FEBS Letters 580 (2006) 2301–2305

Extensive adenosine-to-inosine editing detected in *Alu* repeats of antisense RNAs reveals scarcity of sense–antisense duplex formation

Yukio Kawahara[a,b,*], Kazuko Nishikura[a]

[a] *The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA*
[b] *Department of Neurology, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*

Received 24 December 2005; revised 3 March 2006; accepted 13 March 2006

Available online 24 March 2006

Edited by Takashi Gojobori

Abstract **One type of RNA editing converts adenosine residues to inosine in double-stranded regions. Recent transcriptome analysis has revealed that numerous *Alu* repeats, present within introns and untranslated regions of human transcripts, are subject to this A → I RNA editing. Furthermore, it revealed global transcription of antisense RNAs. Here, we demonstrate that antisense RNAs are also edited extensively but only in their *Alu* repeat sequences, and editing does not extend to the surrounding sequence. Our findings imply that sense and antisense RNAs form two separate intramolecular double-stranded RNAs consisting of inversely oriented *Alu* repeats, but rarely form intermolecular duplexes.**

Keywords: Antisense RNA; ADARs; *Alu* repeats; RNA editing; Human transcriptome

1. Introduction

RNA editing refers to a post-transcriptional modification of the base sequence of pre-mRNA, and is recognized as a genetic mechanism for generating RNA and protein diversity [1–4]. Adenosine deaminases that act on RNA (ADARs) are involved in the mechanism of RNA editing that specifically modifies adenosine residues of double-stranded RNAs (dsRNAs) to inosine (A → I RNA editing). Three distinct ADAR genes (ADAR1–3) have been identified in mammals [1–4]. ADARs recognize and edit specific adenosine residues through their interaction with the complete or incomplete dsRNA structure formed between the editing site located within an exon sequence and its complementary sequence usually located in an intron sequence [5–7]. RNA editing of protein-coding sequences could alter codons and functions of target genes [5–7]. Although several cases of A → I editing of protein-coding RNA targets has been identified, bioinformatics analysis of human mRNA and EST (expressed sequence tags) databases revealed a large number of A → I RNA editing sites in non-coding intron and untranslated-region RNA sequences [8–11]. Surprisingly, a majority of editing sites (more than 10 000 sites mapped in ~2000 different genes) resides within repetitive sequences, such as *Alu* and LINE [8–11]. *Alu* repeats, comprising more than 10% of the human genome, are about 300 nucleotides in length and have relatively high homology among subfamilies [12,13]. An *Alu* element is likely to form an intramolecular RNA duplex with a nearby inverted *Alu* repeat sequence, which then could serve as a substrate for A → I RNA editing. In agreement with this prediction, the extent of RNA editing in *Alu* repeats depends on the distance between two inverted *Alu* repeats [9,11].

Recent transcriptome analysis indicates that a large fraction ($\geqslant 70\%$) of human transcripts are derived from non-coding antisense strand DNA sequences [14–17]. Among naturally occurring antisense RNA transcripts, *cis*-encoded antisense RNAs are transcribed from the opposite strand of the same genomic locus and frequently have a long and perfect complementarity to the sense transcript. Notably, sense–antisense RNA pairs may be coexpressed more frequently than the rate expected by pure chance [16], leading to the hypothesis that human genes are regulated by antisense transcripts [14–16]. Although there are a few reported cases of sense–antisense intermolecular duplex formation and consequent regulation of the sense mRNA expression [18], it is not known how frequently coexpression of sense and antisense transcripts and formation of an intermolecular dsRNA occur.

In this study, we demonstrate the extensive editing of *cis*-encoded antisense transcripts coexpressed with their sense strand RNAs in human lung and NT2-N neurons and, most importantly, that antisense RNA editing occurs intramolecularly in the limited regions containing *Alu* sequences, as predicted recently by bioinformatics analysis of the human antisense transcriptome [19]. In addition, our in vitro studies confirmed that *cis*-encoded sense and antisense RNAs remain single-stranded except the region containing *Alu* repeat sequences. Our results imply that intermolecular pairing of sense and antisense RNAs is a rare occurrence. Formation of intramolecular dsRNA structure made of *Alu* repeats may function in some cases to prevent formation of the intermolecular RNA duplex between sense and antisense RNAs that could have detrimental effects on processing and expression of sense mRNAs [18].

2. Materials and methods

2.1. Quantification of RNA editing

Maintenance of NT2 cells and their in vitro induction to NT2-N neurons were described previously [20]. Total RNA extracted from human adult lung was obtained from Clontech (Palo Alto, CA). 1 μg of total RNA (human adult lung or NT2-N neurons) was first treated with 1.0 U of DNase I (Invitrogen Corp., Carlsbad, CA) in 7 μl of the

*Corresponding author. Fax: +215 898 3911.
E-mail address: ykawahara@wistar.org (Y. Kawahara).

0014-5793/$32.00
doi:10.1016/j.febslet.2006.03.042

DNase I reaction mixture at 37 °C for 15 min. DNase-treated RNA was denatured along with 1 μl of 10 μM gene-specific primer (Figs. 1A, 2A and Table 1) and 2 μl of 10 mM dNTP mix at 80 °C for 5 min. Denatured RNA was mixed with 4 μl of 5× cDNA Synthesis Buffer, 1 μl of 0.1 M DTT, 40 U of RNaseOUT and 15 U of Thermo-Script (Invitrogen). First-strand cDNA was synthesized at 55 °C for 60 min and the reaction was terminated by incubating at 85 °C for 5 min, followed by incubation at 37 °C for 20 min with 2 U of RNase H.

The resultant cDNA was then amplified by PCR in a reaction mixture of 50 μl containing 200 nM of each primer (Figs. 1A, 2A and Table 1), 1 mM dNTP Mix, 5 μl of 10 × PCR buffer, and 1 μl of Advantage 2 Polymerase mix (Clontech). The PCR amplification began with a 1-min denaturation step at 95 °C, followed by 35 cycles of denaturation at 95 °C for 10 s, annealing at 64 °C for 30 s, and extension at 68 °C for 60 s. After gel purification, PCR products were subcloned using the TOPO TA cloning kit (Invitrogen), and then the editing frequency was determined by sequencing from both directions more than 24 individual clones containing appropriately sized inserts. Because the frequency of mismatches other than A → G changes between the genomic sequence and the cDNA sequences cloned were less than 4.0%, only adenosine residues that showed A → G changes with >4.0% frequency were accounted as A → I editing sites.

2.2. In vitro RNA editing assay

Plasmid pBS-CNNM3 contains a 1092-base-pair (bp) fragment of a partial *CNNM3* intron 2 including two inverted repeats of *Alu* sequences (Fig. 1A). The DNA fragment was PCR-amplified using human genomic DNA and PCR primers Bam-CFW3 (5′-CATCCGGATCCCCCTAGGTGCTTTTGTTTTTGC-3′) and Xho-CDW3 (5′-ATAAGAATCTCGAGCTTGGAGCTGGCTGAAGAGAAT-3′). Bam-CFW3 contains a *Bam*HI recognition site (underlined), and Xho-CDW3 contains an *Xho*I recognition site (underlined). The PCR products were digested with *Bam*HI and *Xho*I, then inserted into pBluescript II KS vector (Stratagene, La Jolla, CA) linearized with the same restriction enzymes.

pBS-CNNM3, linearized with *Xho*I or *Bam*HI, was transcribed with T7 or T3 RNA polymerase (20 U) in the presence of a trace amount of [α-^{32}P] UTP at 37 °C for 60 min to synthesize the sense or antisense *CNNM3* intron 2 RNA, respectively, as described previously [21].

An equal amount of the sense and antisense *CNNM3* intron 2 RNA (2 or 20 fmol each) was mixed in 4 μl of water with or without pre-incubation at 80 °C for 5 min followed by rapid chill on ice for 5 min. The sense and antisense RNA mixture (0.02 or 0.2 nM) was then subjected to in vitro editing at 30 °C for 60 min with 50 ng of recombinant ADAR1 (p110) or ADAR2 protein as described previously [21]. Quantification of RNA editing frequency was done as already described above.

3. Results and discussion

Recent transcriptome analysis revealed global expression of *cis*-encoded antisense strand RNAs [14–17]. However, formation of sense–antisense intermolecular RNA duplexes and regulation of sense RNA expression have been a subject of debate [15,18,19]. Any long RNA duplexes such as those formed between sense and antisense RNA pairs, if any, would be exten-

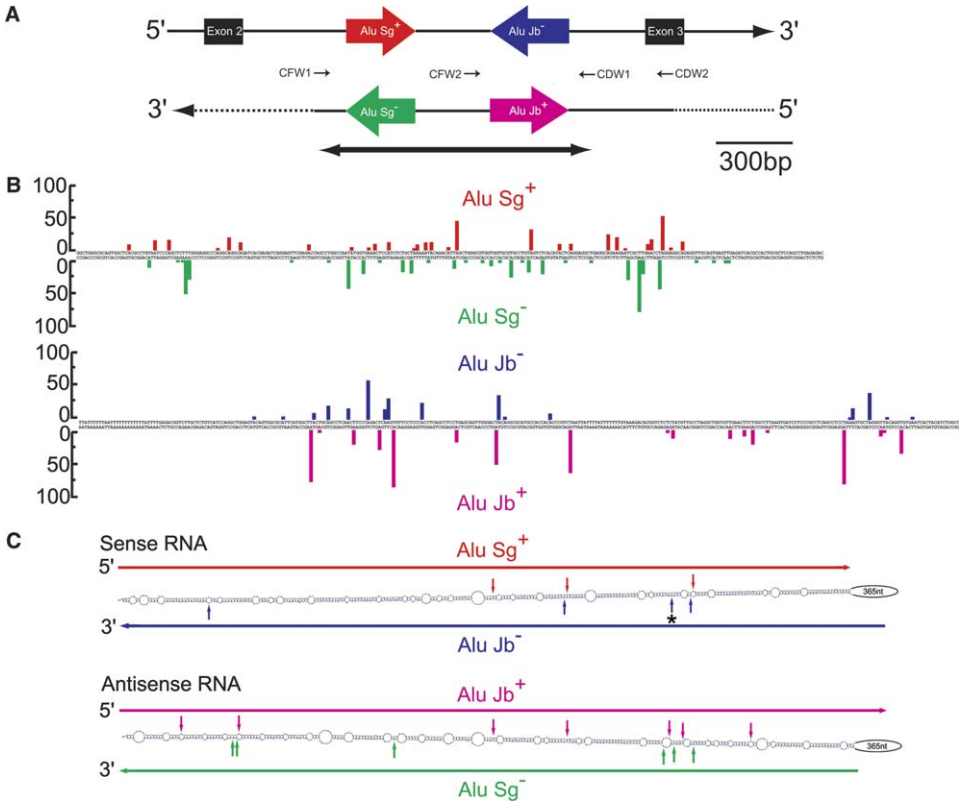

Fig. 1. Sense and antisense *Alu* editing in intron 2 of *CNNM3*. (A) Two inverted *Alu* repeats present in intron 2 of *CNNM3* are shown as red and blue arrows. Antisense RNA corresponding to this region detected in this study is indicated also. The *Alu* repeats of the antisense RNA are presented as green and pink arrows. The small arrows indicate the RT-PCR primers, their directions, and their relative positions. The bold line with inverted arrows encompasses the region that is transcribed to prepare sense and antisense *CNNM3* intron 2 RNAs examined for in vitro RNA editing assay. (B) A → I editing of *Alu* repeats. The editing frequency (%) at individual sites identified within the entire *Alu* Sg (262 bp; *upper panel*) and *Alu* Jb sequences (309 bp; *lower panel*) in sense (*upper side*) and antisense (*lower side*) RNA is summarized. (C) The secondary structure between two inverted *Alu* repeats in sense (*upper*) and antisense (*lower*) RNA is calculated by MFOLD. Highly edited sites (>30%) are indicated by vertical solid arrows. The blue arrow marked*: the site detected in vivo but not edited in vitro by recombinant ADAR proteins (see Fig. 3).