Identification and molecular variations of CAN-SINEs from the ZFY gene final intron of the Eurasian badgers (genus *Meles*)

Sara Tashima1, Yayoi Kaneko2, Tomoko Anezaki3, Minoru Baba4, Shuuji Yachimori5, Alexei V. Abramov6, Alexander P. Saveljev7 and Ryuichi Masuda1,*

1 Department of Natural History Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan
2 Tokyo University of Agriculture and Technology, Fuchu 183-8538, Japan
3 Gunma Natural History Museum, Tomioka 370-2345, Japan
4 Kitakyushu Museum of Natural History and Human History, Kitakyushu 805-0071, Japan
5 Shikoku Institute of Natural History, Susaki 785-0023, Japan
6 Zoological Institute, Russian Academy of Sciences, Saint-Petersburg 199034, Russia
7 Russian Research Institute of Game Management and Fur Farming, Russian Academy of Agricultural Sciences, Kirov 610000, Russia

Abstract. The short interspersed nucleotide elements (SINEs) are specific to the taxa and thought to be one of powerful phylogenetic gene markers. Especially, the SINE sequences, which exist uniquely in genome of order Carnivora, are named CAN-SINEs. Among the Eurasian badgers (genus *Meles*), a member of the family Mustelidae in order Carnivora, the Japanese badger (*M. anakuma*) was previously reported to have an insertion of CAN-SINE in the final intron of the zinc finger protein gene on Y chromosome (*ZFY*). In the present study, we examined occurrence of the CAN-SINE of the *ZFY* final intron in the Eurasian badgers, and three continental and four Japanese haplotypes were identified from a total of 40 male badgers. Among the Eurasian badger CAN-SINEs, a 12-bp deletion specific to the Japanese haplotypes was found, whereas the 12-bp region (non-deletion) in the continental haplotypes consisted of one 6-bp direct repeat and 6-bp microsatellite-like sequences. Moreover, the continental haplotypes were phylogenetically divided into three lineages: eastern Eurasia, Caucasus and western Eurasia. These genetic differentiations supported the classification recently proposing that genus *Meles* are grouped into the European badger (*M. meles*), the Southwest Asian badger (*M. canescens*), the Northwest & Central Asian badger (*M. leucurus*) and the Japanese badger (*M. anakuma*). In addition, the number of adenines in the poly A/T rich tails was polymorphic among all lineages of Eurasian badgers, and geographically variable within the Japanese badgers.

Key words: CAN-SINE, Eurasian badger, *Meles*, molecular variation, *ZFY* gene.

The short interspersed nuclear elements (SINEs), which are the 80–400 base-pair (bp) long nonretroviral retrotransposons, disperse in nuclear genome via RNA with the help of endonuclease or the complex of reverse transcriptase, and they do not encode the reverse transcriptase in their own sequences (Alberts et al. 2001). They acquire the retropositional activity from the corresponding long interspersed nuclear elements (LINEs), which encode reverse transcriptase and which are thought to be the origins of the SINEs (Okada et al. 1997). More than 10⁴ copies of SINEs exist in eukaryotic genome (Shedlock and Okada 2000). The majority of SINEs are reported to have derived from transfer RNAs (tRNAs) with the exceptions of primate Alu and rodent B1, which derived from 7SL RNA (Ullu and Tschudi 1984; Daniels and Deininger 1985; Labuda et al. 1991), and consist of three regions: a tRNA-related region, a tRNA-unrelated region and an A/T-rich region (Kramerov and Vassetzky 2005). The tRNA-related region contains an internal promoter of the RNA polymerase III, and the tRNA-unrelated region includes the LINE derived region, and the A/T-rich region relates to generate microsatellite sequences in Mustelidae (López-Giráldez et al. 2006). The SINE is thought to be one of powerful phylo-
genetic markers because of its wide dispersion in genome and no elimination after insertion to specific sites (Shedlock and Okada 2000). In addition, they are specific to the taxa (Van der Vlugt and Lenstra 1995), as reported in the primate Alu (Weiner et al. 1986; Batzer et al. 1990), cetacean CHR1 and CHR2 (Shimamura et al. 1997) and carnivore-specific SINE (CAN-SINE) (Minnick et al. 1992). Alberts et al. (2001) showed a possibility that the activities of transposable elements are related to speciation. The phylogenetic relationships by SINE markers were reconstructed in many mammalian taxa such as cetaceans (Shimamura et al. 1997), vespertilionids (Kawai et al. 2002) and felids (Pecon-Slattery et al. 2004).

The CAN-SINEs exist in representatives of all carnivore families (Vassetzky and Kramerov 2002). Pecon-Slattery et al. (2000) reported the SINE insertions within the nonrecombining region on the Y chromosome in the six carnivore families: Felidae, Canidae, Mustelidae, Ursidae, Procyonidae and Phocidae. For Mustelidae, Yamada and Masuda (2010) reported the independent insertion of CAN-SINE into the final intron of the zinc finger protein on Y chromosome (ZFY) in the ermine (Mustela ermine) and the Japanese badger (Meles anakuma).

The Eurasian badgers were once thought to comprise three groups of subspecies in the single species Meles meles: subspecies “meles” in Europe westward to the Volga River, subspecies group “arenarius-leptorhynchus” in the east of the Volga River to Siberia, and another subspecies group “amurenensis-anakuma” in Far East and Japan (Heptner et al. 1967). Recently, however, based on morphological features such as the skull, color pattern, baculum structure (Lynch 1994; Abramov 2001, 2002; Abramov and Puzachenko 2006), they are classified into three independent species: the European badger (Meles meles), the Asian badger (M. leucurus) and the Japanese badger (M. anakuma) (Lynch 1994; Abramov 2001; Wozencraft 2005). Based on the mitochondrial DNA (mtDNA) cytochrome b phylogenetic analysis, Kurose et al. (2001) reported the clear genetic distinction between the Japanese badgers and the Eurasian continental badgers. In addition, Marmi et al. (2006) showed the four mtDNA lineages in the Eurasian badgers: Europe, Southwest Asia, North & East Asia, and Japan. Abramov and Puzachenko (2005, 2006) already mentioned a remote position of badgers of Middle East from the others within the European badger Meles meles. Del Cerro et al. (2010) examined mtDNA and nuclear DNA sequences and proposed a new distinct species M. canescens for the Southwest Asian badger, in addition to the previous three species. Tashima et al. (2011) also supported the four lineages using new data of the mtDNA control region and the sex-determining region on Y chromosome (SRY), which is paternally inherited.

In the present study, we identified CAN-SINES from the ZFY final intron of not only the Japanese badgers but also continental Eurasian badgers. Based on a comparison of the haplotype sequences, we presented the molecular structural variations and phylogenetic features of the paternally inherited CAN-SINES, and then discuss molecular evolution of CAN-SINEs related to speciation of the Eurasian badgers.

Materials and methods

Samples and DNA extraction

Tissue samples (muscle, hair, skin and blood) were obtained from 40 male Eurasian badgers. Sampling locations are shown in Fig. 1. The DNA extraction was performed by using the QiAamp DNA Micro Kit (Qiagen) for hairs and the DNeasy Blood & Tissue Kit (Qiagen) for the other tissues. Extracted DNAs (about 1–100 ng/μl) were dissolved in 200 μl of TE buffer and preserved at 4°C until use.

Amplification and sequencing of SINE in the ZFY final intron

Two PCR primers to amplify the SINE inserted in the ZFY final intron were newly designed using the published sequence data of the Japanese badger (accession no. AB491600: Yamada and Masuda 2010): MELSN-F2 (5'-GGAGCAAGACAAATAATTCAG-3') and MELSN-R2 (5'-CATTTAAAGCCACATGTATTGG-3'). A total of 50 μl of the PCR reaction solution consisted of 5 μl of 10 × PCR buffer (Takara), 4 μl of 2.5 mM dNTP mixture, 0.25 μl of rTaq DNA polymerase (5 units/μl, Takara), 0.5 μl of each of the above two primers and 5 μl of the DNA extract (about 1–100 ng/μl). The PCR reactions were carried out in a PCR thermal cycler (TP600, Takara) with the following reaction conditions: one cycle of 94°C for 3 min; 35 cycles of 94°C for 1 min, 57°C for 1 min, 72°C for 1 min; and one cycle of 72°C for 10 min. The PCR products were purified with the QIAquick PCR Purification Kit (Qiagen) and subjected to the sequencing reaction. In the present study, two sequencing primers were newly designed based on the published ZFY final intron sequence of the Japanese badger (acces-
MELSN-SQ1 (5’-GCCACATGTATTTGGATATCTA-3’)
and MELSN-SQ2 (5’-AGACAAATAATTCAGTTGTC
CAT-3’), and 3’-labeled with Texas Red. The cycle PCR
condition was one cycle of 94°C for 3 min; 30 cycles of
94°C for 30 sec; 55°C for 30 sec; 72°C for 1 min; 1 cycle
of 72°C, 7 min, using the Thermo Sequence pre-mixed
cycle sequencing kit (Amersham). The PCR products
were run using an automated DNA sequencer (SQ5500,
Hitachi).

Fig. 1. (A) The sampling locations in the Eurasian Continent. Numerals in circles are the numbers of individuals from those locations. (B) The
precise sampling locations on the Japanese islands. Numerals in parentheses are the numbers of individuals from those locations. The boundaries
of prefectures are shown by lines within three main islands of the Japanese archipelago.
Phylogenetic analysis of nucleotide sequences

A sequence alignment was performed by software CLUSTAL W (Thompson et al. 1994). The published CAN-SINE sequence of the tyrosine aminotransferase gene (AF163863: Leib et al. 2005) and that of the X-chromosome (X52381: Lavrentieva et al. 1991) of the American mink (Neovison vison) were used as references. A median-joining network was constructed by using software NETWORK (fluxus-engineering.com http://www.fluxus-endineering.com/sharenet.htm: Bandelt et al. 1999), where a gap was counted as one nucleotide substitution.

Results

Sequence variations and structures of the SINE insertion in the ZFY final intron

The two PCR primers newly designed for targetting the CAN-SINE sequence of the Japanese badger ZFY final intron successfully amplified the highly homologous CAN-SINE sequences from a total of 40 males of the Japanese and continental populations of the Eurasian badgers (Fig. 2). Of the determined SINE sequences (197–210 bp), 21 nucleotide sites were polymorphic including 12-bp insertions/deletions (indels), and four haplotypes from Japan (SNJ1, SNJ2, SNJ3 and SNJ4) and three haplotypes from the Eurasian Continent (SNE1, SNC1 and SNW1) were identified (Fig. 2). The substitutions among the haplotypes were transitions between T and C at nucleotide sites 102 and 161, except the A/T rich region. A 12-bp deletion occurred at nucleotide sites 135–146 in the Japanese haplotypes, compared from the continental ones (Fig. 2). The other nucleotide differences between the haplotypes were polymorphisms of the numbers (14–17 nucleotides) of the A/T repeats at the A/T rich tail (Fig. 2). Even within the Japanese badgers, the geographical variation in the numbers of the A/T repeats led to classification of the four Japanese haplotypes.

An alignment and components of SINE sequences

![Fig. 2](image_url). The components of the seven CAN-SINE haplotypes (SNJ1 to SNJ4, SNE1, SNC1, and SNW1) identified from the Eurasian badger. Their accession numbers are shown together with the seven CAN-SINE haplotypes. RNA polymerase III boxes A and B and terminator, and the polypyrimidine (C/T shown as Y) tracts are enclosed by solid lines. Direct repeat motifs are enclosed by broken lines. Numerals (end characters) and asterisks above the SNJ1 sequence indicate nucleotide site numbers and polymorphic sites of haplotypes, respectively. AF163863 (Leib et al. 2005) and X52381 (Lavrentieva et al. 1991) are accession numbers of the published CAN-SINE sequences of the American mink (Neovison vison).
obtained in the present study showed the high similarity with the previously reported CAN-SINEs (AF163863 and X52381) from the American mink (Fig. 2). The CAN-SINE haplotype sequences included RNA polymerase III boxes A and B and terminator, two polypyrimidine tracts (C/T shown as Y in Fig. 2), and two direct repeats (GCCTGC and GCCTAC), each of which adhered to the 5' ends of the two polypyrimidine tracts, respectively. In addition, the 12-bp deletion between the two polypyrimidine tracts was found to be specific to the Japanese haplotypes (SNJ1 to SNJ4), whereas the 12-bp region (non-deletion) in the continental haplotypes (SNE1, SNC1 and SNW1) consisted of one direct repeat (GCCTGC) at the 5' end and 6-bp microsatellite-like sequences (TCTCTC) at the 3' end (Fig. 2). Consequently, in the entire sequence of these CAN-SINEs, the four Japanese haplotypes contained two direct repeats, whereas the three continental haplotypes had three.

The CAN-SINE sequences determined in the present study were deposited to DDBJ/GenBank/EMBL databases with the following accession numbers: AB551115–AB551121.

**Molecular phylogeny of CAN-SINE in the badger ZFY final intron**

A median-joining network (Fig. 3) showed that the SINE haplotypes were grouped into four lineages (Lineages 1–4). Lineage 1 consisting of the four Japanese haplotypes (SNJ1, SNJ2, SNJ3 and SNJ4) was apparently separated from the other three lineages, all of which consisted of the continental haplotypes with the 12-bp region (non-deletion at nucleotide sites 135–146) (Figs. 2 and 3).

Table 1 showed the distribution and frequencies of the badger CAN-SINE haplotypes in Eurasia. SNE1 (Lineage 2) was identified from eastern Eurasian badgers of Far East, Transbaikalia and Chelyabinsk Province. SNW1 (Lineage 4) was found in western Eurasian badgers of Kirov and Leningrad Provinces. SNC1 (Lineage 3) was from Armenia (Table 1). The distribution pattern shows the genetic differentiation among the sampling localities in the Eurasian Continent.

Among the four Japanese haplotypes, SNJ1 was most closely related to the continental haplotypes (Fig. 3), and shared by 15 badgers from Honshu Island and three from Shikoku Island (Table 1). SNJ2 was specific to Honshu. SNJ3 and SNJ4, both of which were identified from Kyushu Island, were connected with each other (Fig. 3; see Fig. 1 for geographical locations of the continent and Japanese islands).

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Number of animals</th>
<th>Japan (M. anakuma)</th>
<th>Eurasian Continent</th>
<th>Caucasus (M. meles)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Honshu</td>
<td>Shikoku</td>
<td>Kyushu</td>
</tr>
<tr>
<td>SNJ1</td>
<td>18</td>
<td>15</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>SNJ2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNJ3</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNJ4</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNE1</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>SNW1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNC1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>17</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>
Discussion

Components of CAN-SINEs in the Eurasian badger ZFY final intron

The majority of SINEs consist of three parts (tRNA-related and -unrelated regions and A/T rich regions), and the most common source of SINEs is tRNA of lysine (Shedlock and Okada 2000). Furthermore, the typical CAN-SINEs have the two promoter regions of RNA polymerase III (Pol III boxes A and B) in the tRNA-related region, polypyrmididine tracts, and the A/T rich region overlapping with a polyadenilation signal (AATAAA), and direct repeats in the tRNA unrelated region (Vassetzky and Kramerov 2002; López-Giráldez et al. 2005). The SINEs of the badger ZFY final intron obtained in the present study possessed these characteristic structures. Thus, the structural homologies clearly show that the badger SINEs can be classified in CAN-SINEs.

In both Pol III boxes A and B, there were no substitutions among the haplotypes obtained in the present study, whereas one substitution (at nucleotide site 161) in the second polypyrmididine tract was found between haplotypes of the Japanese/eastern Eurasian badgers and those of the western Eurasian/Caucasian badgers. The polypyrmididine tract was reported in not only the most families of order Carnivora (Vassetzky and Kramerov 2002), but also the hedgehog (Mesechinus daurius: Borodulina and Kramerov 2001) and tree shrew (Tupaia glis: Ten et al. 2006). Moreover, because the sequences and length of the polypyrmididine tract vary even between phylogenetically related species such as the wolf (Canis lupus) and red fox (Vulpes vulpes), it is assumed that the generation of this tract is associated with expansion of microsatellite-like regions (Vassetzky and Kramerov 2002). Especially, we found that the direct repeats (GCCTGC and GCCTAC) adhere to the 5′ end of the first and second polypyrmididine tracts, respectively (Fig. 2). Vassetzky et al. (2003) and Ten et al. (2006) reported that such direct repeats consisting of 3–6 bp could generate internal duplications and deletions in mammal SINEs. They also mentioned that one of the direct repeats was present within the deletions. Actually in the present study, within the 12-bp indels of all the three continental badgers’ CAN-SINEs, one of the direct repeats (GCCTGC) was found to adhere to the 5′ end of 6-bp microsatellite-like sequences (CTCTCT) (Fig. 2). These suggest that the formation of two polypyrmididine tracts and the 12-bp indels found in the badger CAN-SINEs has been mediated by the 6-bp direct repeats.

In addition, the repeat number of A/T in the downstream of the Pol III terminator was variable (Fig. 2). The poly A/T rich tail was known to be microsatellite-like sequences in the human (Shriver et al. 1992) and pig (Ellegren 1993), and it was thought to be associated with generation of microsatellites in mustelids (López-Girárd et al. 2005, 2006). Polymorphisms of the nucleotide repeat number in the A/T rich tail within the species were reported in the European wild cat (Felis silvestris: Pecon-Slattery et al. 2004). The present study first shows the intraspecific and interspecific variations of length in the A/T rich region of the Eurasian badgers.

Phylogenetic features of the CAN-SINE sequences for the Eurasian badgers

The four lineages identified in the present study are separately localized following geographical distributions of the Eurasian badgers. It means that the molecular differentiation of the CAN-SINE in the ZFY final intron corresponds with the morphological differentiation of the Eurasian badgers classified into the four groups: the classification recently proposed by Del Cerro et al. (2010) using an agreement with molecular phylogeny. In addition, the SINE lineages were in agreement with those of mtDNA (Marmi et al. 2006; Del Cerro et al. 2010; Tashima et al. 2011). Yamada and Masuda (2010) reported that the CAN-SINE sequence was inserted to the ZFY final intron independently to the Japanese badger and the ermine, after diversification of at least nine Asian mustelid species. The present study revealed that the CAN-SINE could be inserted once prior to speciation of the Eurasian badgers, because it occurs commonly at the same nucleotide position in the ZFY final intron of genus Meles.

It is likely that the 12-bp region was deleted from the CAN-SINE of the ancestor of the Japanese badgers, rather than this region was inserted into the ZFY final intron of the continental badgers. That is, the 12-bp deletion could have caused secondarily after the SINE insertion into the badger ZFY final intron. The first reason is that the previously reported mustelid CAN-SINEs in addition to the haplotypes identified from the present study also contain the homologous 12-bp sequences (6-bp direct repeat and 6-bp microsatellite-like sequences) as shown in Fig. 2. The second is that the ancestors of the Japanese badgers are thought to have immigrated from the Eurasian continent via land bridges during Pleistocene (Kawamura et al. 1989; Kawamura 1991).
Thus, the finding of the 12-bp deletion within the CAN-SINEs of the Japanese badgers (Fig. 2) suggests their geographical and genetic separation from the continental badgers.

It is interesting that single haplotypes of CAN-SINE were identified from three continental lineages, whereas four haplotypes were found in the Japanese lineage. Tashima et al. (2011) reported the similar frequencies of identified haplotypes of the SRY gene as another paternally inherited gene: 6 types in 33 Japanese badgers; 2 types in 8 badgers of Chelyabinsk Province; 1 type in both 2 badgers of Kirov Province and 2 badgers of Leningrad Province; 1 type in 2 badgers of Armenia. Although this could be attributed to a small number of samples and sampling localities in the continent, the findings suggest a higher genetic variability in the Japanese badgers. The variation of mtDNA haplotypes in the Japanese badgers was higher than that in the continental badgers (Tashima et al. 2011). Thus, the geographic isolation within islands could have resulted in the relatively higher genetic variability in the Japanese badgers.

The four CAN-SINE haplotypes identified from the Japanese badgers were distributed region-specifically: the number of single adenine repeats in the tail part was lower (15 in SNJ3 and 14 in SNJ4) in individuals from Kyushu Island than those from Honshu (16 in SNJ1 or 17 in SNJ2) and Shikoku (16 in SNJ1). Geographical isolation of Honshu, Shikoku and Kyushu Islands by the Seto Inland Sea, which started 7,000–5,000 years ago (Ohshima 1990), would have genetically differentiated the island populations within Japan. Tashima et al. (2010) examined biparentally inherited microsatellite loci of the Japanese badgers and showed the genetic differentiation among Honshu, Shikoku, and Kyushu Islands, suggesting the geographic isolation of the Japanese populations via the Seto Inland Sea. Long uninterrupted mono-nucleotide repeats as seen in the 3′ end of the CAN-SINEs are likely related to the time since its mutation, that is, the ages of longer repeats are thought to be younger (Batzer et al. 1990; Ellegren 1993). The variability in the number of nucleotides in the poly A/T tail of the badger CAN-SINEs could have been generated by nucleotide slippages after the geographical isolation between the continent and the Japanese islands and even among the Japanese islands.

Acknowledgments: We thank T. Tsujimoto, H. Tanaka, S. Dakemoto, V. Sotnikov, O. Kruglova, and I. Khorozyan for supplying badger samples. This study was supported in part by Grants-in-Aid for Scientific Research (Nos. 17405012 and 22405003) from the Japan Society for the Promotion of Science and the Russian Foundation for the Basic Research (No. 09-09-00073).

References


Received 15 September 2010. Accepted 28 December 2010.