Available online at www.sciencedirect.com

SCIENCE DIRECT

Genomics 85 (2005) 221–230

GENOMICS

www.elsevier.com/locate/ygeno

Ancient repeated DNA elements and the regulation of the human frataxin promoter☆

Eriko Greene, Ali Entezam, Daman Kumari, Karen Usdin*

Section on Genomic Structure and Function, Laboratory of Molecular and Cellular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Building 8, Room 202, 8 Center Drive MSC 0830, Bethesda, MD 20892-0830, USA

Received 30 June 2004; accepted 23 October 2004

Abstract

Friedreich ataxia results from frataxin insufficiency caused by repeat expansion in intron 1 of the frataxin gene. Since the coding sequence is unchanged, the potential exists to ameliorate symptoms by increasing frataxin promoter activity. We therefore defined the minimal frataxin promoter in humans. Despite the fact that frataxin is an essential gene, its promoter is not well conserved in mammals, in part because it has been the frequent target of retroelement insertions. Most of the activity of the human frataxin promoter can be attributed to these retroelements, illustrating how these elements, considered parasitic by some, have been co-opted to drive critical genes. Individuals with the milder French Acadian form and those with the classic form of the disease have no biologically relevant sequence differences in the promoter or 3′UTR, suggesting that some other region of the gene, perhaps the repeat itself, is responsible for the difference in disease severity. Published by Elsevier Inc.

Keywords: Frataxin promoter; *Alu*; MIR; L2 elements

Friedreich ataxia (FRDA), a disease found in populations of European, North African, Middle Eastern, and Indian origin, is the most common recessively inherited ataxia [22]. It is a relentlessly progressive gait and limb ataxia, which is accompanied by dysarthria, loss of tendon reflexes, and skeletal abnormalities. Onset is often early with patients on average losing the ability to walk 15 years after the first symptoms appear. Diabetes is a frequent complication that is thought to result at least in part from free radical-induced apoptosis of pancreatic β cells [27]. Early mortality is not uncommon, most frequently due to hypertrophic cardiomyopathy [6,34].

FRDA results from loss-of-function mutations in the frataxin gene on chromosome 9p13 [5]. This gene encodes an essential protein, frataxin, which is thought to play a role in mitochondrial iron metabolism and the biogenesis of iron–sulfur clusters [22]. More than 98% of FRDA patients are homozygous for an expansion of a GAA · TTC repeat tract in the first intron of this gene. In general there is a correlation between the number of repeats in the smaller allele and the disease severity and an inverse correlation with age of onset [22].

Expanded repeats are associated with a deficit in full-length frataxin mRNA. This deficit has been suggested to result from a problem with transcription elongation [4,9,10,28], although an effect on transcription initiation may also be possible [32]. One recent approach to ameliorating FRDA symptoms has been to try and increase frataxin expression [8,31,36]. However, almost nothing is known about the factors normally involved in the regulation of this gene. In an effort to understand better factors important for frataxin expression we have analyzed the region 5′ of the frataxin open reading frame in a number of different primates and rodents and identified the regions of the human sequence that are important for optimal promoter activity. Our data provide a molecular explanation for the

☆ Sequence data from this article have been deposited with the GenBank Data Library under Accession Nos. AY599001–AY599014.

* Corresponding author. Fax: +1 301 402 0053.
E-mail address: ku@helix.nih.gov (K. Usdin).

0888-7543/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.ygeno.2004.10.013

effects of some drugs already shown to increase frataxin expression and also suggest other ways to increase frataxin gene expression. To understand better the molecular basis of the milder FRDA symptoms seen in individuals of French Acadian descent [2,3,14,20,26] we have also compared the regulatory regions of these individuals with those suffering from the classic form of the disease.

Results

Defining the minimal human frataxin promoter

To understand better the regulation of the frataxin gene, five constructs, FRDA4356, FRDA3433, FRDA2541, FRDA1862, and FRDA1255, containing 4356, 3433, 2541, 1862, and 1255 bp, respectively, upstream of the human frataxin open reading frame were generated by cloning PCR-amplified fragments into pGL-Basic. The ability of these constructs to drive luciferase activity was measured after transfection into C2C12 mouse skeletal muscle myoblasts or myotubes. These cells were chosen since frataxin expression is relatively high in skeletal muscle, the pattern of frataxin expression is similar in mouse and human [12], and transgenic mice containing YACs and BACs carrying the human frataxin locus show levels of expression similar to those of endogenous murine frataxin, particularly in skeletal muscle [25,30].

Promoter activity was measured in cells grown in either 10% fetal calf serum, conditions in which the cells remain undifferentiated, or 2% serum, conditions that promote skeletal myotube formation. All five constructs resulted in an ~20-fold increase in luciferase activity above pGL-Basic in 10% serum (Fig. 1). The same result was seen when cells were grown in 2% serum. An increase in frataxin protein has been reported when mouse embryonic carcinoma cells P19 are induced to differentiate into cardiomyocytes [29]. It is unclear at this time whether this effect results from differentiation-related changes at the transcriptional or posttranscriptional level. In any event it does not seem to be a gene-specific phenomenon since a similar increase is seen in a number of other mitochondrial proteins and has been attributed to an increase in the number of mitochondria [29].

Sequence conservation in the minimal frataxin promoter

Since the presence of sequences farther 5′ of base 1255 has little effect on promoter activity over that provided by the FRDA1255 construct, we focused our subsequent analysis on the region of the frataxin 5′ end contained within this construct. Sequence similarities between humans and rodents are limited to two small regions of patchy homology downstream of the transcription start site (TSS). The first region overlaps a region with sequence similarity to the ancient L2 family of retrotransposable elements, whose activity predated the divergence of rodents and primates (Fig. 2). The second region is located immediately upstream of the translation start site. A recent survey of 46 different human promoters showed that ~94% of them had significant sequence similarity with their rodent counterparts [38]. This puts the frataxin promoter into that small category of promoters in which significant amounts of the promoter do not seem to be evolving under strong selective pressure. However, while these promoters lack strict sequence similarity in their 5′ ends, both the rodent and the human promoters contain related SINE elements, *Alu* in the case of human and B1 in the case of mouse (data not shown). Both of these elements originate from the same region of the 7SL RNA, a component of the signal recognition particle that is involved in translation of secreted proteins in eukaryotes [37]. Thus these elements may share regulatory elements originally derived from this RNA. The B1 element in rodents is embedded in a small piece of an old rodent-specific L1 retrotransposable element, Lx (data not shown).

While very little human–rodent homology is seen upstream of the start of transcription, the Catarrhine primates *Homo sapiens* (humans) and *Pan troglodytes* (chimpanzee) and the Platyrrhine primate *Ateles geoffroyi* (spider monkey) share significant sequence similarity (Fig. 3). Both the human and the chimpanzee promoters contain *Alu* retroelements belonging to the AluJb and AluY subfamilies. The activity of AluJb predates the divergence of the Catarrhine lineage from that of the Platyrrhine primates, while AluY is believed to have been active in the Catarrhine lineage after this time. The AluY element is missing in the spider monkey, consistent with fact that the AluY family arose in the Catarrhine lineage. The primate promoters also contain a member of another very ancient repeated DNA family, known as MIR (mammalian-wide interspersed repeats), which was active in the common ancestor of all modern mammals. Despite the early origin of the MIR family, no such element is found in either the rat or the mouse frataxin promoters.

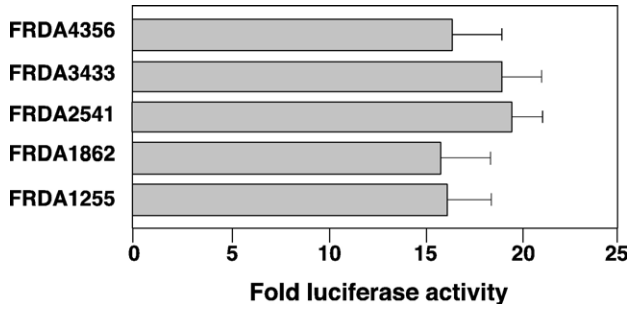

Fig. 1. Effects of sequences at the 5′end of the human frataxin gene on the activity of a reporter gene in mouse myoblasts. Promoter constructs were transfected into C2C12 cells as described under Materials and methods. The luciferase activity was measured, normalized to the activity of a cotransfected plasmid expressing the *Renilla* luciferase, and then expressed relative to the activity of the empty pGL3-Basic vector.