Gene 634 (2017) 22–28

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Activity analysis of LTR12C as an effective regulatory element of the *RAE1* gene

Yi-Deun Jung[a,1], Hee-Eun Lee[b,c,1], Ara Jo[b,c], Imai Hiroo[d], Hee-Jae Cha[e,2], Heui-Soo Kim[b,c,*,2]

[a] Biotherapeutics Translational Research Center, Korea Research Institute of Bioscience & Biotechnology, Daejeon 34141, Republic of Korea
[b] Department of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Republic of Korea
[c] Institute of Systems Biology, Pusan National University, Busan 46241, Republic of Korea
[d] Molecular Biology Section, Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, Inuyama, Aichi, Japan
[e] Department of Parasitology and Genetics, College of Medicine, Kosin University, Busan, Republic of Korea

ARTICLE INFO

Keywords:
Ribonucleic acid export 1 (*RAE1*)
Long terminal repeat
Regulatory element
Nuclear transcription factor Y

ABSTRACT

Ribonucleic acid export 1 (RAE1) plays an important role in the export of mature mRNAs from the nucleus to the cytoplasm. Long terminal repeats (LTRs) became integrated into the human genome during primate evolution. One such repeat element, LTR12C, lies within a predicted regulatory region located upstream of the *RAE1* gene. We examined the transcriptional activity of LTR12C by using the luciferase assay, and showed that the tandem repeat region (TRR) located within LTR12C was required for its regulatory function. A bioinformatics analysis revealed that the LTR12C element had multiple transcription factor binding sites specific for nuclear transcription factor Y (NF-Y), and the promoter activity of LTR12C was significantly decreased after NF-Y knockdown. Additionally, we discovered novel data indicating that LTR12C was initially inserted into the gorilla genome. Taken together, our results reveal that the TRR of LTR12C has powerful regulatory activity due to its NF-Y binding sites, and the integration of the LTR12C element into the primate genome during evolution may have affected *RAE1* transcription.

1. Introduction

The ribonucleic acid export 1 (*RAE1*) gene encodes a nuclear mRNA exporting factor, which is conserved in terms of structure and function (Bharathi et al., 1997). Export of the mature mRNA from the nucleus to the cytoplasm is an important step in regulating gene expression (Brown et al., 1995). Therefore, *RAE1* may influence the translation of other genes by controlling the mRNA export machinery. *RAE1* expression is also linked to several cancers. Lung cancer patients with higher *RAE1* transcript levels had an increased survival rate (Lu et al., 2006), whereas transcriptional aberration of *RAE1* was associated with breast cancer (Chin et al., 2006). For these reasons, it is important to explore the gene regulatory mechanism controlling *RAE1* expression.

Long terminal repeat (LTR) retroposons comprise approximately 8% of the human genome. These sequences are traces of exogenous retroviruses that had infected ancient germ cells and successfully integrated into the host genome (International Human Genome Sequencing Consortium, 2001). Although they remain stable at present, LTR retroposons drove genetic diversity and affected gene regulation during evolution, where they had both beneficial and harmful effects (Kazazian, 2004; Lee et al., 2012). This is because LTRs within the structural components of endogenous retroviruses have multiple regulatory elements, including promoter, enhancer, silencer, and insulator regions (Ha et al., 2007; Gogvadze and Buzdin, 2009). LTRs can act as alternative regulatory elements that influence gene expression with sequence-specific characteristics like transcription factor binding sites (TFBSs) (Gogvadze and Buzdin, 2009; Jung et al., 2013b). Furthermore, TFBSs in LTRs lead to species-specific regulation of gene expression (Gogvadze and Buzdin, 2009). For example, LTR integration into the cytochrome P450, family 19 (*CYP19*) gene causes placental-specific expression in primates (van de Lagemaat et al., 2003), and an ancient LTR represents the only known promoter for the liver-specific bile acid-CoA:amino acid *N*-acyltransferase (*BAAT*) gene in humans, but not in mice (Carlton et al., 2003).

Abbreviations: bp, base pairs; LTR, long terminal repeat; NF-Y, nuclear transcription factor Y; PCR, polymerase chain reaction; *RAE1*, ribonucleic acid export 1; siRNA, small interfering RNA; TFBS, transcription factor-binding site; TRR, tandem repeat region; TSS, transcription start site

* Corresponding author at: Department of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 46241, Republic of Korea.
E-mail address: khs307@pusan.ac.kr (H.-S. Kim).

[1] These authors contributed equally to this work.
[2] These correspondences authors contributed equally.

http://dx.doi.org/10.1016/j.gene.2017.08.037
Received 21 July 2017; Accepted 24 August 2017
Available online 01 September 2017

In this study, the role of one LTR subtype; LTR12C element located upstream of the *RAE1* gene was investigated and to determine its role as a *cis*-regulatory factor. We determined the regulatory activity in vitro, using three types of LTR12C constructs, and we examined TFBSs within the active regulatory region of the element. We also estimated the time of LTR12C element integration during primate evolution, based on a DNA sequence comparison of the human, chimpanzee, gorilla, orangutan, and gibbon genomes.

2. Materials and methods

2.1. PCR amplification, cloning, and sequencing

Three LTR12C constructs were generated from human genomic DNA to determine the regulatory region important for *RAE1* gene transcription. All three constructs were amplified with the same reverse primer (5′-CCG GAA TTG GTG GGT TCT T-3′), whereas the forward primer and annealing temperature for each were different; the original construct primer (5′-CAC ATT GAG AGA TGA CAG CGT-3′) was annealed at 55 °C, the deletion construct-1 primer (5′-CTG AGT CTG GTG GGG AGG-3′) at 57 °C, and the deletion construct-2 primer (5′-AGA CCA CTG GGC TCT ACC AA-3′) at 53 °C. The constructs were inserted into the pGL-4.11 vector (Promega, Fitchburg, WI, USA), and the cloned sequences were confirmed with the pGL-4.11 sequencing primer (5′-CTA GCA AAA TAG GCT GTC CCC AG-3′) (Macrogen, Seoul, Korea).

To determine the integration time of the LTR12C element, we performed nested PCR amplification using genomic DNA from different primates. The first PCR was performed for 30 cycles at 94 °C (40 s), 50 °C (40 s), and 72 °C (90 s), using the following primers: forward (5′-GGA AAG ATA GTA AGA GGA GTT CCA-3′) and reverse (5′-TCC AGA CCA GCC CTA AAA GAT A-3′). The second PCR was performed for 30 cycles at 94 °C (40 s), 57 °C (40 s), and 72 °C (60 s), using the following primers: forward (5′-CAC ATT GAG AGA TGA CAG CGT-3′) and reverse (5′-CCG GAA TTG GTG GGT TCT T-3′). The PCR products were cloned into the pGEM-T Easy vector (Promega). Plasmid DNA was isolated from the putative clones using the Hybrid-Q Plasmid Mini Kit (GeneAll, Seoul, Korea) and sequenced using T7 primers (Macrogen).

2.2. Cell culture and luciferase reporter assays

Seven different human cell lines (viz., HEK293, U251, A549, HCT116, HepG2, AGS, and Tera-1; all from the American Type Culture Collection, Manassas, VA, USA) were cultured in RPMI 1640 medium or Dulbecco's modified Eagle's medium supplemented with 10% (v/v) heat-inactivated fetal bovine serum and 1% (v/v) antibiotic-antimycotic solution (Gibco, Grand Island, NY, USA) at 37 °C in a 5% (v/v) CO_2 incubator. The cells were seeded in 24-well plates at a concentration of 3×10^4 cells per well and grown to 70% confluence. To check the promoter activity of the LTR12C constructs, the cells were transfected with Opti-MEM (Gibco) containing 100 ng/μL of the both empty and cloned pGL-4.11 and *Renilla* luciferase vectors (pGL-4.73), using the Lipofectamine 2000 transfection reagent (Invitrogen, Carlsbad, CA, USA).

For small interfering RNA (siRNA) treatment, cells were transfected with Opti-MEM (Gibco) containing 200 ng/μL of the pGL-4.11 and *Renilla* luciferase vectors (pGL-4.73) and 40 nM of siRNA constructs targeting nuclear transcription factor Y (NF-Y; Bioneer, Dajeon, Korea), using the Lipofectamine 2000 transfection reagent. AccuTarget siRNA (Bioneer) was used as a negative control. After 24 h, the cells were lysed, and the luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega).

2.3. Bioinformatics approach

Reference nucleotide sequences were obtained from the University of California Santa Cruz (UCSC) Genome Browser (http://www.genome.ucsc.edu). The promoter region was predicted by applying the BioInformatics & Molecular Analysis Section Promoter Scan, which uses PROSCAN v1.7. The prediction was based on scoring homologies with putative eukaryote polymerase II sequences (http://www-bimas.cit.nih.gov/molbio/proscan). The TFBSs of the LTR12C constructs were analyzed using MATCH in TRANSFAC v8.0. Threshold values > 0.95 were used for the prediction. The LTR12C sequencing results were aligned using the BioEdit program (Hall, 1999), and the sequences were analyzed using RepeatMasker (Smit et al., 1996).

2.4. Genomic DNA isolation

Genomic DNA was isolated from human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orangutan (*Pongo pygmaeus*), and gibbon (*Hylobates agilis*) blood samples using a DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol. The concentration of the extracted genomic DNA was increased by amplification using an Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare, Little Chalfont, Buckinghamshire, UK). The primate samples were a gift from professor Takenaka.

3. Results

3.1. Regulatory element prediction associated with RAE1

The human *RAE1* gene has two isoforms, NM 003610.3 and NM 001015885.1. Prediction of the *RAE1* regulatory region was performed using human reference sequences (UCSC Genome Browser) ranging from the upstream region of the *RAE1* transcription start site (TSS) to the fourth intron of the gene (Fig. 1). Two significant signals,

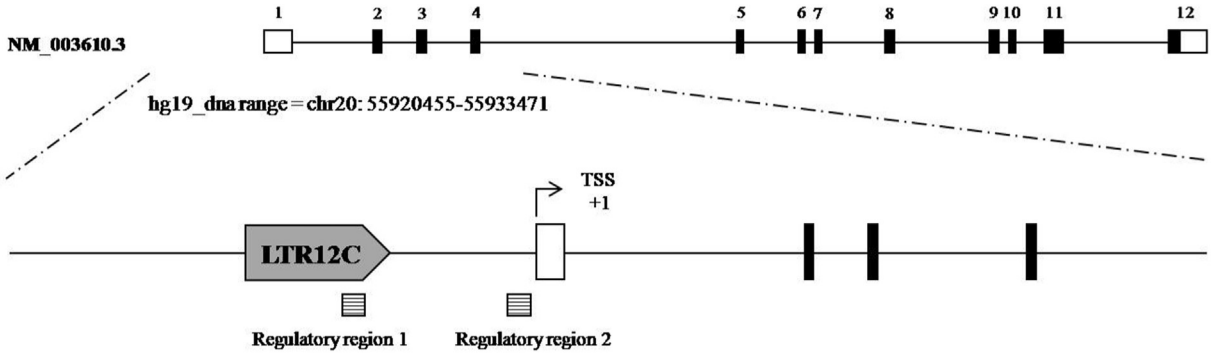

Fig. 1. Illustration of the human *RAE1* gene structure and its predicted regulatory regions. The gene structure is based on *RAE1* NM 003610.3. White and black block boxes indicate the untranslated region (UTR) and coding sequences (CDS), respectively. Numerals above the gene structure stand for the exon number. The regulatory region was predicted by a 13,017-bp sequence. Striped boxes indicate predicted regulatory regions, and the gray box labels the LTR12C element. Two significant promoter signals were detected on the forward strand; the promoter score of the first region was 61.94 and that of the second region was 58.95 from the left (cutoff = 53.0).