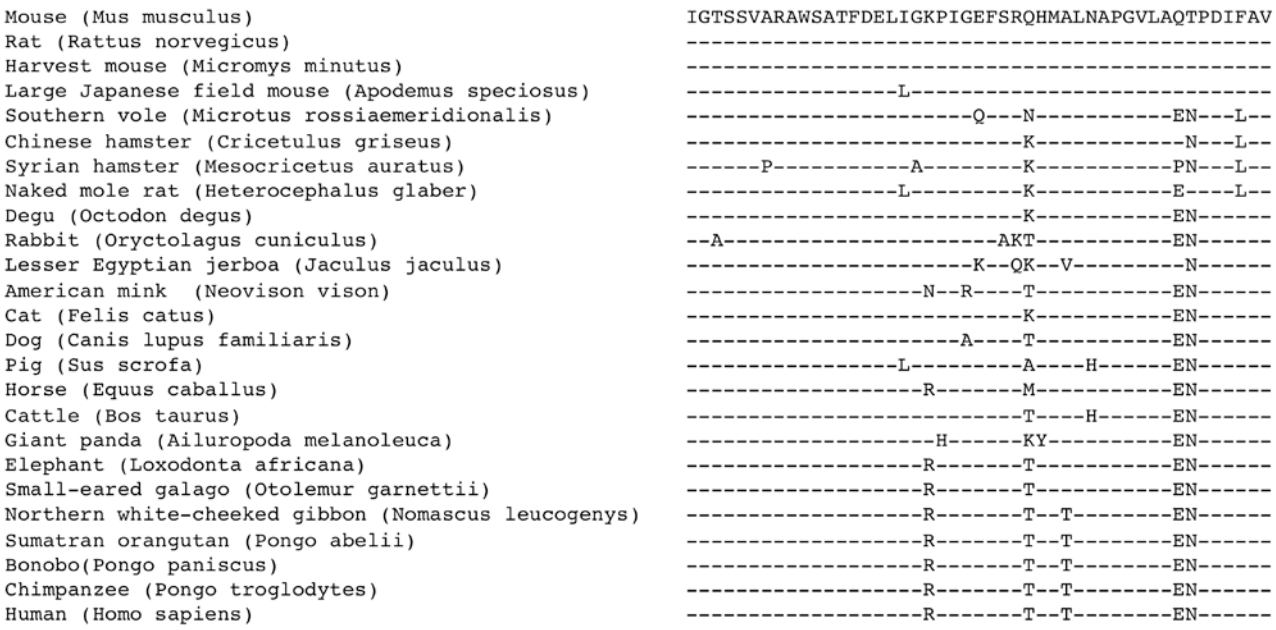
Mouse (Mus musculus) IGTSSVARAWSATFDELIGKPIGEFSRQHMALNAPGVLAQTPDIFAV
Rat (Rattus norvegicus) ----------
Harvest mouse (Micromys minutus) ----------
Large Japanese field mouse (Apodemus speciosus) ----------L----------
Southern vole (Microtus rossiaemeridionalis) ----------Q---N----------EN---L--
Chinese hamster (Cricetulus griseus) ----------K----------N---L--
Syrian hamster (Mesocricetus auratus) ------P----------A--------K----------PN---L--
Naked mole rat (Heterocephalus glaber) ----------L---------K----------E----L--
Degu (Octodon degus) ----------K----------EN------
Rabbit (Oryctolagus cuniculus) --A----------AKT----------EN------
Lesser Egyptian jerboa (Jaculus jaculus) ----------K--QK--V---------N------
American mink (Neovison vison) ----------N--R----T----------EN------
Cat (Felis catus) ----------K----------EN------
Dog (Canis lupus familiaris) ----------A----T----------EN------
Pig (Sus scrofa) ----------L---------A----H------EN------
Horse (Equus caballus) ----------R-------M----------EN------
Cattle (Bos taurus) ----------T----H------EN------
Giant panda (Ailuropoda melanoleuca) ----------H------KY----------EN------
Elephant (Loxodonta africana) ----------R-------T----------EN------
Small-eared galago (Otolemur garnettii) ----------R-------T----------EN------
Northern white-cheeked gibbon (Nomascus leucogenys) ----------R-------T--T--------EN------
Sumatran orangutan (Pongo abelii) ----------R-------T--T--------EN------
Bonobo(Pongo paniscus) ----------R-------T--T--------EN------
Chimpanzee (Pongo troglodytes) ----------R-------T--T--------EN------
Human (Homo sapiens) ----------R-------T--T--------EN------

Fig. 2B continued

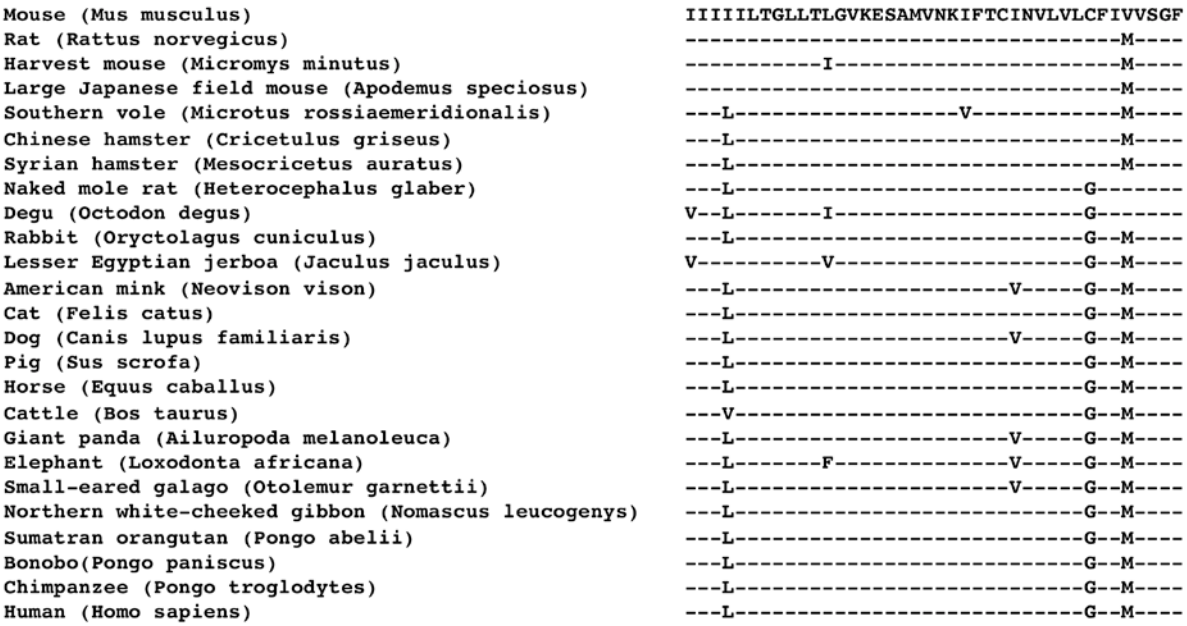

Fig. 3A

	M. musculus	*R.norvegicus*	*M. minutus*	*A. speciosus*	*M. rossiaemeridionalis*	*C. griseus*	*M. auratus*	*H. glaber*	*O. degus*	*O. cuniculus*	*J. jaculus*
M. musculus	100	74	76	76	71	74	74	70	72	64	66
R.norvegicus		100	82	74	82	82	82	75	74	71	69
M. minutus			100	87	76	87	82	80	79	72	69
A. speciosus				100	76	84	79	80	80	70	72
M. rossiaemeridionalis					100	89	94	81	79	68	68
C. griseus						100	94	82	80	72	72
M. auratus							100	78	77	68	68
H. glaber								100	89	72	73
O. degus									100	73	72
O. cuniculus										100	74
J. jaculus											100

Fig. 3B

	M. musculus	*R.norvegicus*	*M. minutus*	*A. speciosus*	*M. rossiaemeridionalis*	*C. griseus*	*M. auratus*	*H. glaber*	*O. degus*	*O. cuniculus*	*J. jaculus*
M. musculus	100	98	96	98	88	92	90	88	86	85	88
R.norvegicus		100	98	100	90	94	92	85	84	83	87
M. minutus			100	98	90	94	92	86	86	84	89
A. speciosus				100	90	94	92	84	83	84	89
M. rossiaemeridionalis					100	96	96	87	87	83	85
C. griseus						100	98	85	88	85	89
M. auratus							100	84	86	83	86
H. glaber								100	92	87	85
O. degus									100	87	89
O. cuniculus										100	86
J. jaculus											100

Fig. 3C

	F. catus	*C. lupus*	*S. scrofa*	*E. caballus*	*B. taurus*	*A. melanoleuca*	*L. africana*	*O. garnettii*	*N. leucogenys*	*P. abelii*	*P. paniscus*	*P. troglodytes*	*H. sapiens*
F. catus	100	79	87	79	87	82	82	82	82	82	82	82	82
C. lupus		100	79	77	82	85	82	77	79	79	79	79	79
S. scrofa			100	90	95	79	82	92	90	90	90	90	90
E. caballus				100	92	77	90	95	85	85	85	85	85
B. taurus					100	85	87	95	92	92	92	92	92
A. melanoleuca						100	79	79	77	77	77	77	77
L. africana							100	87	87	87	87	87	87
O. garnettii								100	90	90	90	90	90
N. leucogenys									100	100	100	100	100
P. abelii										100	100	100	100
P. paniscus											100	100	100
P. troglodytes												100	100
H. sapiens													100

Fig. 4

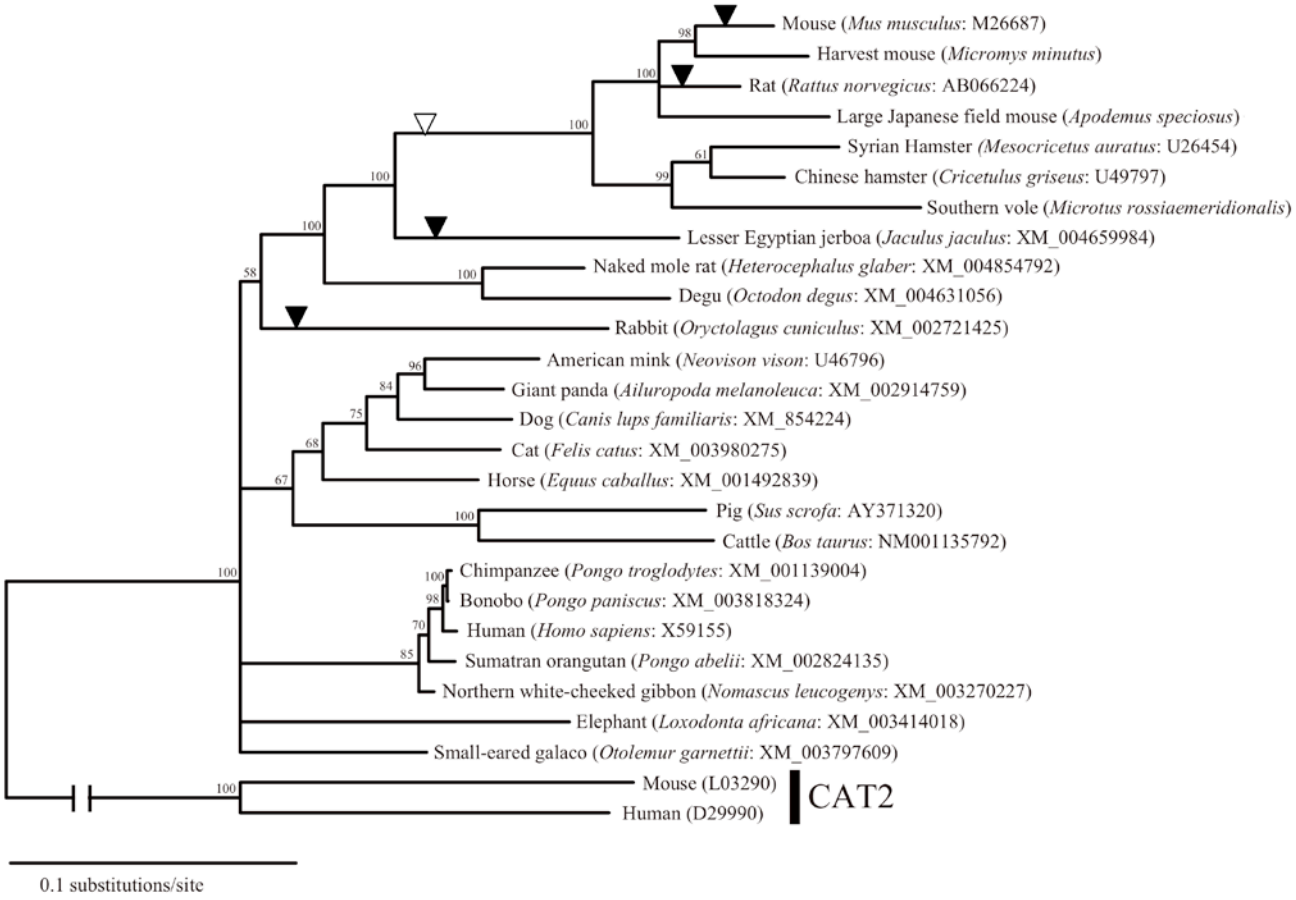
Mouse (*Mus musculus*: M26687)
Harvest mouse (*Micromys minutus*)
Rat (*Rattus norvegicus*: AB066224)
Large Japanese field mouse (*Apodemus speciosus*)
Syrian Hamster (*Mesocricetus auratus*: U26454)
Chinese hamster (*Cricetulus griseus*: U49797)
Southern vole (*Microtus rossiaemeridionalis*)
Lesser Egyptian jerboa (*Jaculus jaculus*: XM_004659984)
Naked mole rat (*Heterocephalus glaber*: XM_004854792)
Degu (*Octodon degus*: XM_004631056)
Rabbit (*Oryctolagus cuniculus*: XM_002721425)
American mink (*Neovison vison*: U46796)
Giant panda (*Ailuropoda melanoleuca*: XM_002914759)
Dog (*Canis lups familiaris*: XM_854224)
Cat (*Felis catus*: XM_003980275)
Horse (*Equus caballus*: XM_001492839)
Pig (*Sus scrofa*: AY371320)
Cattle (*Bos taurus*: NM001135792)
Chimpanzee (*Pongo troglodytes*: XM_001139004)
Bonobo (*Pongo paniscus*: XM_003818324)
Human (*Homo sapiens*: X59155)
Sumatran orangutan (*Pongo abelii*: XM_002824135)
Northern white-cheeked gibbon (*Nomascus leucogenys*: XM_003270227)
Elephant (*Loxodonta africana*: XM_003414018)
Small-eared galaco (*Otolemur garnettii*: XM_003797609)
Mouse (L03290)
Human (D29990)
CAT2
0.1 substitutions/site

Fig. 5

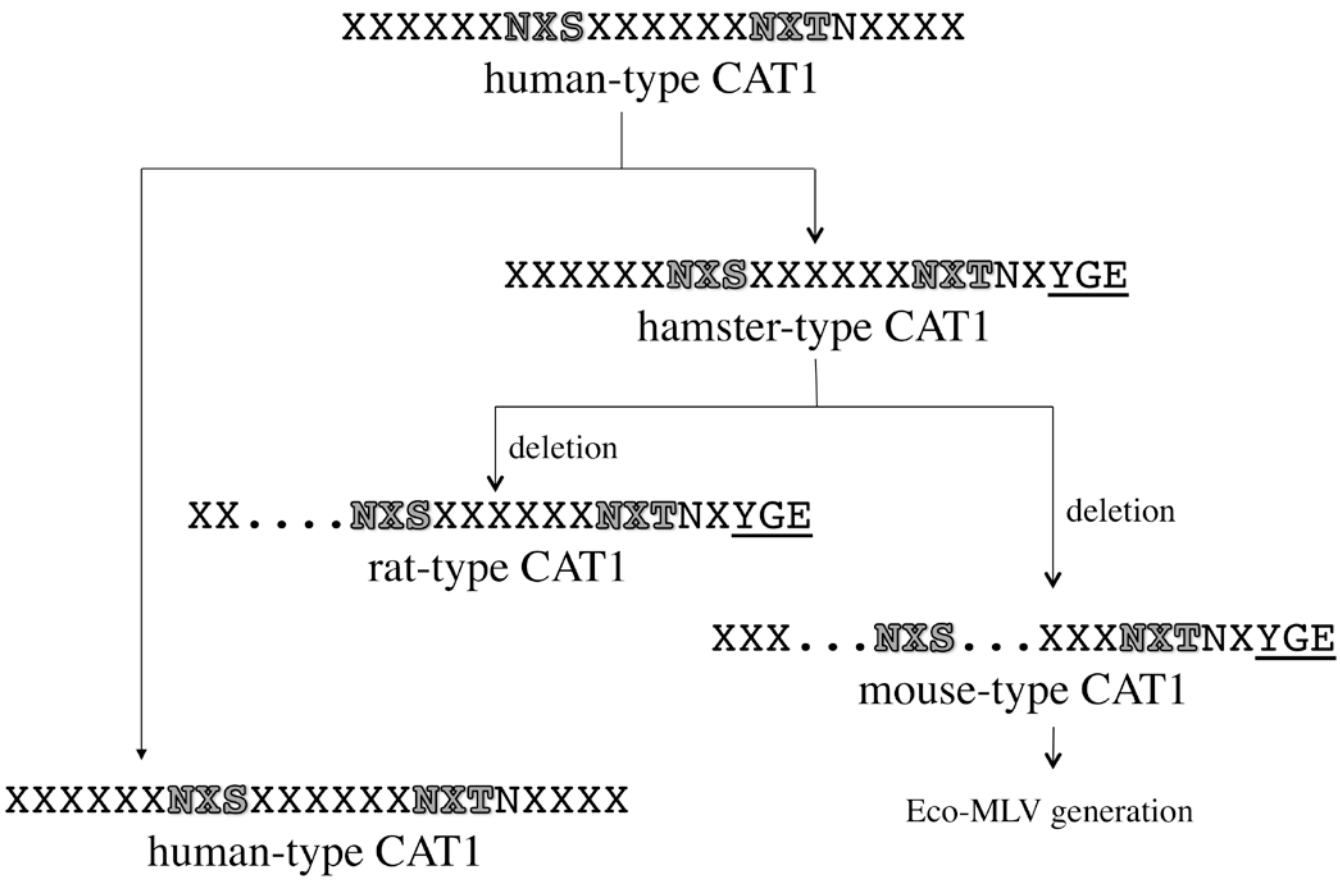
XXXXXXNXSXXXXXXNXTNXXXX
human-type CAT1
XXXXXXNXSXXXXXXNXTNXYGE
hamster-type CAT1
deletion
deletion
XX....NXSXXXXXXNXTNXYGE
rat-type CAT1
XXX...NXS...XXXNXTNXYGE
mouse-type CAT1
Eco-MLV generation
XXXXXXNXSXXXXXXNXTNXXXX
human-type CAT1

Table 1 Changes of first, second, and third bases of codons in mammalian CAT1s compared to mouse CAT1

species	Ratios of changes in upstream of ECL3 (%)			Ratios of changes in ECL3 (%)		
	first	second	third	first	second	third
R.norvegicus	0.8	0	6.9	4.0	3.4	5.1
M. minutus	0.8	0	0.5	5.1	4.5	6.2
A. speciosus	1.1	0	8.0	4.5	4.5	7.3
C. griseus	1.9	1.1	8.4	5.6	4.5	11.3
M. auratus	1.5	0.4	8.0	5.1	5.1	9.6
M. rossiaemeridionalis	3.4	0.4	9.6	5.6	5.1	11.3
H. glabe	2.7	0	9.2	6.8	6.8	16.4
O. Degus	2.7	0.4	10.7	7.3	6.8	13.6
O. cuniculus	3.8	1.1	10.0	7.9	9.0	18.6
J. jaculus	2.3	1.1	8.4	9.6	9.8	16.9