A novel cryptic exon identified in the 3' region ofintron 2 of the human dystrophin gene

Tran, Van Khanh

(Degree)
博士（医学）

(Date of Degree)
2006-03-25

(Resource Type)
doctoral thesis

(Report Number)
甲3577

(URL)
https://hdl.handle.net/20.500.14094/D1003577

【 98 】

氏　　名・（本　籍）	Tran Van Khanh　（　ベトナム　）
博士の専攻分野の名称	博士（医学）
学　位　記　番　号	博い第1720号
学位授与の　要　件	学位規則第５条第１項該当
学位授与の　日　付	平成18年3月25日

【 学位論文題目 】

A novel cryptic exon identified in the 3′ region of intron 2 of the human dystrophin gene
（ヒト・ジストロフィン遺伝子イントロン２ には未知のクリプティックエクソンが存在する）

審　査　委　員

主 査　教　授　西尾　久英
教　授　千原　和夫
教　授　熊谷　俊一

Introduction

The human dystrophin gene, which is defective in patients with Duchenne or Becker muscular dystrophy (DMD/BMD), spans approximately 3,000 kb of the X-chromosome and encodes a 14-kb transcript consisting of 79 exons. Genomic structural analysis disclosed at least eight alternative promoters over the entire dystrophin gene, producing tissue-specific dystrophin isoforms. Consequently, more than 99% of the gene sequence is comprised of introns and has been considered functionless. Intron 2, the second largest intron with 170-kb-long, has been shown to contain a cryptic exon, exon 2a, in its 5' region but the physiological role of exon 2a is still unknown. Recently, a part of the 5' region of intron 2 in its 5'region was shown to be incorporated in dystrophin mRNA due to an activation mutation in the splice donor site of an embedded weak exon (exon p2a). Splicing is the process that removes introns from pre-mRNA thereby producing mature mRNA consisting of only exons. The presence of well-defined cis elements, namely, the 5' and 3'splice sites and the branch point, is necessary but not sufficient to define intron–exon boundaries in pre-mRNA. Unconventional splicing defects often occur at exons with weak homology to canonical splicing sequences, leading to dystrophinopathies. Here, we identify an unknown sequence inserted into a dystrophin transcript in a case with exon 2 duplication of the dystrophin gene, and we find that the sequence is a novel cryptic exon (exon 2b) located in the 3' region of intron 2 of the dystrophin gene. Exon 2b is incorporated into mRNA in a promoter- or tissue-specific manner. This provides a clue to a novel cause of dystrophinopathy.

Patient and methods

Case: A 5-year-old Japanese boy was referred to the Kobe University Hospital for the genetic diagnosis of DMD. At age 4, he was shown to have an extremely high level of serum CK (13,750IU/l, normal: 56–248IU/l) and was clinically diagnosed as DMD.

Analysis of genomic DNA: Genomic DNAs were isolated from lymphocytes of DMD patients and a normal male individual. To examine the entire dystrophin gene, Southern blot analysis of the patient's genomic DNA was performed using *Hae*III-digested cDNA fragments as probe. A genomic region encompassing the 98-bp inserted sequence (exon 2b) was amplified using primers derived from the flanking sequences. The copy number of exons (exon 1; 1a; 2, p2a, 2a; 2b) was assessed by semiquantitative, multiplex PCR (Agilent 2001 Bioanalyzer with DNA 1000 Lab Chips, Agilent Technologies, Palo Alto, CA, USA)

Analysis of dystrophin transcripts: Total RNA was isolated from peripheral lymphocytes. A fragment encompassing exon 1-5 of dystrophin mRNA was analyzed by Reverse Transcription (RT) seminested PCR. To examine the promoter specificity of exon 2b incorporation, fragments stretching from promoter-specific exon 1(L, M, C, P) to exon 2b were amplified from lymphocyte cDNA. To examine the efficiency of exon 2b activation in different tissues, fragments spanning from exon 1 to exon 2b and from exon 1 to exon 5 were amplified from cDNA prepared from total RNA from 20 human tissues. DNA sequencing: the purified DNA was subcloned into vector pT7 and inserted DNA was sequenced

Results

PCR quantification of the region encompassing exon 2 showed the duplication of exon 2 in the genome of the index case. The duplication exon 2 was also detected in the lymphocyte dystrophin mRNA by RT nested PCR encompassing exon 1-5. Remarkably, one barely visible, weak band and as well as two major, equally dense bands were obtained. Each of the bands was sequenced after subcloning. Sequencing of the smallest fragment revealed a sequence of tandem exon 2 sequences between exons 1 and 3. Sequencing of the middle-sized band revealed an insertion of exon 1a between exon 1 and 2 in addition to duplication of exon 2. The sequence of the largest fragment revealed the same exons as in the middle-sized one, but, remarkably, an unidentified 98-bp sequence was found to be inserted precisely between tandem exon 2 and exon 3. Since all three dystrophin mRNAs contained tandem exon 2 sequences, we concluded that the index case had a duplication of exon 2. This mutation created a premature stop codon at 15th codon of the duplicated exon 2 sequence and was determined as a cause of DMD. A BLAST search of the 98-bp sequence revealed an identical sequence in the 3'region of intron 2 (bases 10151–10054 of GenBank AL121880). The 98-bp sequence was located 82 kb downstream from exon 2a and 29 kb upstream of exon 3. Remarkably, the 98-bp-inserted sequence exhibited all of the characteristics typical of a genomic exon and was inserted between authentic dystrophin exons, we refer to it as the novel exon 2b.

The genomic sequence encompassing exon 2b failed to show any mutation, indicating that no genomic mutation contributed to the activation of exon2b. Although tandem exon 2 sequences were identified in the case's dystrophin mRNA, only one

exon 2b was identified. To explore this, the copy number of exons (exon 1; 1a; 2, p2a, 2a; 2b) was assessed by semiquantitative, multiplex PCR and the results showed the presence of single copy number of exon 2b in his genome, indicating that only single exon 2b downstream of the duplicated exon 2 was activated. Since we failed to detect the incorporation of exon 2b in another case with exon 2 duplication, we concluded that exon 2b was specific to the index case. The RT- PCR amplification encompassing promoter-specific exon 1 to exon 2b of mRNA either from the case's lymphocytes or from 20 different human tissues indicated that exon 2b incorporation is dependent upon the muscle-specific promoter or tissue-specific factors.

Discussion

The functional diversity of the dystrophin gene is now becoming apparent, but the role of its unusually huge introns is still unknown. The identification of exons within the introns may shed light on the diverse functions of dystrophin. In this report, the novel cryptic exon 2b was identified in the 3'region of intron 2; exon 2b maintains all of the characteristic sequences necessary for exon recognition and is incorporated into dystrophin mRNA. This is the second cryptic exon discovered within the 170-kb long intron 2, but nearly 170 kb still lack any described function. Although exon 2b has a structure similar to the real exon, exon 2b had not been previously described. This may be due to its low Shapiro splicing probability score or its tissue-specific incorporation. The concomitant incorporation of exon 1a and exon 2b suggests a common regulatory system for the two cryptic exons, but exon 1a incorporation is not always accompanied by exon 2b incorporation, indicating that exon 2b incorporation is regulated by a different mechanism from that of exon 1a. Among four alternative promoters at the 5'end of the dystrophin gene, transcripts containing exon 2b were initiated only at the muscle-specific promoter, indicating that exon 2b incorporation is promoter specific. However, exon 2b incorporation was detected in mRNAs from only five of the 20 tissues in which the muscle-specific promoter-driven transcript could be detected. These findings indicate that exon 2b incorporation is under the control of the muscle-specific promoter, but that this not sufficient for exon 2b incorporation, which requires another tissue-specific factor. This complex pattern of regulation of exon 2b incorporation strongly suggests a physiological role for this cryptic exon. Cryptic exons have been shown to be activated by intron mutations that either create or strengthen splice sites or create a branch site .In addition, an intracryptic exon deletion has been shown to cause erroneous splicing. These observations suggest that cryptic exons are targets for human genetic diseases. Since the typical sequence characteristics of exons have been maintained in exon 2b, we suggest that a sequence acting as a splicing silencer inhibits the incorporation of exon 2b into mRNA. Future experiments may reveal intronic mutations that either disrupt a splicing silencer or activate a splicing enhancer to cause exon 2b incorporation.

論文審査の結果の要旨			
受付番号	甲 第 1722 号	氏 名	Tran Van Khanh
論文題目 Title of Dissertation	A novel cryptic exon identified in the 3' region of intron 2 of the human dystrophin gene ヒト・ジストロフィン遺伝子イントロン２には未知のクリプティックエクソンが存在する		
審査委員 Examiner	主 査 Chief Examiner 西尾久英 副 査 Vice-examiner 千原和夫 副 査 Vice-examiner 熊谷俊一		
審査終了日	平成 18 年 2 月 16 日		

（要旨は１，０００字～２，０００字程度）

Introduction: The human dystrophin gene, which is defective in patients with Duchenne or Becker muscular dystrophy (DMD/BMD), spans approximately 3,000 kb of the X-chromosome and encodes a 14-kb transcript consisting of 79 exons. Genomic structural analysis disclosed at least eight alternative promoters over the entire dystrophin gene, producing tissue-specific dystrophin isoforms. Consequently, more than 99% of the gene sequence is comprised of introns and has been considered functionless. Intron 2, the second largest intron with 170-kb-long, has been shown to contain a cryptic exon, exon 2a, in its 5' region but the physiological role of exon 2a is still unknown. Recently, a part of the 5' region of intron 2 in its 5'region was shown to be incorporated in dystrophin mRNA due to an activation mutation in the splice donor site of an embedded weak exon (exon p2a). Splicing is the process that removes introns from pre-mRNA thereby producing mature mRNA consisting of only exons. The presence of well-defined cis elements, namely, the 5' and 3'splice sites and the branch point, is necessary but not sufficient to define intron–exon boundaries in pre-mRNA. Unconventional splicing defects often occur at exons with weak homology to canonical splicing sequences, leading to dystrophinopathies. In this study, Khanh et al. identified an unknown sequence inserted into a dystrophin transcript in a case with exon 2 duplication of the dystrophin gene, and they found that the sequence is a novel cryptic exon (exon 2b) located in the 3' region of intron 2 of the dystrophin gene. Exon 2b is incorporated into mRNA in a promoter- or tissue-specific manner.

Patient and methods: A 5-year-old Japanese boy was referred to the Kobe University Hospital for the genetic diagnosis of DMD. At age 4, he was shown to have an extremely high level of serum CK (13,750IU/l, normal: 56–248IU/l) and was clinically diagnosed as DMD.

Genomic DNAs were isolated from lymphocytes of DMD patients and a normal male individual. To examine the entire dystrophin gene, Southern blot analysis of the patient's genomic DNA was performed

using *Hae*III-digested cDNA fragments as probe. A genomic region encompassing the 98-bp inserted sequence (exon 2b) was amplified using primers derived from the flanking sequences. The copy number of exons (exon 1; 1a; 2, p2a, 2a; 2b) was assessed by semiquantitative, multiplex PCR (Agilent 2001 Bioanalyzer with DNA 1000 Lab Chips, Agilent Technologies, Palo Alto, CA, USA)

Total RNA was isolated from peripheral lymphocytes. A fragment encompassing exon 1-5 of dystrophin mRNA was analyzed by Reverse Transcription (RT) seminested PCR. To examine the promoter specificity of exon 2b incorporation, fragments stretching from promoter-specific exon 1(L, M, C, P) to exon 2b were amplified from lymphocyte cDNA. To examine the efficiency of exon 2b activation in different tissues, fragments spanning from exon 1 to exon 2b and from exon 1 to exon 5 were amplified from cDNA prepared from total RNA from 20 human tissues. DNA sequencing: the purified DNA was subcloned into vector pT7 and inserted DNA was sequenced

Results: PCR quantification of the region encompassing exon 2 showed the duplication of exon 2 in the genome of the index case. The duplication exon 2 was also detected in the lymphocyte dystrophin mRNA by RT nested PCR encompassing exon 1-5. Remarkably, one barely visible, weak band and as well as two major, equally dense bands were obtained. Each of the bands was sequenced after subcloning. Sequencing of the smallest fragment revealed a sequence of tandem exon 2 sequences between exons 1 and 3. Sequencing of the middle-sized band revealed an insertion of exon 1a between exon 1 and 2 in addition to duplication of exon 2. The sequence of the largest fragment revealed the same exons as in the middle-sized one, but, remarkably, an unidentified 98-bp sequence was found to be inserted precisely between tandem exon 2 and exon 3. Since all three dystrophin mRNAs contained tandem exon 2 sequences, Khanh et al. concluded that the index case had a duplication of exon 2. This mutation created a premature stop codon at 15th codon of the duplicated exon 2 sequence and was determined as a cause of DMD. A BLAST search of the 98-bp

sequence revealed an identical sequence in the 3'region of intron 2 (bases 10151–10054 of GenBank AL121880). The 98-bp sequence was located 82 kb downstream from exon 2a and 29 kb upstream of exon 3. Remarkably, the 98-bp-inserted sequence exhibited all of the characteristics typical of a genomic exon and was inserted between authentic dystrophin exons, Khanh et al. referred to it as the novel exon 2b.

The genomic sequence encompassing exon 2b failed to show any mutation, indicating that no genomic mutation contributed to the activation of exon2b. Although tandem exon 2 sequences were identified in the case's dystrophin mRNA, only one exon 2b was identified. To explore this, the copy number of exons (exon 1; 1a; 2, p2a, 2a; 2b) was assessed by semiquantitative, multiplex PCR and the results showed the presence of single copy number of exon 2b in his genome, indicating that only single exon 2b downstream of the duplicated exon 2 was activated. Since they failed to detect the incorporation of exon 2b in another case with exon 2 duplication, Khanh et al. concluded that exon 2b was specific to the index case. The RT- PCR amplification encompassing promoter-specific exon 1 to exon 2b of mRNA either from the case's lymphocytes or from 20 different human tissues indicated that exon 2b incorporation is dependent upon the muscle-specific promoter or tissue-specific factors.

Discussion: The functional diversity of the dystrophin gene is now becoming apparent, but the role of its unusually huge introns is still unknown. The identification of exons within the introns may shed light on the diverse functions of dystrophin. In this report, the novel cryptic exon 2b was identified in the 3'region of intron 2; exon 2b maintains all of the characteristic sequences necessary for exon recognition and is incorporated into dystrophin mRNA. This is the second cryptic exon discovered within the 170-kb long intron 2, but nearly 170 kb still lack any described function. Although exon 2b has a structure similar to the real exon, exon 2b had not been previously described. This may be due to its low Shapiro splicing probability score or its tissue-specific

incorporation. The concomitant incorporation of exon 1a and exon 2b suggests a common regulatory system for the two cryptic exons, but exon 1a incorporation is not always accompanied by exon 2b incorporation, indicating that exon 2b incorporation is regulated by a different mechanism from that of exon 1a. Among four alternative promoters at the 5'end of the dystrophin gene, transcripts containing exon 2b were initiated only at the muscle-specific promoter, indicating that exon 2b incorporation is promoter specific. However, exon 2b incorporation was detected in mRNAs from only five of the 20 tissues in which the muscle-specific promoter-driven transcript could be detected. These findings indicate that exon 2b incorporation is under the control of the muscle-specific promoter, but that this not sufficient for exon 2b incorporation, which requires another tissue-specific factor. This complex pattern of regulation of exon 2b incorporation strongly suggests a physiological role for this cryptic exon. Cryptic exons have been shown to be activated by intron mutations that either create or strengthen splice sites or create a branch site .In addition, an intracryptic exon deletion has been shown to cause erroneous splicing. These observations suggest that cryptic exons are targets for human genetic diseases. Since the typical sequence characteristics of exons have been maintained in exon 2b, Khanh et al. suggest that a sequence acting as a splicing silencer inhibits the incorporation of exon 2b into mRNA. Future experiments may reveal intronic mutations that either disrupt a splicing silencer or activate a splicing enhancer to cause exon 2b incorporation.

The candidate, having completed studies on Duchenne muscular atrophy, with a specialty in molecular genetics, and having advanced the field of knowledge in the area of exon recognition during the RNA splicing procedures, is hereby recognized as having qualified for the degree of Ph.D.(Medicine).