Gene 681 (2019) 36–44

Contents lists available at ScienceDirect

Gene

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

Research paper

A natural "GA" insertion mutation in the sequence encoding the 3′UTR of CXCL12/SDF-1α: Identification, characterization, and functional impact on mRNA splicing

Xiuying Zhao[a,*], Dong Zhu[a], Haojie Zhang[b], Hongyan Sui[b], Vincent Poon[b], Shibo Jiang[c], Bojian Zheng[b]

[a] Beijing Tsinghua Changgung Hospital, School of Clinic Medicine, Tsinghua University, Beijing, China
[b] The HIV/AIDS Research Laboratories, Department of Microbiology, the University of Hong Kong, Hong Kong
[c] Lindsley F. Kimball Research Institute, the New York Blood Center, New York, USA

ARTICLE INFO

Keywords:
Mutation
Chemokine CXCL12
G-quadruplexes
3′ untranslated region (3′UTR)
Transcription
Minor transcript isoform (MTI)

ABSTRACT

The CXCL12 gene produces a series of transcript variants through alternative splicing at the 3′ end of its pre-mRNA. This study explores the biological activities of these alternative transcripts and the mechanisms involved in the regulation of CXCL12 transcription and RNA splicing. We identified a "GA" insertion mutation in the region of CXCL12α DNA encoding the conserved 3′UTR. This variant transcript was named CXCL12-3′GA+. The mutation occurred at a frequency of 13.2% in healthy Chinese individuals. However, its frequency in healthy Caucasians was 22.6%, significantly higher than what was observed in the Chinese. Genomic analysis indicated that the GA+ mutation likely encodes a G-quadruplex structure in close proximity to a cluster of important AU-rich elements (AREs) that are well-established regulators of mRNA stability at the 3′UTR. Experiments using molecular constructs encoding the 3′UTR of CXCL12 revealed that the GA+ allele can significantly increase gene expression compared to the WT allele. Further studies uncovered that the WT allele was associated with the production of a 225-bp minor transcript isoform (MTI) through alternative splicing resulting in the deletion of exon 2. ARMS-PCR using samples collected from cultured PBMCs of WT/GA+ genotype carriers indicated that the GA+ allele was preferentially transcribed compared to the WT allele. In summary, the study demonstrates that a GA insertion in the region encoding the 3′UTR of CXCL12α may affect gene expression through alternative mRNA splicing. This finding provides a basis for understanding how multiple elements in the sequence encoding the 3′UTR of the CXCL12 gene regulates its transcription and may lead to insights about diseases involving abnormal CXCL12α expression.

1. Introduction

The C-X-C motif chemokine ligand (CXCL12), also known as stromal cell derived factor (SDF)-1, and pre-B-cell growth stimulating factor precursor (PBSF), was originally isolated from a murine bone marrow stromal cell line (Nagasawa et al., 1996; Aiuti et al., 1997). This chemokine, along with its receptors CXCR4 and CXCR7, is believed to play a critical role in HIV-1 infection, the metastatic spread of tumor cells, and the development and maintenance of the nervous, hematopoietic, and cardiovascular systems (Burger and Kipps, 2006; Petit et al., 2007; Sierro et al., 2007; Wang et al., 2008). A previous study showed that the CXCL12 gene is located on human chromosome 10q and spans over 88 kb. In contrast, other CXC chemokine genes are clustered in human chromosome 4q. The amino acid (aa) sequence of the CXCL12 protein

Abbreviations: aa, amino acid(s); AREs, AU-rich elements; ARMS, amplification refractory mutation system; bp, base pair(s); cDNA, DNA complementary to RNA; cpm, counts per minute; CXCL, C-X-C motif chemokine ligand; CXCR, C-X-C chemokine receptor; d, deoxyribo; DMSO, dimethyl sulfoxide; dNTP, deoxyribonucleoside triphosphate; HIV, human immunodeficiency virus; MTI, minor transcript isoform; MU, mutant; ORF, open reading frame; PBL, peripheral blood lymphocytes; PBMC, peripheral blood mononuclear cells; PBSF, pre-B-cell growth stimulating factor precursor; PBS, phosphate buffer saline; PCR, polymerase chain reaction; PHA, phytohemagglutinin; PTC, premature termination codon; QGRS, G-rich quadruplex sequences; RIPA, radio immune-precipitation assay; RT, reverse transcription; SDF, stromal cell derived factor; Tat, trans-activator of transcription; UTR, untranslated region(s); WT/wt, wild type; Δ, deletion

* Corresponding author at: Department of Clinical Laboratories, Beijing Tsinghua Changgung Hospital, School of Clinic Medicine, Tsinghua University, Beijing, China.
E-mail address: zhaoxiuying2001@163.com (X. Zhao).

https://doi.org/10.1016/j.gene.2018.09.045
Received 23 April 2018; Received in revised form 19 September 2018; Accepted 24 September 2018

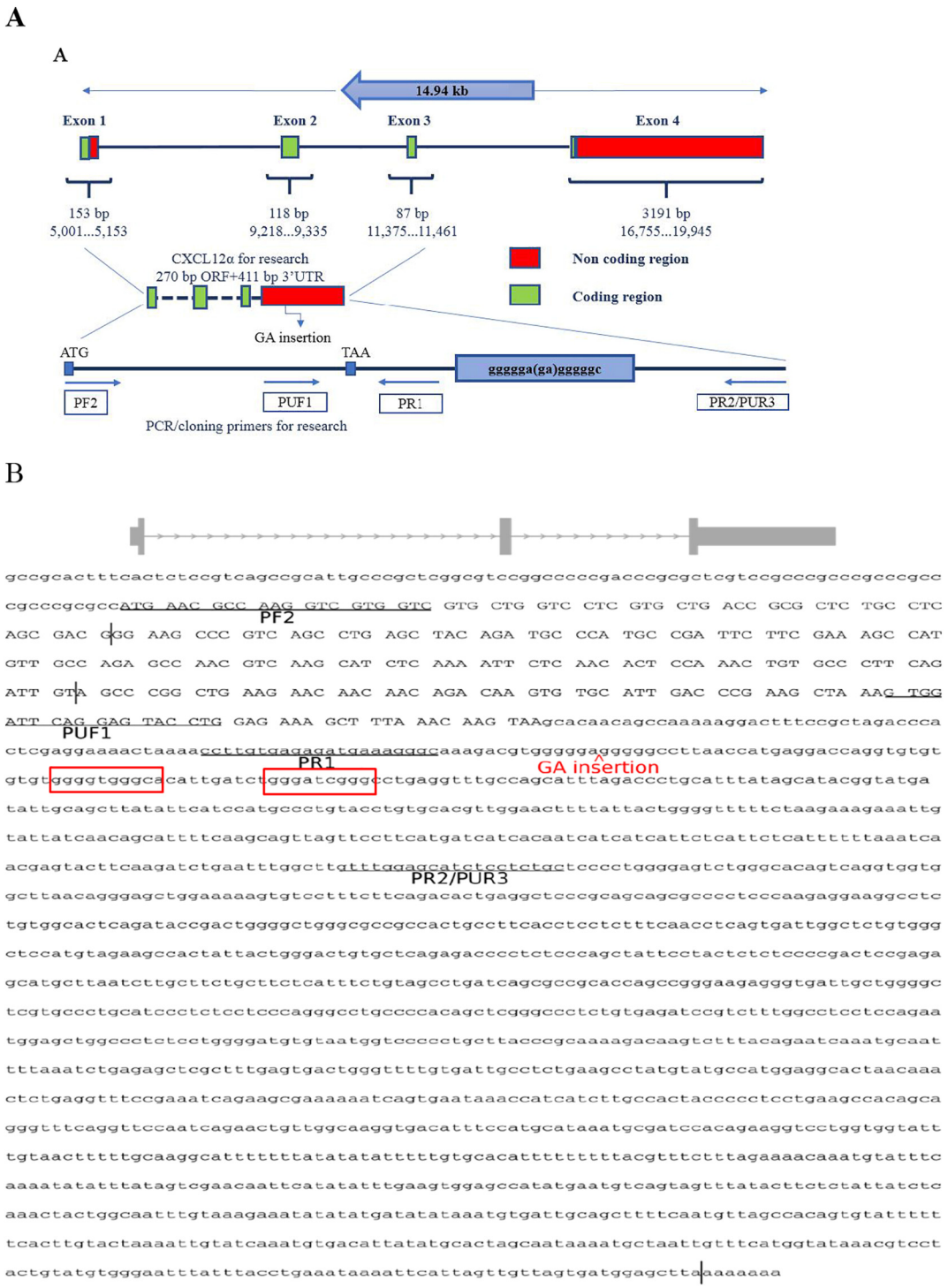

Fig. 1. A natural "GA" insertion mutation was identified in the sequence encoding the 3′UTR of the CXCL12α transcript. (A) Cartoon structure of CXCL12 gene with the location of "GA" insertion. (B) Sequence of CXCL12α transcript, with cloning primers denoted by underline and exon/intron borders indicated by vertical lines. Boxed regions corresponding to two putative G-rich quadruplex sequences by QGRS.

has > 90% homology between mice and humans. The unique chromosomal localization and the evolutionary conservation of the CXCL12 gene suggest that it may have functions distinct from those of other members of the CXC chemokine family (Shirozu et al., 1995).

To date, six functional transcript variants of the CXCL12 gene have been identified. The cDNA of CXCL12α encodes a protein that is 89 aa long. The remaining variants, CXCL12β, CXCL12γ, CXCL12ε, CXCL12δ, and CXCL12θ, are generated by alternative splicing, with different exons spliced with the same first three exons (Yu et al., 2006; Altenburg et al., 2007; Laguri et al., 2007). This implies that the first 88 residues (encoded by the first 3 exons) of these splicing variants of CXCL12 are identical and that they are likely essential for binding to CXCR4 (Altenburg et al., 2007; Laguri et al., 2007; Sierro et al., 2007). These CXCL12 splicing variants have been detected in a variety of tissues, including the liver, pancreas, spleen, and heart. The specific mechanisms governing their biological activities, such as in hematopoiesis, chemotaxis, and HIV-1 blocking, remain under investigation (Schols et al., 1997; Wang et al., 2000; Altenburg et al., 2007; Petit et al.,