Available online at www.sciencedirect.com

SciVerse ScienceDirect

Neuromuscular Disorders 23 (2013) 4–14

www.elsevier.com/locate/nmd

Review

The medical genetics of dystrophinopathies: Molecular genetic diagnosis and its impact on clinical practice

Alessandra Ferlini *, Marcella Neri, Francesca Gualandi

Section of Medical Genetics, Department of Medical Sciences, University of Ferrara, Italy

Received 6 April 2012; received in revised form 23 August 2012; accepted 4 September 2012

Abstract

A large variety of mutations in the dystrophin gene cause Duchenne and Becker muscular dystrophies, diseases affecting predominantly the striated muscles (skeletal and cardiac). Rare mutations also account for the allelic disorder isolated X-linked dilated cardiomyopathy. Dystrophin protein is encoded by a huge gene located on the X chromosome and the understanding of its complex genomic architecture has unraveled general key functions in gene expression regulation. Dystrophin also exists as a number of other tissue specific isoforms, some exclusively or predominantly expressed in the brain and/or in other tissues. Genotype definition of the dystrophin gene in patients with dystrophinopathies has taught us much about functionally important domains of the protein itself and has also provided insights regarding several regulatory mechanisms governing the gene expression profile. This review focuses on the current understanding of the dystrophin mutations heterogeneity, genotype-phenotype correlations, as well as interpretation of the functional significance of mutations that often require non routine genetic studies. It also explores the impact of genetic diagnosis on clinical definition and on the discovery of biomarkers and personalized therapies.

Our aim is to offer an overview of the medical genetic approach on the dystrophin gene and dystrophinopathies with implications for clinical practice and therapeutic perspectives.

Keywords: Duchenne and Becker muscular dystrophies; Dystrophin; Molecular genetics; Mutations; Genotype phenotype correlation

1. The dystrophin gene and protein: structure, architecture and regulation

The dystrophin gene was the first gene isolated by positional cloning and has proved to be the most complex genetic locus still identified.

Dystrophin and titin are the largest disease genes in the human genome [1,2]. Dystrophin spans 2.2 megabases at Xp21 on the human X chromosome and it is composed by 79 constitutive exons. Exon 79 is the longest (2.7 kb), being the one containing the full 3′ UTR region. The gene is driven by three main promoters, the upstream and ancient Brain (B), the Muscular (M) and the Purkinje (P), which guides the full length dystrophin transcription in a tissue/development specific way [3,4]. Furthermore, at least other four first exons have been recognized as adjacent to promoters driving short dystrophin isoforms, and localized within introns 29 (retinal isoforms or Dp260, R), 44 (Brain specific isoform or Dp140, B3), 55 (Schwann cells isoform or Dp116, S) and 62 (General isoform or Dp71, G) [1].

Several alternatively utilized exons, always located within introns, have also been described [1].

The architecture of the dystrophin gene, reporting the exon composition, isoforms, relevant domains and structural organization is reported in Fig. 1.

The DMD Leiden pages (http://www.dmd.nl/) report all the known intronic sequences and their relative accession numbers in the HGMP GeneBank.

* Corresponding author. Address: Section of Medical Genetics, Department of Medical Sciences, University of Ferrara, Via Fossato di Mortara, 74, 44121 Ferrara, Italy. Tel.: +39 0532 974439; fax: +39 0532 974406.
E-mail address: fla@unife.it (A. Ferlini).

0960-8966/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.nmd.2012.09.002

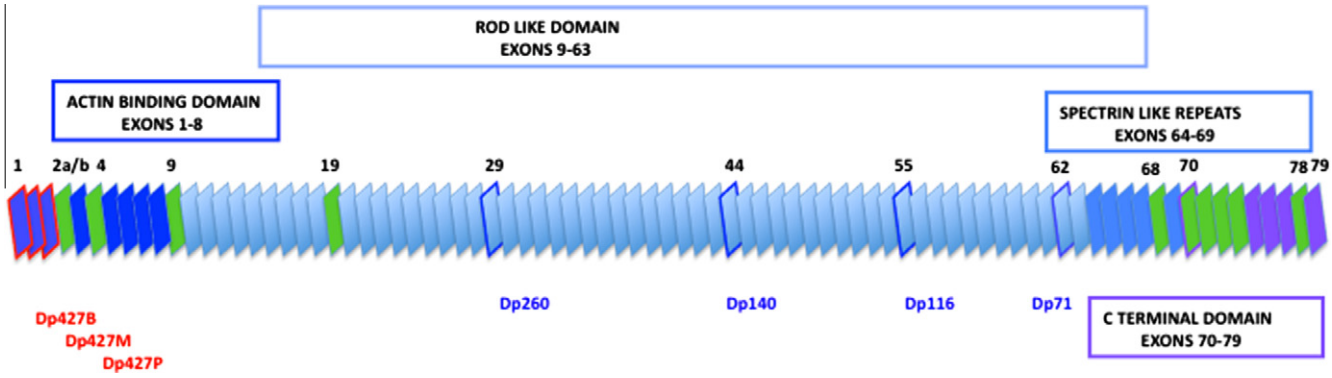

Fig. 1. Dystrophin gene structure and protein domains. Schematic representation of 79 exons of dystrophin gene with isoforms and protein domains. Lines in red represent the 5′ full length promoters and their first exon (isoforms Dp427B-M-P). Lines in blue represent the 3′ promoters and their first exons of isoforms: Dp260 (retinal), Dp140 (brain 3), Dp116 (Schwann cells), Dp71 (general). In green are represented exon alternatively spliced or skipped. Boxes' different blue/violet colors explain the protein domains corresponding to the different exonic regions.

The full sequencing of dystrophin introns has revealed that very large and in some cases huge introns are very common in the architecture of this gene. This is particularly true for the introns close to alternative transcription starting sites (as introns 1-Muscle, intron 1-Brain and intron 1-Purkinje, as well as intron 44). This fact together with the high occurrence of alternative splicing events in this gene raised the question why such large introns have been maintained during evolution. The dystrophin gene's unusually large intron size has been claimed as one of the major causes of the high mutation rate known to occur in some regions of the gene which give rise to two well known mutation hot spots [1,5,6]. Furthermore, the detailed analysis of the large introns, emerging from the full Genome Project sequencing, have revealed interesting characteristics and focused attention on possible roles and functions they may play in gene regulation, especially transcription and splicing [7,8]. Among dystrophin gene introns, the introns 7 and 44, which experience the highest recombination rates and are known to be mutational hot spots, seem to be sites of positive directional selection [9]. This suggests that these two introns may contain relevant regulatory motifs. Intron 7 is adjacent to a region of "exceptions to the rule" of the Monaco open-reading-frame theory [10] (exons 3–7); in fact a restarting dystrophin ATG located within exon 6 has been postulated as mechanism to rescue dystrophin translation in mutations located within exon 2–6 [11]. These mutations cause a BMD phenotype despite of being out-of-frame, pointing out again that this region might be involved in regulatory processes. Introns are the segments involved in the DNA duplication process and duplication forks formation, fact that is mechanistically linked to the gene disruption due to non-allelic homologous recombination (NAHR), non-homologous end joining (NHEJ), and microhomology-mediated replication-dependent recombination (MMRDR) mechanisms known to explain DNA rearrangements associated with genomic disorders. Recently, Ankala et al., [12] have shown that non-recurrent "de novo" rearrangements within the dystrophin gene are associated with microhomologies or short insertions within introns. In many of the DMD cases studied the sequences upstream of the deletion breakpoints were repeatedly replicated. This suggests that an aberrant miss-firing of replication origins may explain non-recurrent rearrangements within the DMD.

Other regulatory regions have been characterized [13–18] as having a role in enhancing gene expression often in a tissue specific way.

The dystrophin coding regions share clear homology with at least three other classes of genes. The entire dystrophin coding sequence is similar to that of utrophin [19] while the 5′ end and central portions share homology with members of the spectrin gene family, including α-actinin. The 3′ end of the dystrophin gene is homologous with an 87 kD post-synaptic protein (dystrobrevin) characterized from the Torpedo electric organ. These observations suggest that dystrophin and utrophin may have arisen during evolution by juxtaposition of ancestral spectrin and 87 kD-like genes to form a larger transcription unit.

The cloning of the complete cDNA sequences encoding dystrophin [20] and utrophin [21] has given information on the predicted structure of these large cytoskeletal proteins. From their deduced amino-acid sequences both proteins share a surprisingly high number of identical residues (46.3%). Dystrophin is a 427 kD protein predicted to fold into several distinct structural domains. The amino terminus contains a major actin binding domain encoded by exons 1–8; however two further actin interactive sites have been mapped further 3′. The majority of the molecule is encoded by exons 9–63 forming the large rod-like domain composed of 24 "spectrin like" repeats interrupted by a few spacer regions. Located downstream of these "spectrin like" repeats is the cystein-rich domain (exons 64–69) which is followed by the C-terminus (exons 70–79), that binds to a group of transmembrane proteins as a complex originally named the dystrophin-associated-glycoprotein complex (DAG) because of its link to dystrophin. These proteins have been re-named dystrophin associated