Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 365 (2006) 67–73
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Probing the secondary structure of salmon SmaI SINE RNA

Hiroko Kawagoe-Takaki [a,1], Nobukazu Nameki [a,1,2], Masaki Kajikawa [a], Norihiro Okada [a,b,*]

[a] *Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259-B-21 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan*
[b] *Department of Evolutionary Biology and Biodiversity, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan*

Received 13 June 2005; received in revised form 11 August 2005; accepted 7 September 2005
Available online 13 December 2005

Abstract

SmaI is a short interspersed element (SINE) of the salmon genome, and is derived from $tRNA^{Lys}$. We probed the secondary structure of SmaI SINE RNA by enzymatic cleavage and found that the RNA structure comprises three separate domains. The 5′-terminal region (the 5′ domain) forms a tRNA-like cloverleaf structure, whereas the 3′-terminal region (the 3′ domain) forms an extended stem-loop. The loop region is thought to be recognized by the reverse transcriptase (RT) encoded by the long interspersed element (LINE). The two structural domains are linked by a single-stranded region (the linker domain). Our melting profile analyses indicated the presence of two structural domains having different thermal stabilities, thus supporting the domain composition described above. Based on these results, we discuss the structural generality and evolutionary advantage of the domain composition of SINE RNA.

Keywords: Transposable element; Retroposon; Retrotransposon; Retrotransposition; LINE; Enzymatic probing

1. Introduction

Short interspersed elements (SINEs) and long interspersed elements (LINEs; also called non-long terminal repeat (non-LTR) retrotransposon) are transposable elements in eukaryotic genomes that mobilize through an RNA intermediate. These elements are first transcribed into RNA, which is then reverse-transcribed into complementary DNA (cDNA) that is subsequently integrated at a new location within the host genome. This "copy and paste" mechanism is called retrotransposition, a process that expands the number of SINEs and LINEs such that they often occupy a considerable portion of a eukaryotic genome. For example, there are ~1,500,000 copies of SINEs and ~850,000 copies of LINEs in the human haploid genome, which comprise ~13% and ~21% of the genome, respectively (Lander et al., 2001). Furthermore, these elements also are thought to have a significant impact on the evolution and complexity of eukaryotic genomes (Moran and Gilbert, 2002; Kazazian, 2004).

LINEs are approximately 4–7 kilobase pair (kbp) long and encode an endonuclease (EN) and a reverse transcriptase (RT), both of which are required for LINE retrotransposition (Xiong and Eickbush, 1988a; Xiong and Eickbush, 1988b; Martin et al., 1995; Feng et al., 1996; Moran et al., 1996). The LINE protein is thought to bind its own RNA and return to the nucleus. The LINE EN creates a nick in the DNA of the host genome, and the LINE RT initiates reverse transcription of its own RNA from the 3′ hydroxyl group generated by this nick (Luan et al., 1993; Cost et al., 2002). Thus, reverse transcription is coupled with the target site cleavage, a process called target-primed reverse transcription (TPRT) (Luan et al., 1993). The newly synthesized LINE copy is integrated into the host genome by TPRT. In contrast, SINEs are relatively short (approximately 100–500 bp) and do not encode proteins required for their own retrotransposition. Typical SINEs are composed of two parts, a tRNA-related region and a tRNA-unrelated region, except for human SINE Alu and mouse SINE B1, which have a

Abbreviations: cDNA, DNA complementary to RNA; kbp, kilobase pair; LINE, long interspersed element; non-LTR, non-long terminal repeat; PAGE, polyacrylamide gel electrophoresis; SINE, short interspersed element; TPRT, target-primed reverse transcription.

* Corresponding author. Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259-B-21 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan. Tel.: +81 45 924 5742; fax: +81 45 924 5835.
E-mail address: nokada@bio.titech.ac.jp (N. Okada).
[1] These authors made equal contributions to this work.
[2] Present address: Department of Biological and Chemical Engineering, Faculty of Engineering, Gunma University, Kiryu, Gunma 376-8515, Japan.

0378-1119/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2005.09.027

7SL RNA-related region instead of the tRNA-related region (Weiner, 1980; Ullu and Tschudi, 1984; Okada, 1991a,b). Recently, a novel class of SINEs, which has a 5S rRNA-related region instead of the tRNA-related region, was also found in the zebrafish genome (Kapitonov and Jurka, 2003). The tRNA-related region of SINEs derives from a host tRNA, and includes an RNA polymerase III-dependent internal promoter. In the tRNA-unrelated region, the 3′ tails of some SINEs are almost identical to those of certain LINEs (Ohshima et al., 1996; Okada et al., 1997).

Recent experiments provide evidence for the SINE amplification mechanism. We previously isolated a SINE and LINE from the eel genome, named UnaSINE1 and UnaL2, respectively, which have a common 3′ tail of ~60 bp (Kajikawa et al., 2005). Using these elements, we showed that UnaL2 can retrotranspose in cultured HeLa cells and that the conserved 3′ tail of UnaL2 is required for its own retrotransposition (Kajikawa and Okada, 2002). We also showed that UnaL2 can recognize the conserved 3′ tail of UnaSINE1 in *trans* and mobilize a transcript that has the 3′ tail of UnaSINE1 (Kajikawa and Okada, 2002). These results indicate that SINEs are mobilized by LINEs through their conserved 3′ tails. Human LINE L1 can mobilize human SINE Alu via the poly A tail at the 3′ end, although L1 and Alu have no such common 3′ tails (Dewannieux et al., 2003).

During amplification, the UnaL2 RT appears to bind to the conserved 3′ tail of UnaSINE1 (and UnaL2) RNA, recognizing a specific structure (Kajikawa and Okada, 2002; Baba et al., 2004). In addition, SINE RNAs bind to proteins other than the LINE protein in cells. For example, SRP9/14 proteins of the signal recognition particle bind to Alu RNA (Sarrowa et al., 1997), and this binding is proposed to increase the retrotransposition efficiency of Alu SINE (Boeke, 1997; Dewannieux et al., 2003). Thus, SINE RNA structure dictates protein binding specificity, thereby influencing SINE retrotransposition efficiency. Little is known, however, about SINE RNA structure. To address this issue, we used RNA of the salmon SINE, SmaI, the 3′ tail of which is quite similar to that of UnaL2 and UnaSINE1 (Kajikawa et al., 2005). We used enzymatic cleavage to probe the RNA secondary structure of SmaI. We also obtained additional structural information by measuring thermal melting profiles of several SINE RNAs.

2. Materials and methods

2.1. RNA preparation

The full-length SmaI sequence was amplified from chum salmon (*Oncorhynchus keta*) genomic DNA by PCR using primers SmaT7F (5′-CTAATACGACTCACTATAGGTCCTTCTGTAGCTCAGT-3′) and SmaAseIR (5′-TTTTGTTATTAATATGCCATTTAGCAGACG-3′). The PCR product (T7 promoter and full-length SmaI) was subcloned into pUC18/SmaI, and the resulting plasmid containing the SmaI consensus sequence (Kido et al., 1991) was designated pSmaT7. The pSmaT7 plasmid was used to generate full-length SmaI RNA. We altered the C to T at position 80 relative to the 5′ end of SmaI in pSmaT7 using QuikChange Site-directed Mutagenesis kit (Stratagene, USA), producing an NdeI site (CATACG to CATATG). The resulting plasmid was designated pSmaH and was used to prepare the 5′ domain RNA of SmaI. The 3′ domain of SmaI was amplified from the pSmaT7 DNA by PCR using primers Sma1Fsl3F (5′-CTAATACGACTCACTATAGTATGCACACATGACTGTAAG-3′) and SmaAseIR. The PCR product was cloned in the pUC18/SmaI vector. The resulting plasmid was designated pSmasl3 and was used to prepare the 3′ domain RNA of SmaI. Using primers M13M4 (5′-GTTTTCCCAGTCACGAC-3′) and M13RV (5′-CAGGAAACAGCTATGAC-3′), we amplified the SmaI sequences (and the T7 promoter) by PCR from pSmaT7, pSmaH or pSmasl3 and then digested the PCR products with AseI (pSmaT7 and pSmasl3) or NdeI (pSmaH). SmaI RNAs were transcribed from the resulting DNA fragments using the AmpliScribe T7 Transcription kit (Epicentre Technologies, USA) according to the manufacturer's instructions. The transcribed RNA was purified by 7 M urea/10% polyacrylamide gel electrophoresis (PAGE).

The full-length HpaI SINE sequence was amplified from masu salmon (*Oncorhynchus masou*) genomic DNA by PCR using primers HpaF-T7 (5′-CTAATACGACTCACTATAGGGGGCGGCAGGGTAG-3′) and HpaStuR (5′-TTTAGGCCTTTATTTAACTAGGCAAGTC-3′). The PCR product was cloned into the pUC18/SmaI vector, and the resulting plasmid containing the HpaI general consensus sequence (Kido et al., 1995) was designated as pHpaIT7Gcons. The HpaI sequence (and the T7 promoter) was amplified from pHpaIT7Gcons DNA by PCR using primers M13M4 and M13RV, and the PCR product was digested with StuI. The full-length HpaI RNA was prepared from the resulting fragment by T7 transcription as described above.

The full-length UnaSINE2 sequence was amplified from the plasmid UnaSINE2-8 (Accession number AB179634; Kajikawa et al., 2005) DNA by PCR using primers UnaS2FT7 (5′-CTAATACGACTCACTATAGGGGGCGATATAGCTCAG-3′) and UnaS2DraR (5′-ACTTTAAATGGTAAATGGACTGCA-3′). The PCR product was cloned in the pUC18/SmaI vector, and the resulting plasmid was designated as pUnS2T7. The UnaSINE2 sequence (and the T7 promoter) was amplified from pUnS2T7 DNA by PCR using primers M13M4 and M13RV and was digested with DraI. The full-length UnaSINE2 RNA was prepared from the resulting DNA fragment by T7 transcription as described above.

2.2. Enzymatic probing

The enzymatic reactions were carried out essentially as described by Felden et al. (1997). 5′ and 3′ end-labeled RNAs were prepared as described Felden et al. (1997) and purified by 7 M urea/10% PAGE. Prior to the enzymatic digestion, the labeled RNA was heated at 85 °C for 3 min and cooled slowly to 20 °C in the reaction buffer as below. Digestion with the various ribonucleases (V1, S1, and T1) was performed on both 5′- and 3′-labeled RNAs (10,000–20,000 cpm per reaction, depending on the experiments) supplemented with 1 μg