ANALYSIS OF PHENOTYPE EXPRESSIONS OF DELETIONS IN THE DYSTROPHIN GENE IN TERMS OF EFFICIENCY OF EXON SKIPPING AS A METHOD FOR TREATMENT OF HEREDITARY DYSTROPHINOPATHIES

Zotova ED1,2, Reshetov DA1, Zhernovkov VE1, Vlodavets DV2, Dimitrieva TV1, Deykin AV1,2

1 Marlin Biotech, Moscow, Russia
2 Molecular Biology Core Facility, Institute of Gene Biology, Russian Academy of Sciences, Moscow, Russia
3 Department of Neurology, Neurosurgery and Medical Genetics, Faculty of Pediatrics, Progov Russian National Research Medical University, Moscow, Russia

Duchenne muscular dystrophy (DMD) is a common genetic disease caused by a mutation of the dystrophin gene. It leads to death in childhood. At the time of writing this paper, patients had access to supportive therapy only. However, DMD treatment methods are actively being developed. Exon skipping is a promising method. Exon skipping involves restoration of the reading frame within a gene by inducing alternative splicing. This leads to synthesis of truncated but still functional dystrophin. The paper assesses the functionality of the truncated forms of dystrophin resulting from correction of nonsense mutations and internal exon indels by exon-skipping technique. The assessment was made based on data on the phenotype of carriers of mutations in the dystrophin gene taken from the Leiden Open Variation Database (LOVD). It was revealed that the same mutation could manifest itself as a variety of phenotypes. This, perhaps, is as a result of the patients having different genetic background. For example, deletion of exon 48, for which there is 97 records in LOVD, resulted in asymptomatic diseases in 2 % of cases, Duchenne muscular dystrophy in 60 %, Becker muscular dystrophy (characterized by milder symptoms than DMD) in 12 % and intermediate phenotype in 26 % of cases. High phenotypic variability of mutations of the dystrophin gene raises the issue of limits of applying exon skipping for treatment of inherited myopathies.

Keywords: muscular dystrophy, Duchenne muscular dystrophy, Becker muscular dystrophy, exon skipping

Acknowledgement: authors thank the Shared Resource Center of the Institute of Gene Biology of Russian Academy of Sciences for the equipment provided for this research.

Correspondence should be addressed: Evgeniya Zotova
ul. Vavilova, d. 34/5, Moscow, Russia; 119334; zotova@ibb.msu.ru
Received: 17.06.2016 Accepted: 24.06.2016

АНАЛИЗ ФЕНОТИПИЧЕСКИХ ПРОЯВЛЕНИЙ ДЕЛЕЦИЙ В ГЕНЕ ДИСТРОФИНА В КОНТЕКСТЕ ЭФФЕКТИВНОСТИ ПРОПУСКА ЭКЗОНОВ КАК МЕТОДА ТЕРАПИИ НАСЛЕДСТВЕННЫХ ДИСТРОФИНОПАТИЙ

Е. Д. Зотова1,2, Д. А. Решетов1, В. Е. Жерновков1, Д. В. Влодавец1, Т. В. Димитриева1, А. В. Дейкин1,2
1 Марлин Биотех, Москва
2 Центр коллективного пользования, Институт биологии генов Российской академии наук, Москва
3 Кафедра неврологии, нейрохирургии и медицинской генетики, педиатрический факультет, Российский национальный исследовательский медицинский университет имени Н. И. Пирогова, Москва

Мышечная дистрофия Дюшенна (МДД) — распространенное наследственное заболевание, развивающееся вследствие мутации в гене дистрофина и приводящее к смерти в детском возрасте. На момент написания статьи пациентам было доступна только поддерживающая терапия, однако подходы к лечению МДД активно разрабатываются, и перспективным является метод пропуска экзонов. Его суть заключается в восстановлении рамки считывания гена путем индукции альтернативного сплайсинга. В результате синтезируется укороченный дистрофин, который в той или иной степени сохраняет функциональность. В работе дана оценка функциональности укороченных форм дистрофина, получающихся при коррекции нонсенс-мутаций и внутривариантных инделов по методике пропуска экзонов. Оценка производилась по данным о фенотипе носителей мутаций в гене дистрофина, взятых из базы LOVD (Leiden Open Variation Database). Было обнаружено, что одни и те же мутации способны проявляться как различные фенотипы, что, возможно, объясняется разным генетическим фоном пациентов. Так, экзона 48, для которой в LOVD есть 97 записей, в 2 % случаев приводила к бессимптомному течению заболевания, в 60 % — к миодистрофии Дюшенна, в 12 % — к миодистрофии Беккера (отличается более мягкой симптоматикой, чем МДД), в 26 % случаев — к промежуточному фенотипу. Высокая фенотипическая вариабельность мутаций в гене дистрофина ставит вопрос о границах применения методики пропуска экзонов для терапии наследственных миопатий.

Ключевые слова: мышечная дистрофия, миодистрофия Дюшенна, миодистрофия Беккера, пропуск экзонов

Благодарности: авторы выражают благодарность Центру коллективного пользования Института биологии гена РАН за предоставленное для экспериментов оборудование.

Для корреспонденции: Зотова Евгения Дмитриевна
119334, г. Москва, ул. Вавилова, д. 34/5; zotova@ibb.msu.ru
Статья поступила: 17.06.2016 Статья принята к печати: 24.06.2016

BULLETIN OF RSMU | 3, 2016 | VESTNIKRGMU.RU


Duchenne muscular dystrophy (DMD) is an X-linked recessive hereditary disease that occurs in about 1 out of every 3,600 male infants and it is the most common hereditary disease that causes death in childhood [1, 2]. Skeletal, cardiac, and smooth muscle failure can occur in DMD. The disease can also cause damage to the digestive and excretory systems. But, the skeletal muscle is affected the most. Apart from progressive muscle weakness, development delay, speech and breathing problems and heart failure are all observed in patients [3]. While diagnosed at 5 years average, DMD progresses to wheelchair dependency by the age of 13 years (95 % of patients). The mean age at which the patient dies not receiving special treatment is 19 years [4].

Since DMD is an X-linked disorder, it affects men in most cases, who are hemizygous by X-chromosome. The disease is usually inherited by the son from the carrier mother with X-chromosome. However, it may also occur de novo as a result of mutation in the dystrophin gene (DMD) [5]. The dystrophin gene is considered the longest in the human genome, measuring 2.4 million base pairs [6], and has 79 exons. The encoded protein contains approximately 3,700 amino acid residues [7]. Dystrophin mRNA requires about 16 hours to be synthesized [8].

Dystrophin is a structural protein which makes part of the costamere complex, responsible for connecting acto-myosin complexes (sarcomeres) to plasma membrane (sarcolemma) and proteins of extracellular matrix [9]. Dystrophin has an elongated shape. One of its ends binds actin, and the other — the membrane dystroglycan complex, also anchored to the spectrin cytoskeleton inside and to extracellular matrix proteins outside the cell. In the absence of dystrophin, complexes of proteins associated with it lose their stability. It is known that two protein complexes are part of costameres: dystrophin-glycoprotein complex and vinculin-talin-integrin complex, apparently performing similar functions. Damage to various links of this complex and non wholly understood system can lead to different types of hereditary myopathies [10, 11]. However, apparently due to the complex structure of the system and duplication of the functions among costamere proteins, even complete absence of dystrophin expression leading to DMD does not completely break the link between the actomyosin complexes, membrane and extracellular matrix, but only strongly weakens its strength. Because of this, fragile sarcolemma is subjected to mechanical damage during muscle contraction. Loss of sarcolemmal integrity leads to muscle fiber necrosis and inflammation. In the patient’s first years of life, his muscle fibers are regenerated through pool of myosatellites — muscular differon stem cells. However the pool of myosatellites gradually get depleted, resulting in muscle degeneration and fibrosis [12].

The human dystrophin contains 4 structural domains: N-terminal actin-binding domain; central rod domain containing 24 spectrin-like repeats and 4 hinges; cysteine-rich domain, binding β-dystroglycan; C-terminal domain binding dystrobrevin and syntrophin. The central rod domain has a strongly extended form, and since each spectrin-like repeat consists of three α-helices, it is apparently sufficiently flexible and elastic [12]. According to data obtained from analysis of known mutants, a small change in the number of spectrin repeats caused by frameshif deletions does not have much impact on the normal operation of the protein [13].

A shift in the reading frame in the dystrophin gene typically leads to DMD. The shift causes premature termination of translation and nonsense-mediated decay of mRNA, nonsense mutations, as well as large deletions in the gene regions encoding the N- and C-terminal domains of dystrophin. This results in completely disrupted binding of either with actin or with membrane dystroglycan complex. Non-frameshift deletions of average size in the middle of the gene are usually associated with other myopathy: Becker muscular dystrophy, which has less severe symptoms — many patients retain in adulthood the ability to walk independently [14].

A range of DMD therapies aimed not only at suppressing inflammation and fibrosis and reducing the toxic effects of excess of calcium in the cytoplasm, but also at restoring dystrophin expression are being developed [15]. One of the most promising therapies is exon skipping. In this method, natural oligonucleotides or their synthetic analogs are used to eliminate some exons from mature mRNA during splicing due to steric hindrance in spliceosome. Thus, with the presence of frameshif deletions or nonsense mutations, exclusion of one or more exons can translate all the subsequent part of the gene in the correct reading frame [16]. Translation of truncated but partially functional protein can significantly improve the condition of DMD patients, especially if treatment is started at early age [1].

One of the first studies on exon skipping as a correction therapy for DMD involved four patients — carriers of the dystrophin gene with frameshift deletion of one or more exons: 50, 52, 49–50 and 48–50. In all the four cases, exon 51 skipping induction is able to restore the reading frame in theory. Therapy was carried out by intramuscular injection of antisense oligonucleotides, and although exon skipping efficiency was 46–90 %, results of standard physiological tests have not been improved [17].

However, according to Aartsma-Rus and van Ommen [18], exon skipping can help most DMD patients. Exceptions will be mutations located between exons 64 and 70, which appear to be necessary for the functioning of dystrophin, as well as deletions disrupting actin-binding domains in the N-terminal region or affecting the first or last exon, and large chromosomal rearrangements such as translocations. However, the above-mentioned mutations are rare and together account for less than 10 % of all the described mutations in the dystrophin gene. Thus, theoretically, expression of functional or partially functional protein can be restored via exon skipping in about 90 % of DMD patients.

Prediction and analysis of the efficiency of exon skipping therapy are an important applied task since due to the large size of gene, the frequency of mutation in it is very high and about every third DMD patient has a de novo mutation [13]. An important theoretical aspect of the applicability of the method is assessment of the functionality of the dystrophin, truncated by exon skipping. Such assessment was the aim of this work. To make such assessment, it was necessary to decide which mutations can be analyzed, how they can be corrected (i.e. skipping of which exons would restore dystrophin expression) and which methods of analysis of phenotypic expression of mutations can be used. We studied and summarized literature data on these aspects of the study.

The simplest case for DMD gene correction is the presence of a nonsense mutation or frameshift insertions and deletions (indels) within a single exon. Theoretically, such mutations can be corrected by skipping one or a group of adjacent exons whose removal from mature mRNA will not lead to frameshift. In this case, 7 exons at most can be skipped — this is the exact number allowed to be excluded from the mature mRNA for the method to remain efficient [18].

Out of 79 exons of the dystrophin gene, 34 can be removed without disrupting the reading frame. Mutations in 29 exons
can be corrected by removing exons together with one of its neighbors. Moreover, 8 out of these 29 exons may be removed together with an upstream or a downstream neighboring exon (fig. 1). Errors in the remaining 14 exons can be corrected by removing a greater number of exons: exons 6–8 without frameshift are removed only together, as well as exons 76–78; exons 61 and 71 are removed only with two upstream exons, while exon 67 is removed with both upstream and downstream neighboring exons; exons 72 and 73 are removed only together with all the upstream exons up to exon 69 (four and five exons respectively); exon 74 can be removed in the same way or in a unit consisting of exons 70–75 (six exons in both cases), while exon 75 — only with the five upstream exons. Correction of exon 2 requires removal of six exons. However, Wein et al. [19] claim that exon 2 deletion leads to a very weak or asymptomatic disease course since translation starts from the internal ribosome entry site in exon 5, and in this case, truncated dystrophin retains functionality. Correction of mutations in the first and last exons is considered impossible (though this may be incorrect considering the fact that an internal ribosome entry site exists in exon 5). Thus, theoretically, mutations in 75 exons can be corrected by skipping no more than 6 exons, corresponding to the stipulated limit for the exon skipping method.

The number of possible corrections of single nucleotide substitutions in the dystrophin gene is extremely high. Experimental checking of all combinations in animal models is hardly possible. Besides, mouse model of DMD does not reproduce human DMD phenotype too well [20]. A more realistic approach to solving this problem is to analyze a large amount of clinical data of patients with mutations in the dystrophin gene that lead to DMD or BMD. Obtaining a sufficient sample from literary sources is a time-consuming process complicated by the fact that not all mutations have been published. Fortunately, there are databases of genetic variations. One of them — Leiden Open Variation Database (LOVD) [21] — contains records of different mutations in the dystrophin gene, information on the gender and population affiliation of their carriers, analysis method and most importantly, information on phenotype to which this or that mutation in a given patient leads to: DMD, BMD or asymptomatic course. Users around the world can add records to LOVD on their own. This is both an advantage (data vastness) and a disadvantage (possible errors) for the database. As of June 2016, LOVD contained over 16,000 non-unique records of mutations in the dystrophin gene relating to one or more exons. Deletions account for 83 % among them.

METHODS
Data on mutations were found using LOVD’s built-in text search system [22]. A search query was introduced in the field ‘Exon’. The query was created according to the database nomenclature: N-1i_Ni for exon N, which means that the mutation should affect the introns adjacent to the exon. For example, to search for exon 3 deletion, the query 2i_3i was created. For deletion of an exon group, the query was made in similar fashion. Only records containing information about full deletions of targeted exons (exons with which nonsense mutation or internal exon indels can be corrected) were taken for analysis. Duplications, as well as mutations carried by women were excluded from this list.

RESULTS
The distribution of phenotypes corresponding to the exon combinations selected for reading frame correction is shown in fig. 2. For 25 deletions no record was found in LOVD. Most of such deletions were located in the central domain region, as well as in the region of exons 65–78. Only in 11 combinations the number of cases found in the database exceeded 5. Distribution of these 11 combinations by dystrophin domains was very uneven: 7 of them were combinations correcting exons 45–55 (C-terminal part of the central rod domain), 1 combination — exon 13 deletion — affects the N-terminal region of the central domain, while 3 combinations affect the N-terminal actin-binding domain.

**Fig. 1. Exon structure of the dystrophin gene**

Green blocks indicate exons; exon size (in base pairs) is indicated in the bottom right and its sequence number in the middle. Three types of joints (straight, concave and convex) indicate the three reading frames within which exon boundaries fall. Arrows indicate the gene region borders encoding the structural domains of dystrophin: CH1 and CH2 — elements of the N-terminal actin-binding domain, R1–R24 and H1–H4 — spectrin repeats and hinges of the central rod domain, respectively, WW, ZZ, EFH1 and EFH2 — structural elements of the cysteine-rich domain, and C term — a C-terminal domain unique for dystrophin and its homologues (Nicolas et al., 2012 [29]).
BMD or asymptomatic course were observed in 45 % of all the cases (corresponding to the deletions we chose for analysis) represented in the database, whereas larger frameshift deletions of exons manifest as BMD in less than 4 % of cases, indicating the milder nature of muscular dystrophy in patients with non-frameshift deletions.

It was also established that the same mutation could manifest itself as a variety of phenotypes. Specifically, exon 48 deletion, which is found 97 times in the database, may be asymptomatic (2 % of cases) or result in BMD (60 % of cases), DMD (12 % of cases) or intermediate phenotype (26 % of cases). Even more interesting are deletions of exons 50–51 and 51–52, for which 14 and 11 cases are described respectively: there are 6 (43 %) total of BMD or asymptomatic course for the first, and none for the second, although both deletions do not affect anything except the hinge region and one of the spectrin repeats of the central domain, and as it would seem, should not have significant effect on the protein structure. Exon 47 deletion is another frequent non-frameshift deletion (13 cases), which also does not affect anything except spectrin repeats, but manifests as DMD or mixed phenotype in 100 % of cases.

**DISCUSSION**

Records on patients (carriers of different deletions) in the database quite strongly vary in number. The reasons for this may be both unequal frequency of mutations in different regions of the gene and under-representation of a group of mutations in the database. Indeed, such distribution can be explained by the presence of meiotic recombination hot spots in the regions of exons 7 and 44 [23]. Absence of information on deletions in the central domain region can be attributed largely to asymptomatic manifestation of such deletions which is likely to greatly reduce the probability of such mutations being detected. However, it is surprising that the database does not contain any combinations for correction of mutations in exons 65–78. One possible explanation for that is a large size of exon combinations needed to be deleted to correct mutations at C-end.

According to Aartsma-Rus et al. [24], 91 % of mutations in the LOVD database are in agreement with the following rule: frameshift mutations and translation of truncated non-functional dystrophins are the cause of DMD, while non-frameshift mutations usually almost do not affect the functionality of dystrophin and they cause BMD, if they do not affect the key domains located on the protein’s N- and C-ends.

A confirmation of this rule apparently is the case described in the article Passos-Bueno et al. [25]. It is about a patient with BMD induced by a major non-frameshift deletion (from exon 13 to 48). This deletion captures 18 out of 24 spectrin repeats of the central domain, but does not touch the N- and C-ends. Of protein amino acid sequence only about a half remains, which, according to the authors, is enough for dystrophin to remain partially functional. However, this rule has some interesting exceptions. For example, Takeshima et al. [26] describe a case of large N-terminal non-frameshift deletion (from exon 3 to 41), which manifests itself in the form of an intermediate phenotype between DMD and BMD. The absence of an N-terminal domain makes dystrophin dysfunctional. Nevertheless, the patient’s phenotype is not consistent with the classical picture of DMD course.

Another exception is nonsense mutation in exon 76 described in Suminaga et al. [27], which leads to synthesis of truncated form of dystrophin, detectable by immunohistochemistry on muscle biopsy preparations. At the same time, except elevated levels of creatine phosphokinase in the blood (which is one of the key markers of DMD and BMD), no other muscular dystrophy symptom was detected in a patient. The authors write that they cannot explain the picture obtained: another nonsense mutation described in the literature, which cuts the protein only into 10 amino acids closer to the beginning than the mutation found in [27], leads to a typical picture of DMD course [28]. The work tested different assumptions on the possible causes of such phenotype: somatic mosaicism, restoration of reading frame through alternative splicing, compensatory high level of expression of utrophin, the dystrophin homologue protein. However, all the assumptions were rejected.

As seen from analysis of the database and the cases mentioned above, same mutations in the dystrophin gene can lead to different phenotypes. The reasons for this phenomenon are not fully understood, but it is probably due to patients’ different genetic background. This is a problem since even if effective skipping of exon or exons, resulting in reading frame restoration, is induced in the patient and synthesis of truncated dystrophin is observed, there is still no guarantee that the patient’s condition will improve. Besides, due to insufficient knowledge of the problem, there is no reliable method of

---

**Fig. 2.** Distribution of phenotypic manifestations of deletions in the dystrophin gene in exon skipping to restore the reading frame (based on LOVD data)

The vertical axis indicates the numbers of exons that were included in the deletion, while the horizontal axis is the number of records containing data on a particular deletion, which are contained in the LOVD database. Sections of horizontal bars correspond to representation of individual phenotypes.
predicting whether frame restoration would help a particular patient. Statistical analysis in this case is also difficult because phenotypic data are sufficient for only a small number of deletions.

Our work did not consider some variants of truncated dystrophins obtained by exon skipping correction of mutations that are more complex for analysis: large deletions affecting at least one exon entirely and inducing frame shift, and duplications of one or more exons. Some of them can be corrected by skipping the same exons as those used for correction of intragenic mutations, for example, duplication of exons, deletion of which causes no frame shift, or deletion of a whole exon, which can be corrected by deletion of a neighboring one. So, in skipping both copies of exon, duplication of exon 13 is reduced to deletion of exon 13, while exon 51 deletion can be corrected by skipping exon 50 or exon 52. Correction of some duplications of two or more exons and large deletions, leading to a frameshift and affecting two or more exons, is also feasible. So, deletion from exon 7 to exon 34 can be corrected by skipping exon 6. The resulting deletion will not disrupt the reading frame. However, it is unknown whether this form of dystrophin would be functional. Moreover, LOVD contains no data on this particular mutation. As a result, it was decided not to carry out detailed analysis of such mutations due to the large number of combinations and lack of data on the functionality of the dystrophin forms resulting from correction of the mutations.

CONCLUSIONS

Analysis of data on phenotypic mutations in the dystrophin gene showed that the same mutations can manifest various phenotypes. This indicates that the success of exon skipping for treatment of Duchenne muscular dystrophy is probabilistic in nature even in the case of highly efficient induction of alternative splicing. Accurate assessment of the probability of success for a particular therapy is difficult due to lack of data. However, one can hope that the situation will change with the emergence of higher-throughput and cheaper DNA sequencing methods.

References

Литература