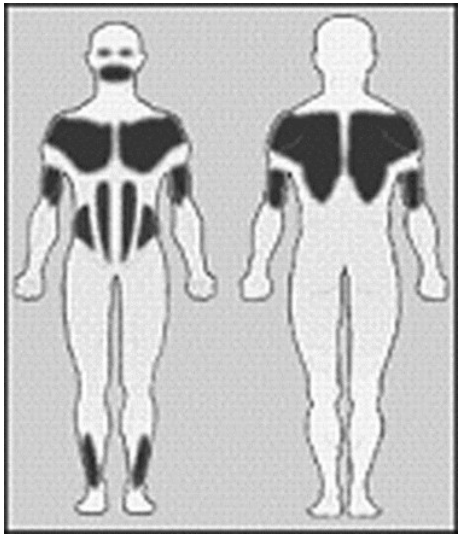
- **Facioscapulohumeral muscular dystrophy**
- **Spinal muscular atrophy**

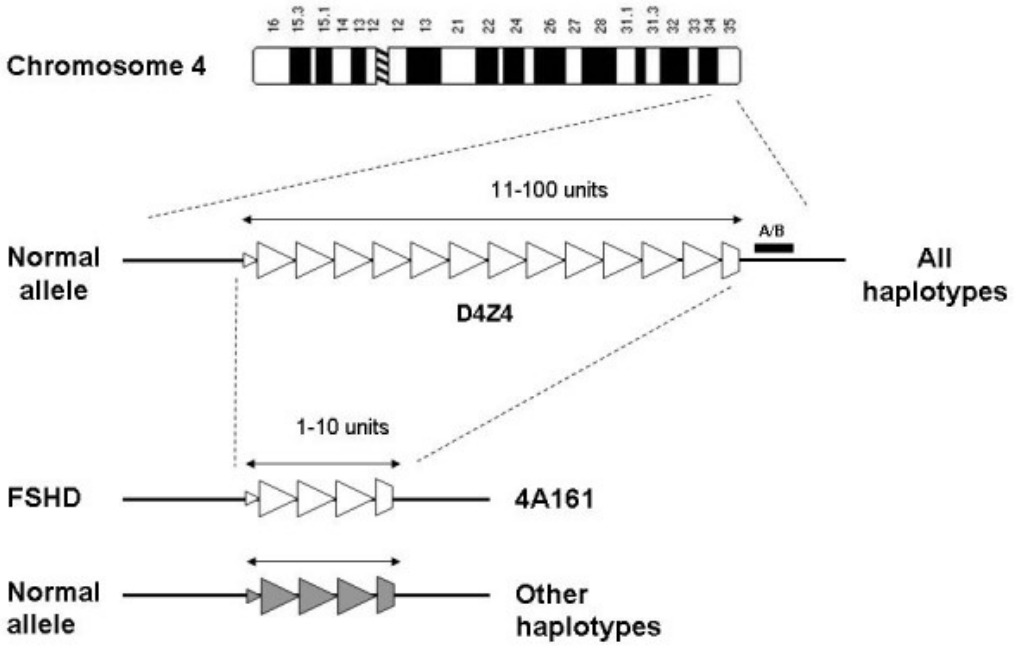
~ **Half of the human genome consists of repetitive DNA**, significant proportion is organized in **tandem arrays.**

- Repeat unit sizes **1- 4 nucleotides** and spanning less than 100 bp are typically defined as **microsatellite repeats**.
- Repeat unit sizes **10 - 40 nucleotides** covering several hundreds of bp are referred to as **minisatellite repeats**.
- The term **midisatellite repeats** has been proposed for loci containing repeat units of **40 - 100 nucleotides** that can extend over distances of 250–500 kb.
- **Macrosatellite repeats** are the largest class of repeat arrays with unit sizes of **>100 nucleotides** and can span hundreds of kb of DNA.

Facioscapulohumeral muscular dystrophy, FSHD

- The third most common muscular dystrophy after Duchenne muscular dystrophy and myotonic dystrophy
- The prevalence ~1:8 000 – 22 000
- AD inheritance
- FSHD locus: 4q35, deletion of macrosatelite repetitive sequence D4Z4





D4Z4 size: 3300 bp

• Normal DNA:

11-100 D4Z4

(36 000-330 000 bp)

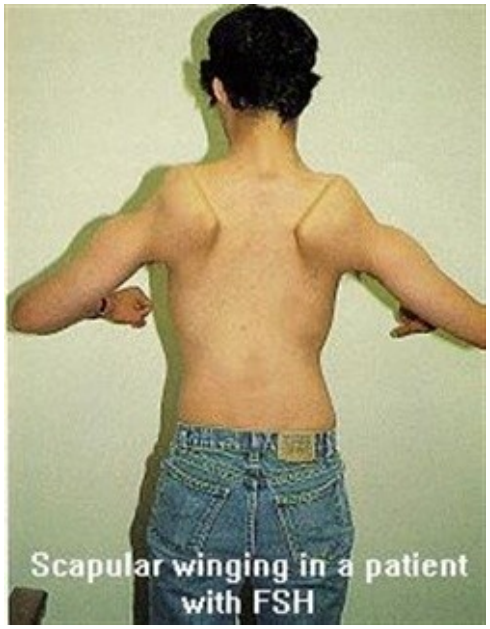
• FSHD patients:

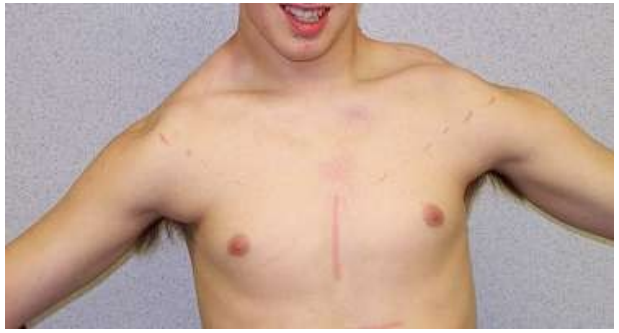
1-10 D4Z4

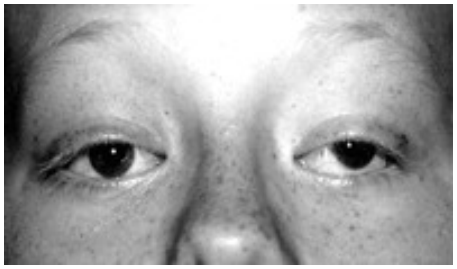
(<33 000 pb)

- Disease onset - **the second decade of life - initially weakness of shoulder and facial muscles.**

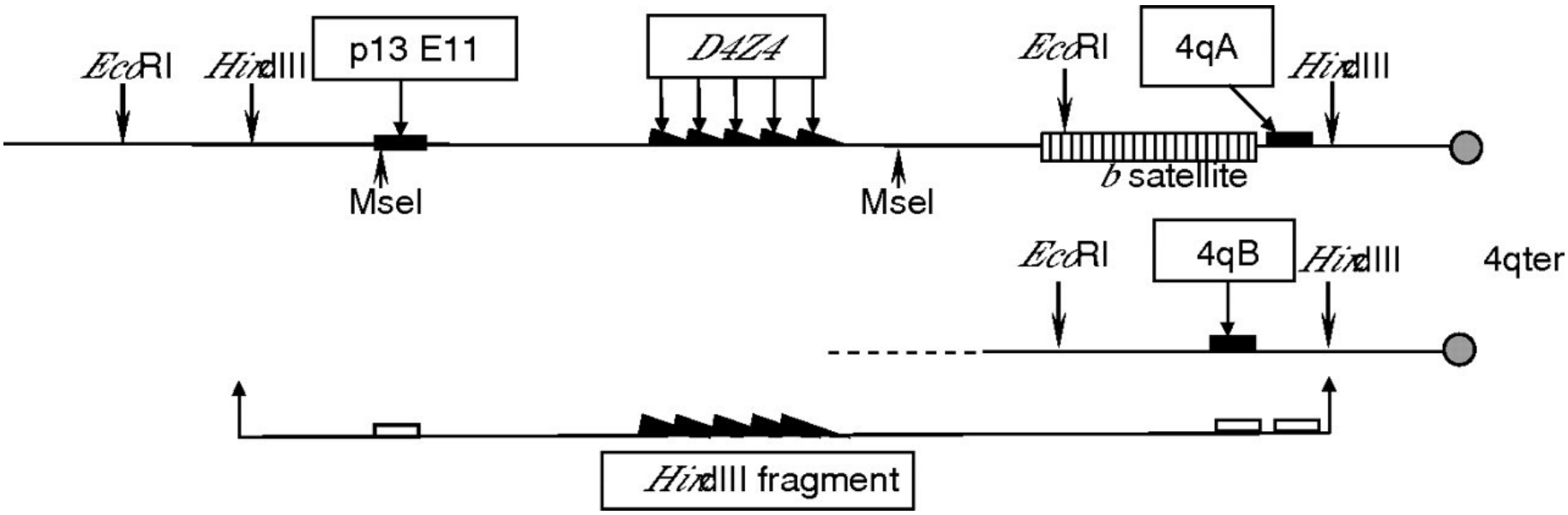
- The spectrum of disease severity is wide, ranging from **mildly affected to severely affected wheelchair-bound individuals**.







FSHD, 4qA and 4qB alleles

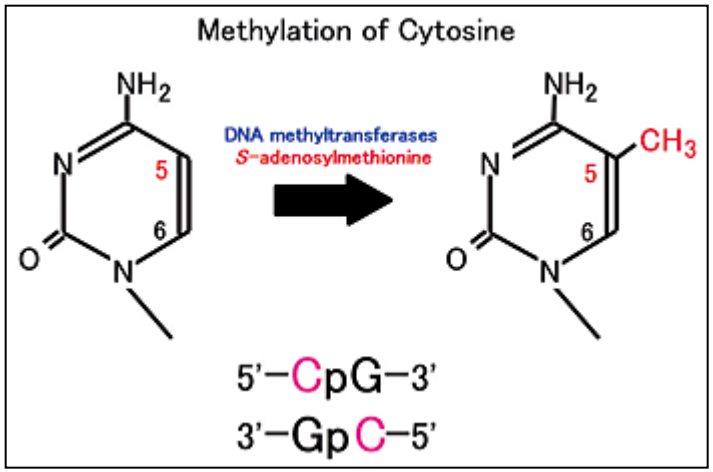


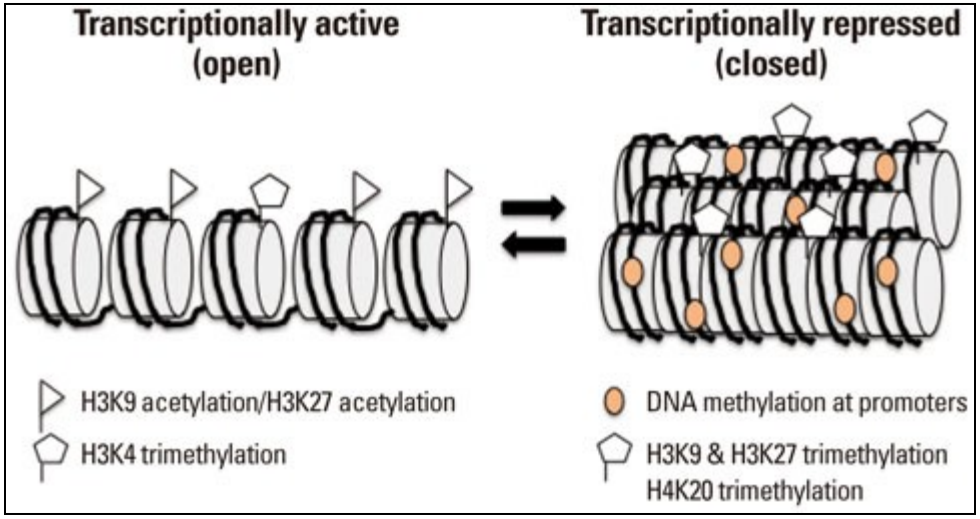
- Two allelic variants of 4q35 were identified - **4qA and 4qB**.
- **FSHD is associated with deletion of D4Z4 on 4qA (FSHD1).**

- **5% of FSHD patients do not have D4Z4 deletion on 4q35 but 4qA allele is present (FSHD2).**

D4Z4, 290 CpG - candidates for **DNA methylation**

(modification of DNA associated with chromatin condensation and gene silencing)



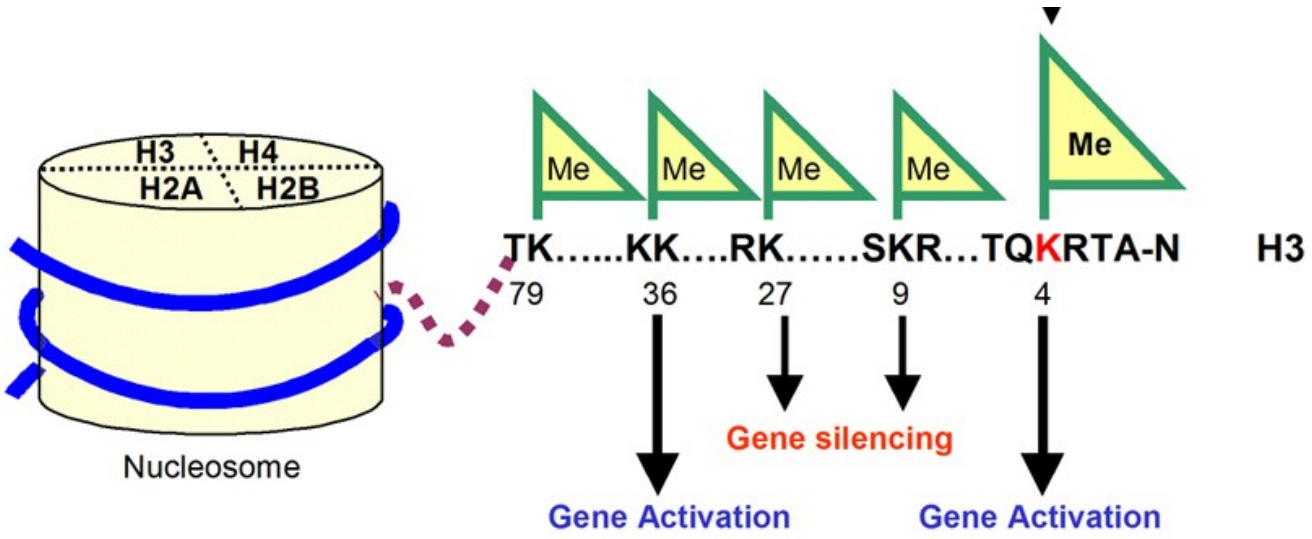


DNA methylation of D4Z4 is significantly reduced in

- **FSHD1 chromosome (D4Z4 deletion)**
- **FSHD2 chromosome (without D4Z4 deletion, both 4q35 chromosomes)**

Histone modifications

- Chromatin - DNA, histones and other chromosomal proteins; a major function of chromatin is packaging of DNA in the nucleus.
- Histones may undergo posttranslational modifications (acetylation, methylation, phosphorylation and ubiquitination).
- Specific histone modifications are associated with either transcriptional activation or transcriptional repression. Methylation at lysine residues 4, 36 and 79 of histone H3 has been correlated with transcriptional activation. In contrast, methylation at lysine residues 9 and 27 of histone H3 has been linked to heterochromatin and gene repression.



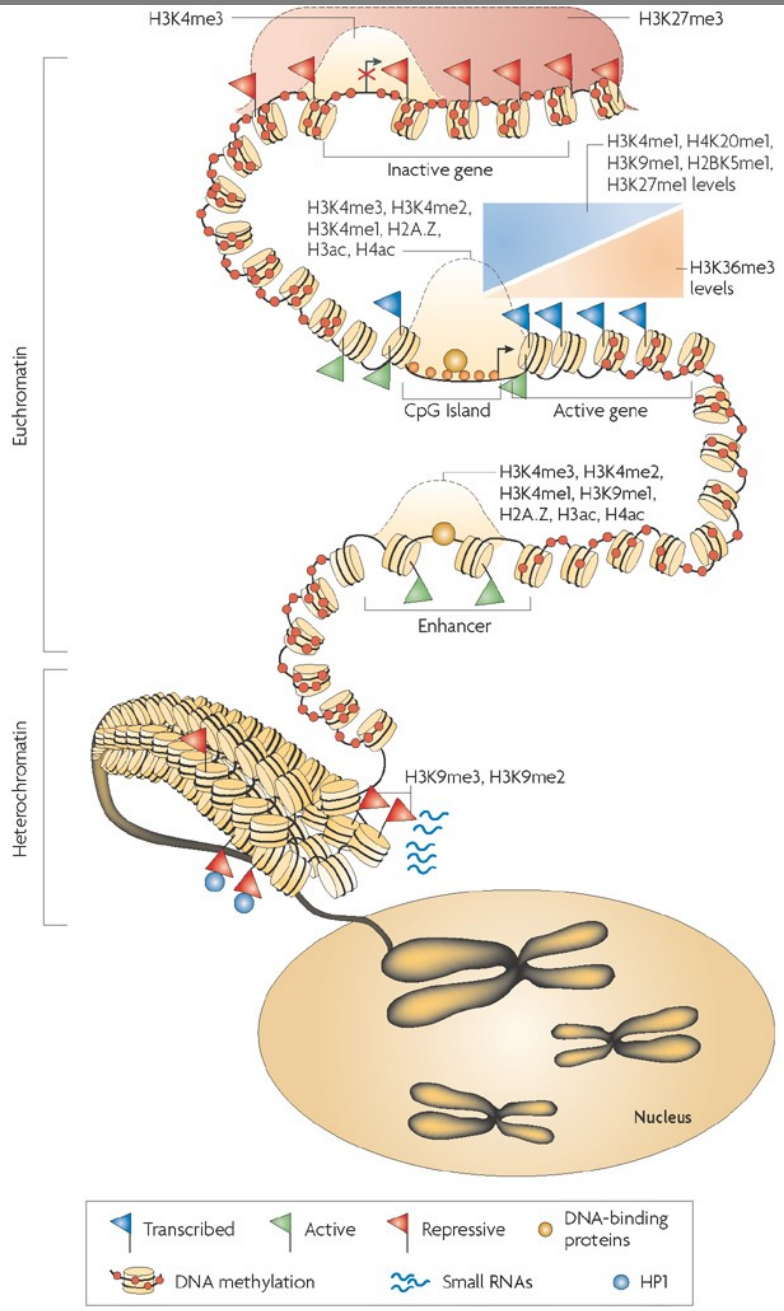
FSHD, epigenetic modifications

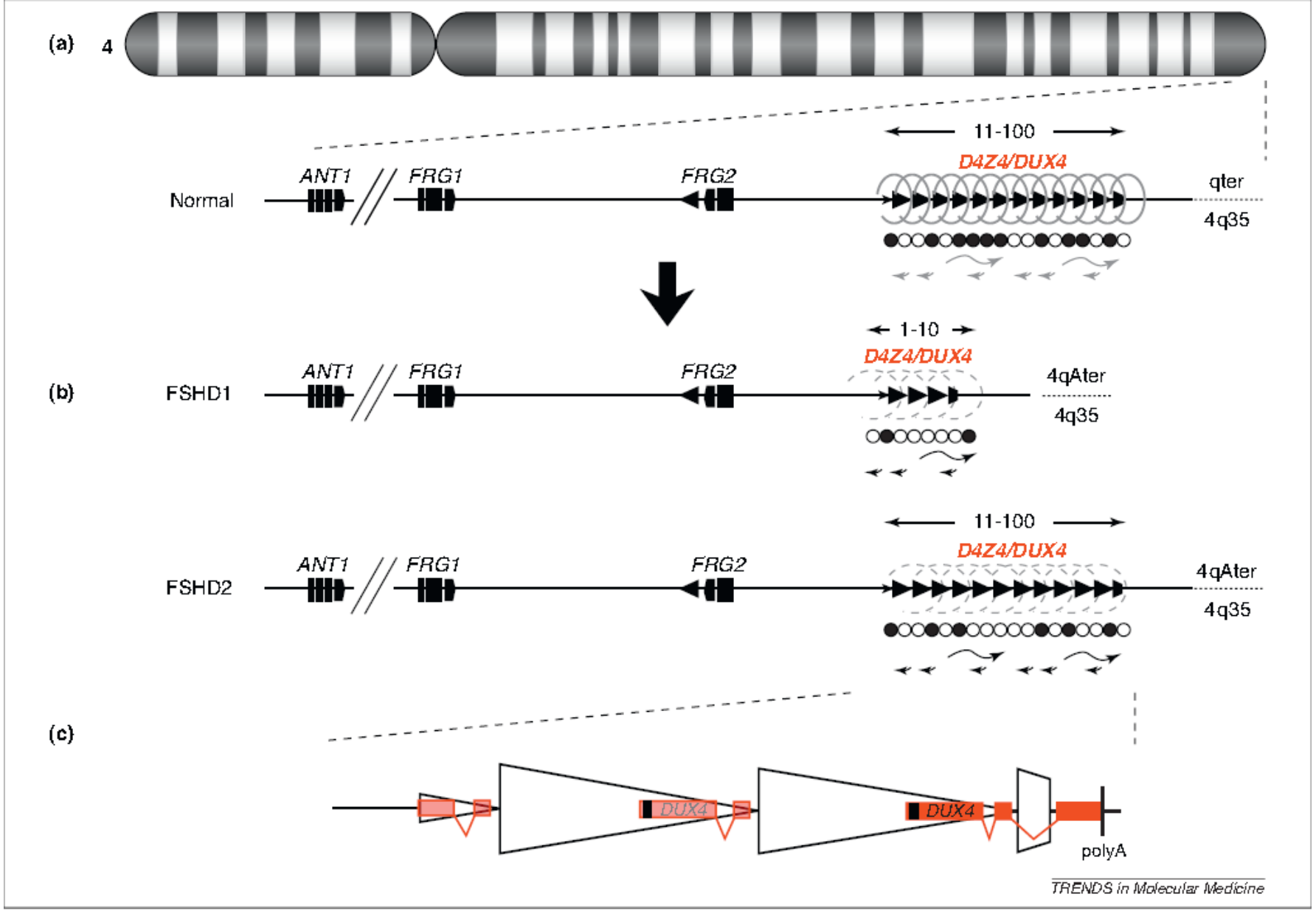
FSHD is associated with changes of DNA methylation and histone modification (H3K9me3).

- Normal allele 4q35:
DNA methylation, H3K9me3

- **FSHD1 chromosome (D4Z4 deletion):**
DNA hypomethylation, H3K9me3 ↓

- **FSHD2 chromosome:**
DNA hypomethylation, H3K9me3 ↓ (both chromosomes)

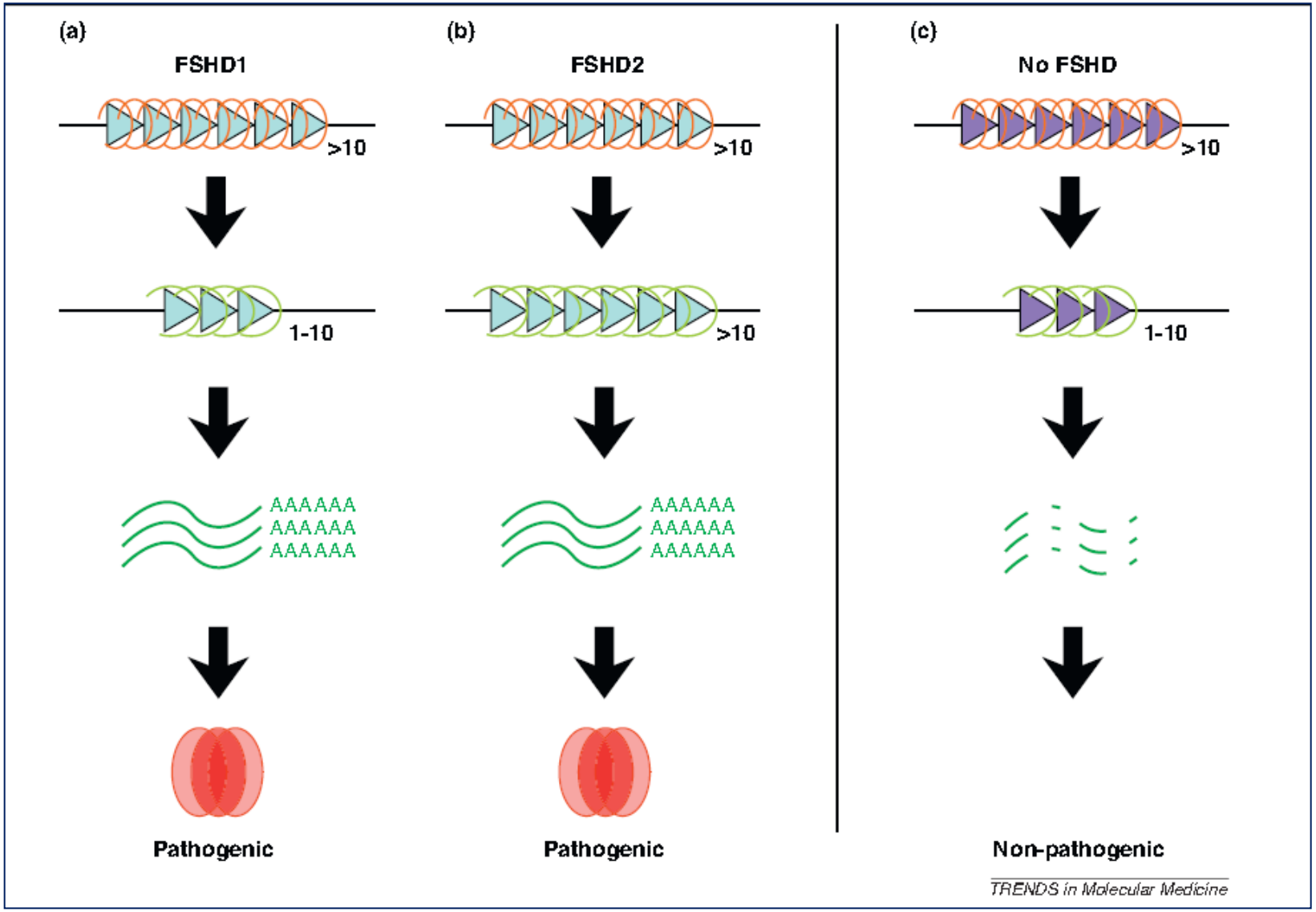




The *DUX4* gene located within D4Z4 unit. FSHD chromosomes, the last copy of *DUX4* splices to a third exon located in the region immediately flanking the repeat and stabilizing the transcript owing to the presence of a polyadenylation signal (association of FSHD and 4qA).

Schematic of the FSHD locus.
(a) The D4Z4 repeat (triangles) is located in the subtelomere of chromosome 4q and can vary between 11 and 100 copies in the unaffected population. This repeat structure has a closed chromatin structure characterized by heterochromatic histone modifications (dense springs), high DNA methylation levels (closed circles) and complex bidirectional transcriptional activity (gray arrows). Candidate genes DUX4, FRG2, FRG1 and ANT1 are indicated.
(b) In patients with FSHD, the chromatin structure of D4Z4 adopts a more open configuration (open springs and open circles) leading to inefficient transcriptional repression (black arrows) of the D4Z4 repeat.
(c) The DUX4 gene is located within each D4Z4 unit. On permissive chromosomes, the last copy of the DUX4 genes splices to a third exon located in the region immediately flanking the repeat and stabilizing the transcript owing to the presence of a polyadenylation (polyA) signal.

Silvère M van der Maarel, Trends in Molecular Medicine May 2011



FSHD1, FSHD2: the D4Z4 repeat array adopts a more open chromatin configuration leading to the expression of *DUX4* mRNA. This mRNA is stabilized owing to the presence of a polyadenylation signal immediately distal to the D4Z4 repeat array. The *DUX4* mRNA encodes for a nuclear double-homeobox protein that when expressed in muscle induces cell death.

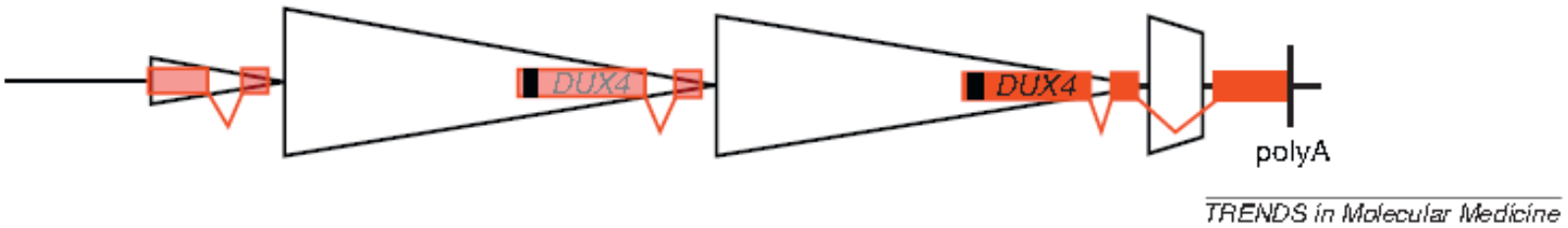
A unifying mechanism for FSHD.
Upon (a) contraction of the D4Z4 repeat (FSHD1) or by (b) a yet unknown mechanism (FSHD2, phenotypic FSHD), the D4Z4 repeat array (triangles) adopts a more open chromatin configuration (orange > green dots) leading to the leaky expression of DUX4 mRNA. On permissive chromosomes, this mRNA is stabilized owing to the presence of a canonical polyadenylation signal immediately distal to the D4Z4 repeat array. (c) Nonpermissive chromosomes do not have this polyadenylation signal and therefore DUX4 mRNA becomes rapidly degraded. The DUX4 mRNA encodes for a nuclear double-homeobox protein that when expressed in muscle induces apoptosis.

Silvère M van der Maarel, Trends in Molecular Medicine 2011

• D4Z4 contains the open reading frame of a double-homeobox transcription factor DUX4.

• 4qA, 4qB; the difference between 4qA and 4qB - the presence of a *DUX4* polyadenylation signal; full-length *DUX4* is produced from the last D4Z4 unit in early development and suppressed during cellular differentiation; in differentiated tissues, the D4Z4 array is associated with DNA methylation and H3K9me3.

• **In FSHD, the expression of the full-length *DUX4* transcript is not completely suppressed in skeletal muscle and results in expression of the full-length *DUX4* mRNA and protein inducing death of muscle cells.**

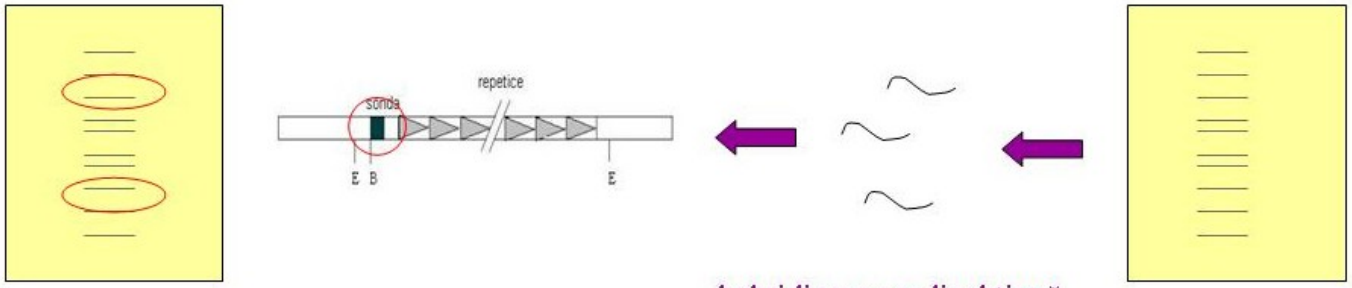


FSHD, DNA diagnostics

izolace genomové DNA

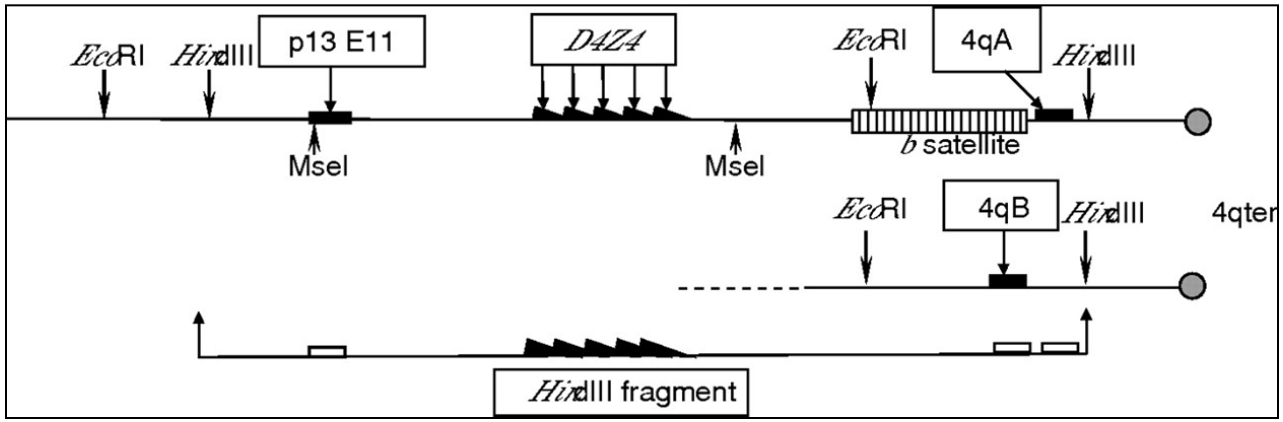
restrikční štěpení

pulzní gelová elektroforéza

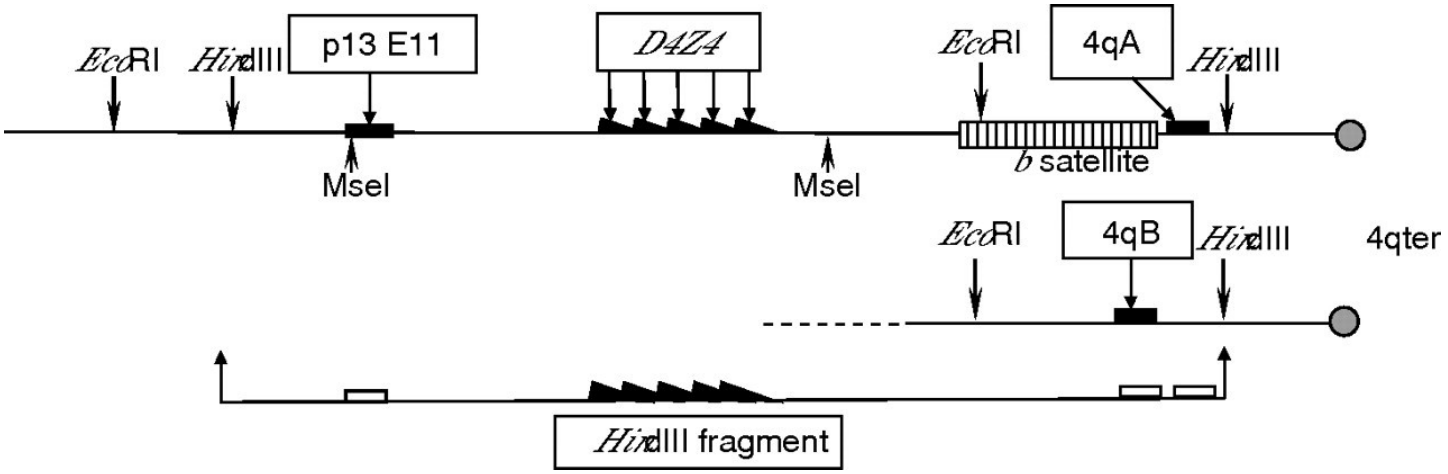


hybridizace s radioaktivně značenou sondou

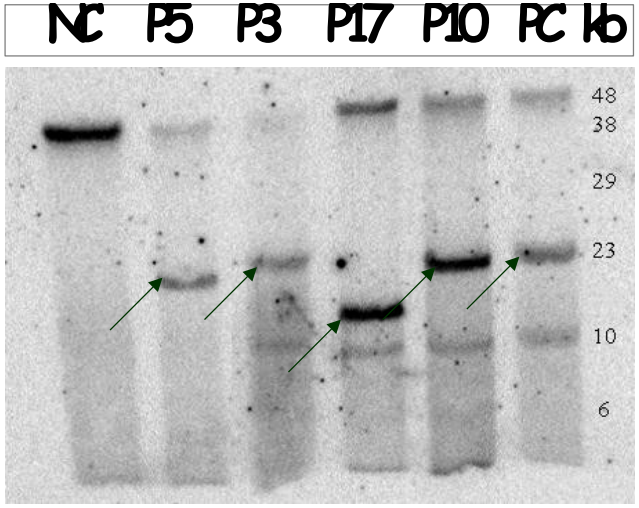
Southern blot

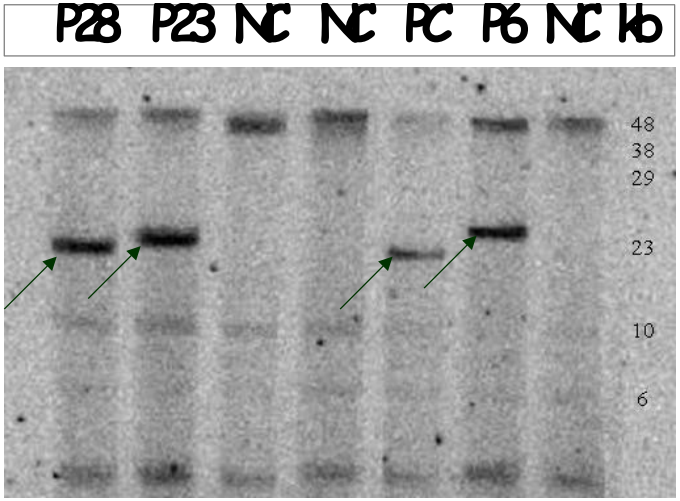


Genomic map of the facioscapulohumeral muscular dystrophy locus region containing 4qA-defined and 4qB-defined 4qter subtelomeres.



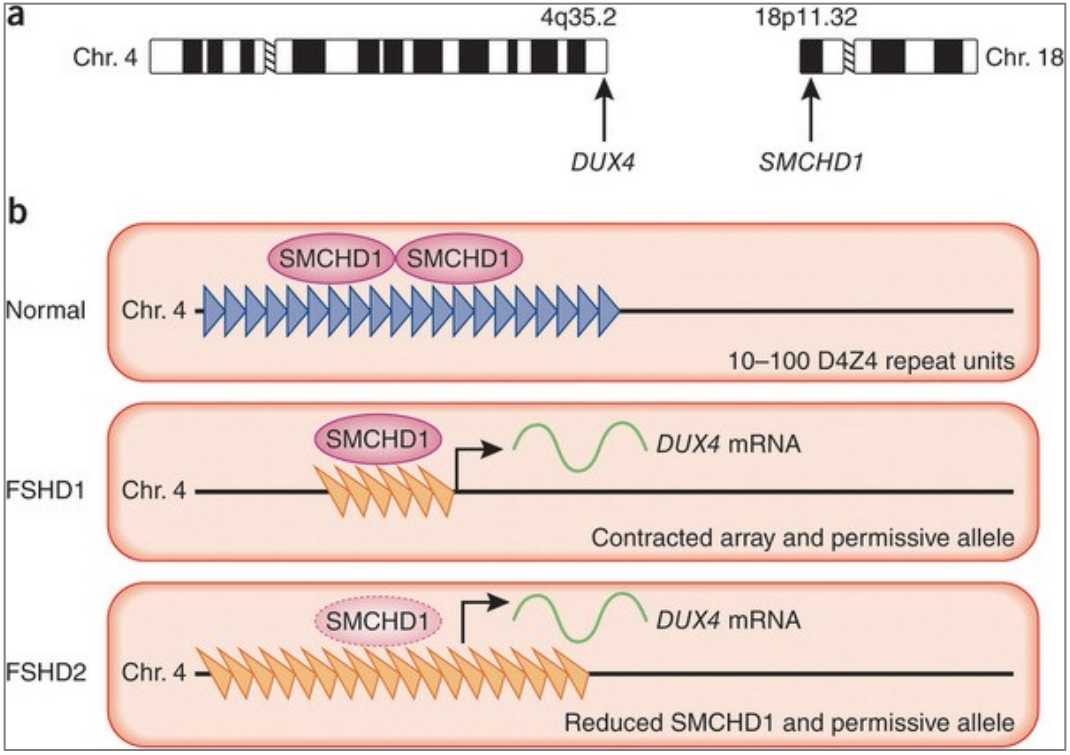
J Med Genet 2007;44:215-218





- **Pacienti s FSHD: detekce *Eco*R1 fragmentu < 43 000 b (< 10 repetic)**
- **Kontrola: detekce *Eco*R1 fragmentu 43 000 - 340 000 (11-100 repetic)**

- **The *SMCHD1* gene** - Structural Maintenance of Chromosomes flexible Hinge Domain containing 1
- The SMCHD1 protein is associated with the D4Z4 array in skeletal muscle cells, this association is required to maintain array silencing.
- **FSHD2 occurs in individuals who inherited normal-sized D4Z4 array on 4qA and a *SMCHD1* mutation.**

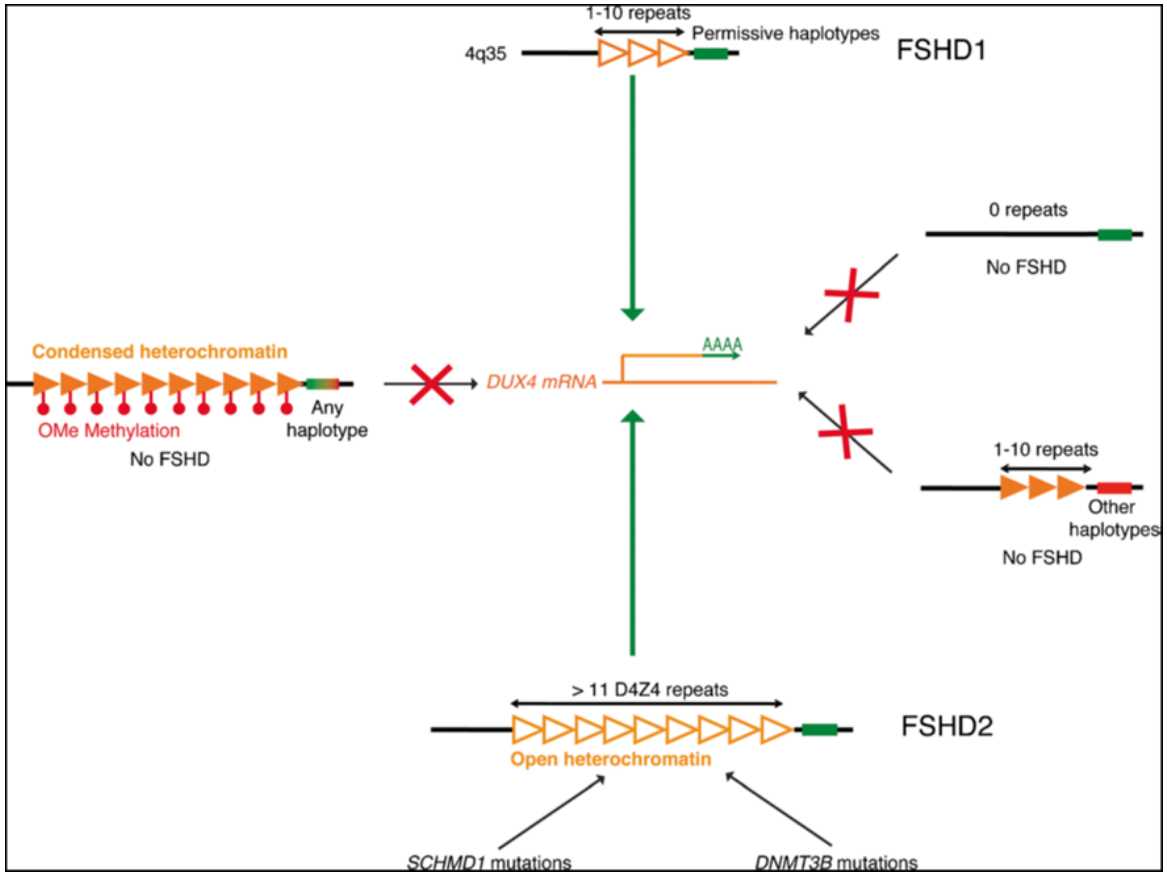


J.R. Lupski, Nature Genetics 2012

- The disease mechanisms of FSHD1 and FSHD2 - **D4Z4 chromatin relaxation and *DUX4* expression**
- FSHD1 and FSHD2 require inheritance of two independent genetic variations: **the *DUX4* gene with a polyadenylation signal (4qA)** and **a second genetic variant.**

- For **FSHD1**, a second genetic variant is **contraction of the D4Z4 array.**

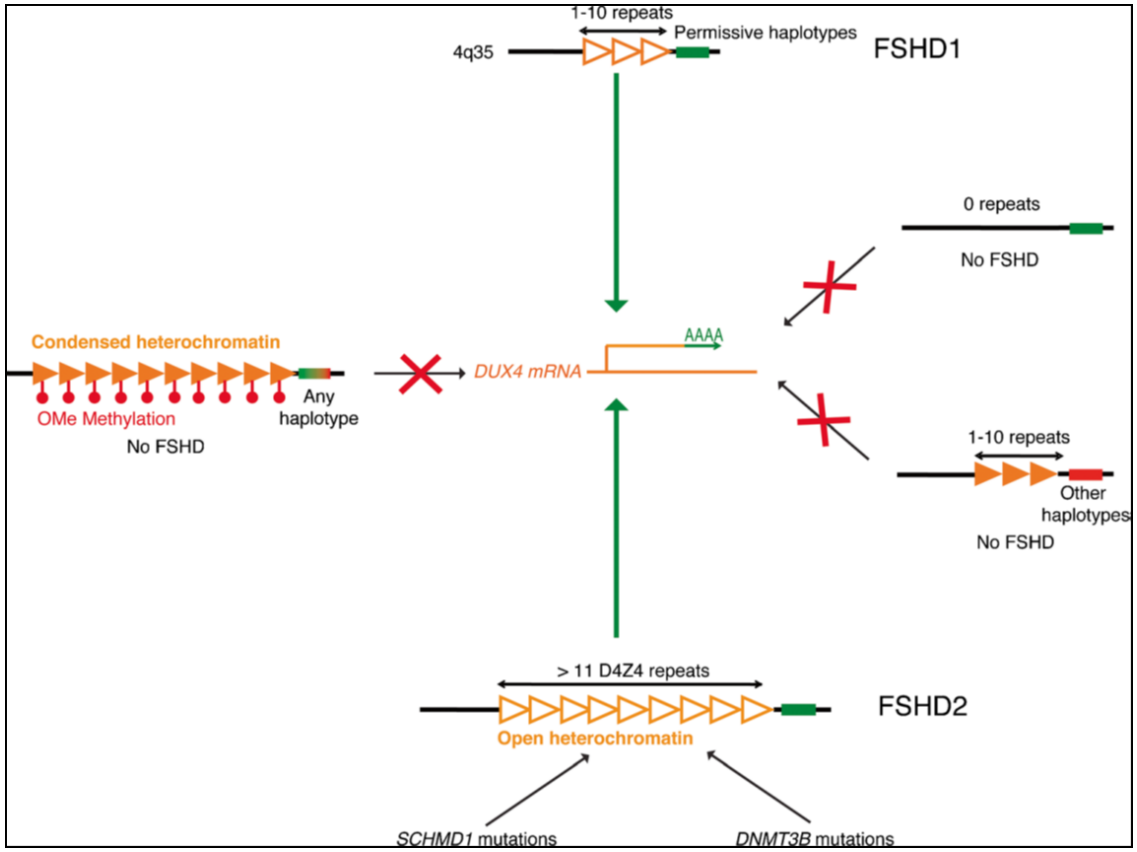
- For **FSHD2**, a second genetic variant is



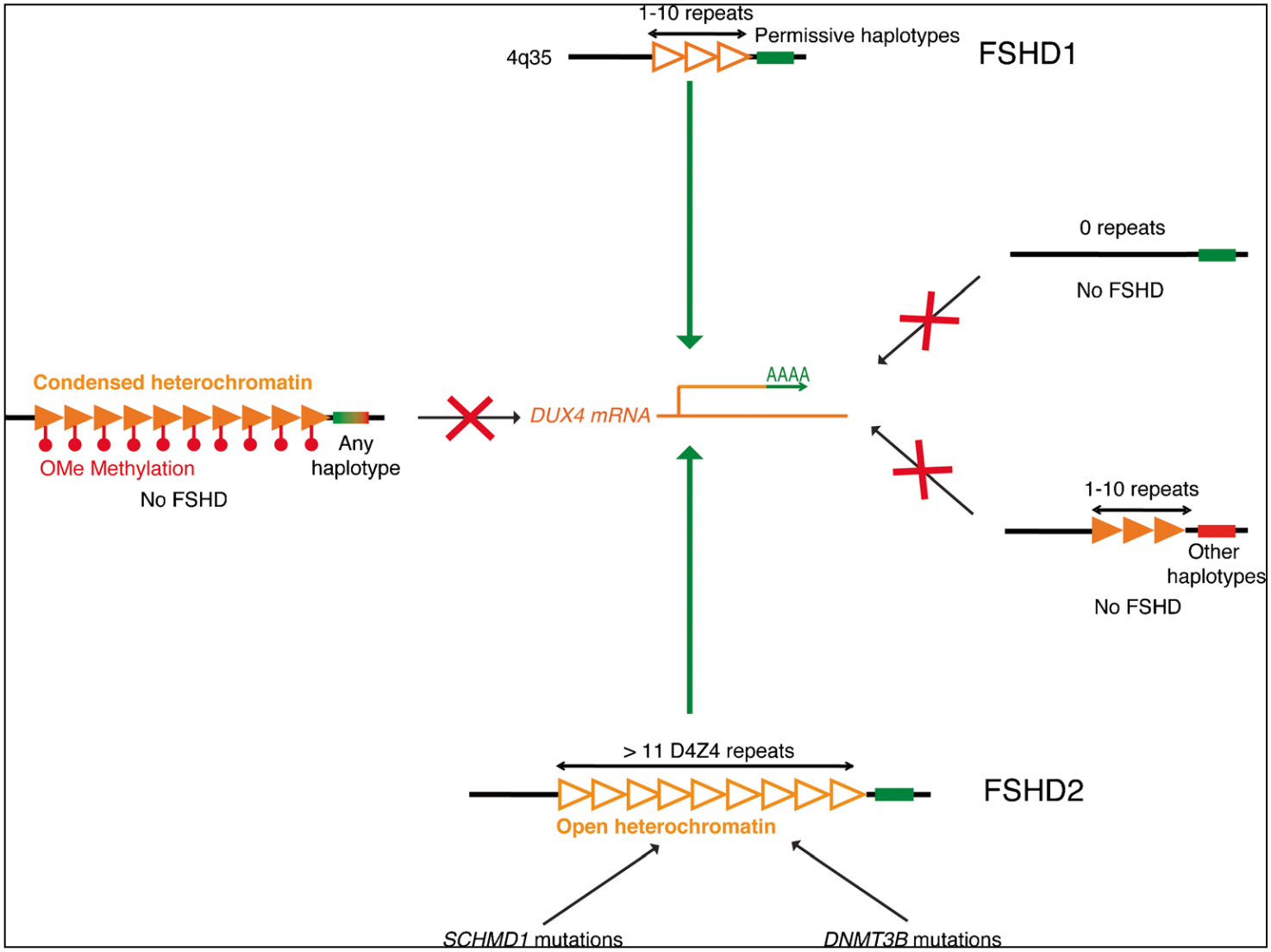
L.H.Wang, Neurol Neurosci Rep (2016) 16: 66

FSHD, history

- **In 1990**, the gene locus responsible for FSHD was mapped to 4q35.
- **In 1992**, contraction in the number of macrosatellite repeats at 4q35 was identified as the genetic defect.
- **In 2010**, the DUX4 gene, a copy of which is located in each macrosatellite repeat D4Z4, was found to be aberrantly expressed.
- **In 2012 and 2016**, mutations in two other genes (*SMCHD1, DNMT3B*) on different chromosomes resulting in decreased heterochromatin condensation and subsequent DUX4 de-repression were described in small percentage of FSHD families (5%) with a normal D4Z4 repeat size.
- **The molecular mechanism leading to FSHD is complex and not yet adequately understood.**



L.H.Wang, Neurol Neurosci Rep (2016) 16: 66



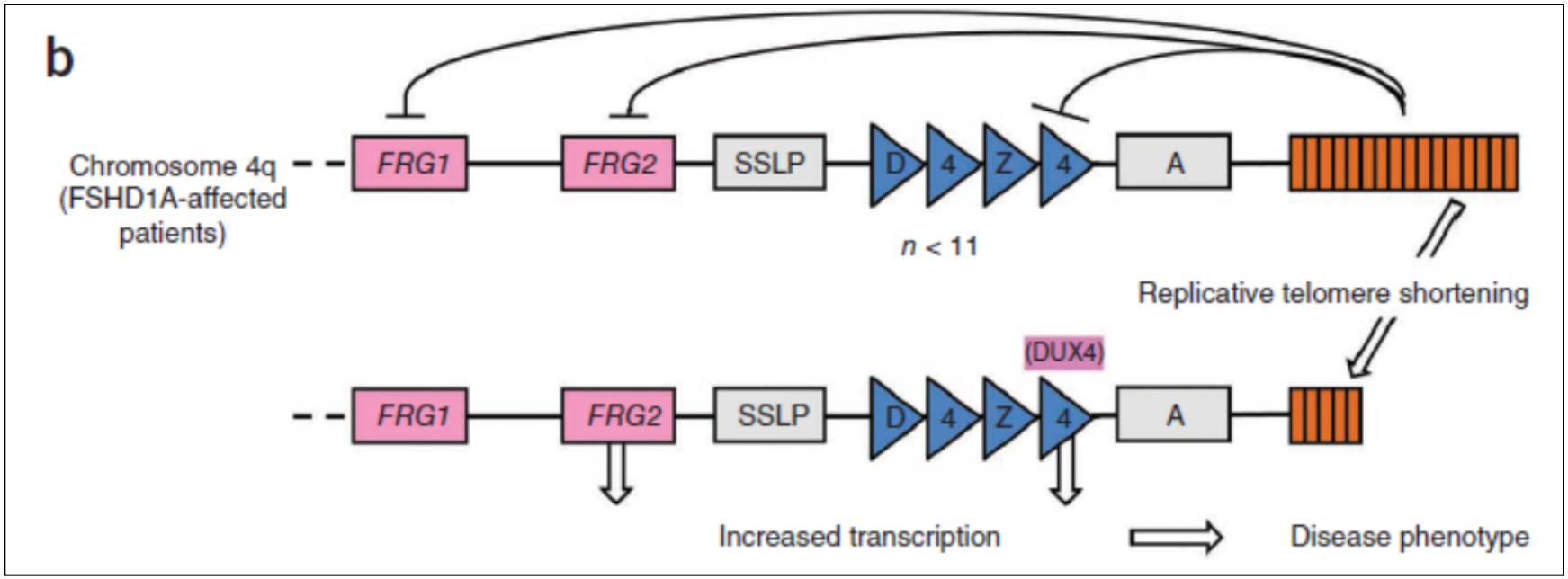
L.H.Wang, Neurol Neurosci Rep (2016) 16: 66

The central component of FSHD molecular pathogenesis is the de-repression of the DUX4 transcription factor, which is present in each D4Z4 repeat (orange triangle). At least one D4Z4 repeat must be present. FSHD is only manifested in chromosomes carrying a haplotype that contains a functional polyadenylation site (green rectangle) just distal to the last D4Z4 repeat. Decreased DNA methylation (red circles) is associated with a more open chromatin structure (open orange triangle). These factors allow for transcription of the DUX4 gene and stabilization of DUX4 mRNA and ultimately expression of the transcription factor that is normally repressed.

FSHD, Telomere position effect

- Marked telomere shortening occurs in myoblasts between the neonatal and adult growth stages (in patients with muscular dystrophy more extensive from the reason of pathological degeneration and regeneration of skeletal muscles).
- Telomere position effects - reversible silencing of genes near telomeres.

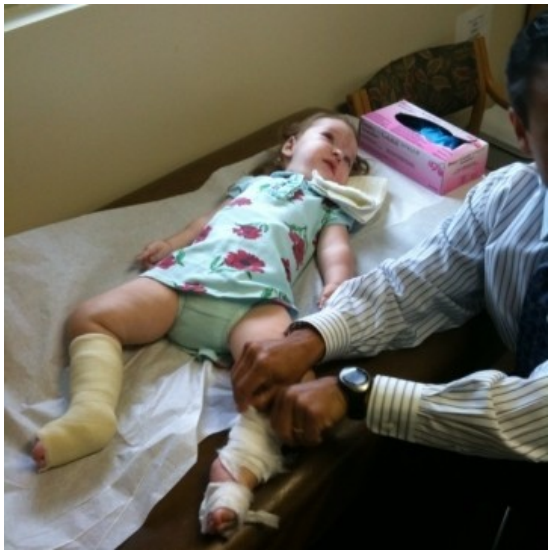
- ***DUX4* expression is upregulated by telomere shortening in cells from patients, in some cases by 100-fold between cells with long and short telomeres. The effect is progressive with decreasing telomere length.**

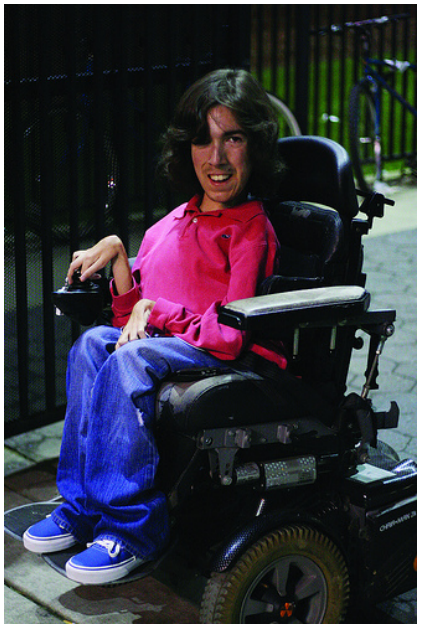


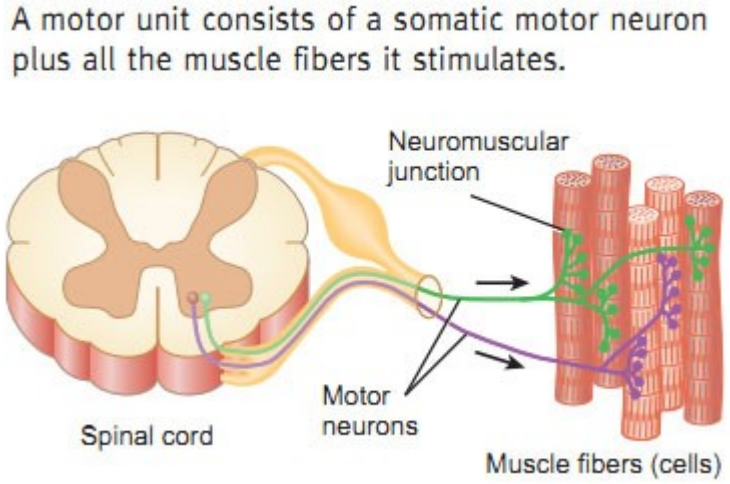
Stadler G., Nature structural & molecular biology 2013

Spinal muscular atrophy (SMA)

- Autosomal recesive disease
- The prevalence: 1: 6 000 – 10 000
- The second most frequent fatal disease with autosomal recesive inheritance (after cystic fibrosis)
- Characterised by degeneration of alpha-motor neurons

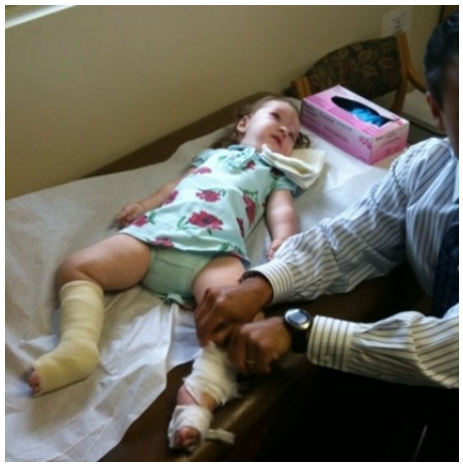




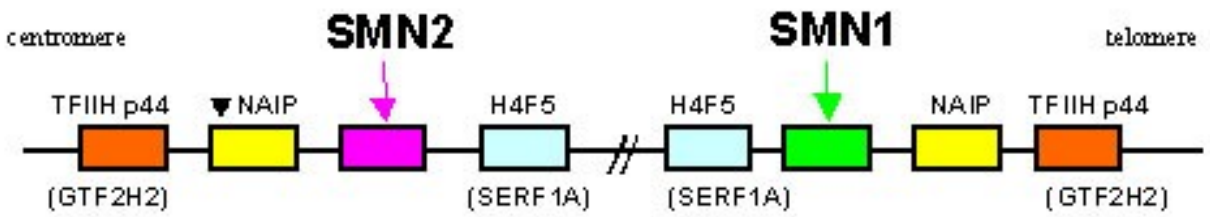


4 clinical groups on the basis of age of onset and clinical course:

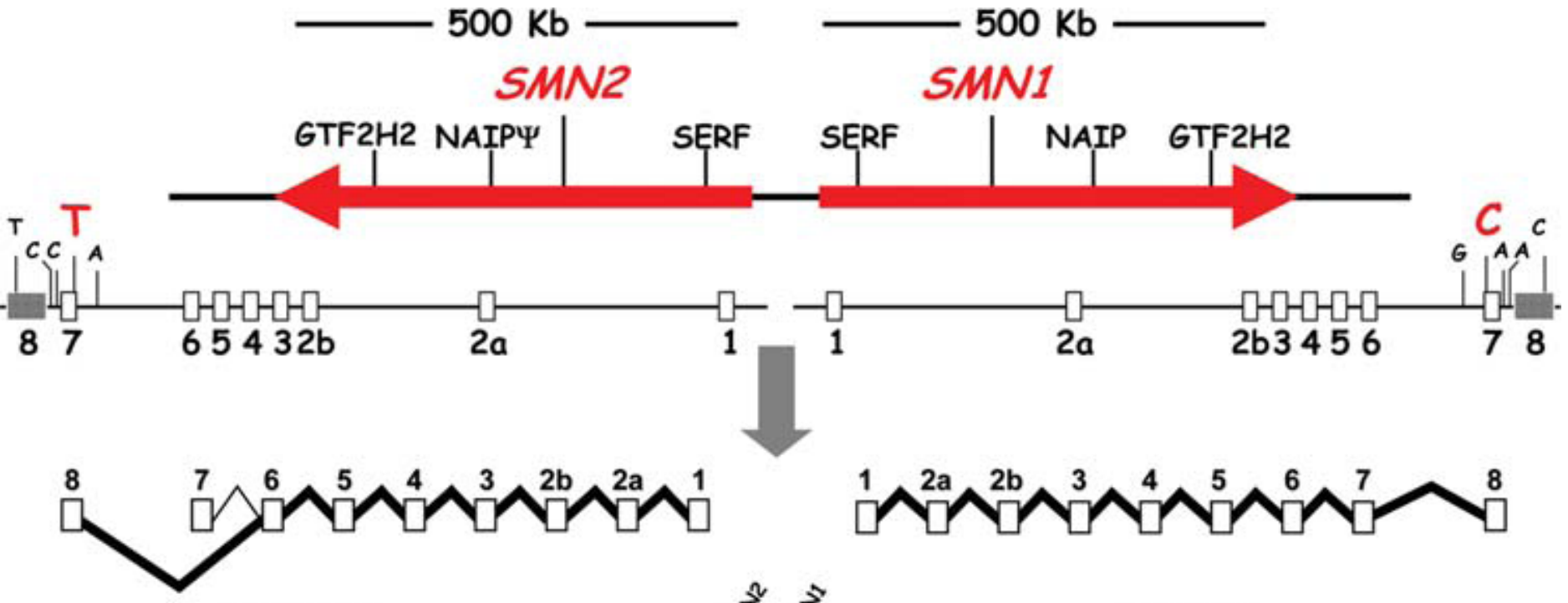
- **Type I (Werdnig-Hoffmann)**, characterized by severe, generalized muscle weakness and hypotonia at birth or within the first three months. Death from respiratory failure usually occurs within the first two years.
- **Type II (intermediate form),** with clinical manifestation starting 6–18 months after birth and life expectancy of 2–30 years; children with type II SMA are able to sit, although they cannot stand or walk unaided.
- **Type III (Kugelberg-Welander disease),** first impacts are typically observed after the second year of life; patients often get wheelchair-bound within or after adolescence.
- **Type IV**, symptoms emerging during adulthood.



- **More than 95% of all SMA cases are due to a homozygous deletion of the *SMN1* gene (Survival of Motor Neuron 1)** located on chromosome region 5q12-5q13.38. This region contains a 500-kb inverted duplication.
- **The *SMN1* and *SMN2* genes** differ in only a few nucleotides (none of which affect the encoded protein sequence).
- The number of *SMN1* and *SMN2* genes may be different.



SMA, the *SMN1* and *SMN2* genes



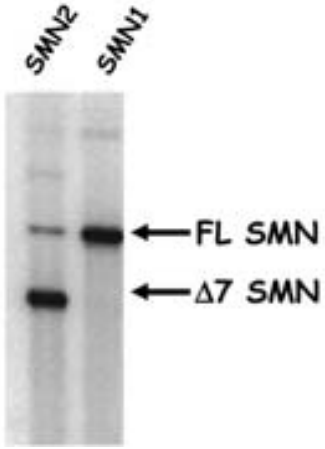
SMN2

Low full-length levels

Alternatively spliced product
~85% Δ7 (lacks exon 7)

Coding sequence identical to SMN1

Only present in humans



SMN1

~100 % full-length SMN

***SMN1*: TTC (Phe)**

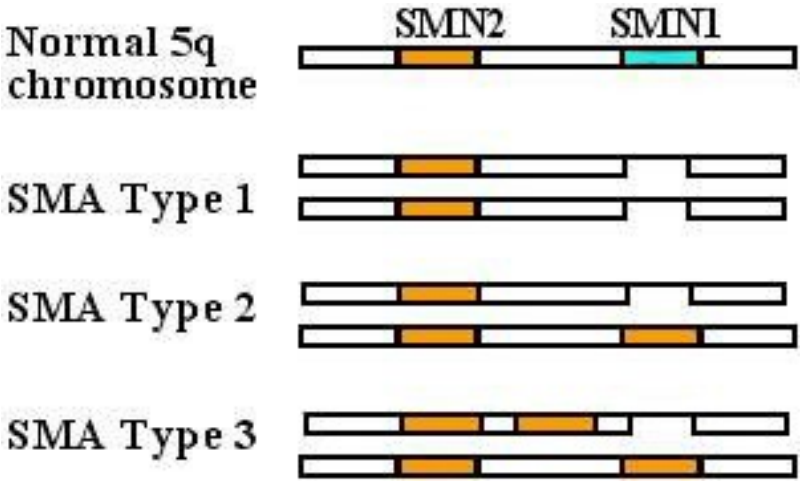
***SMN2*: TTT (Phe)**

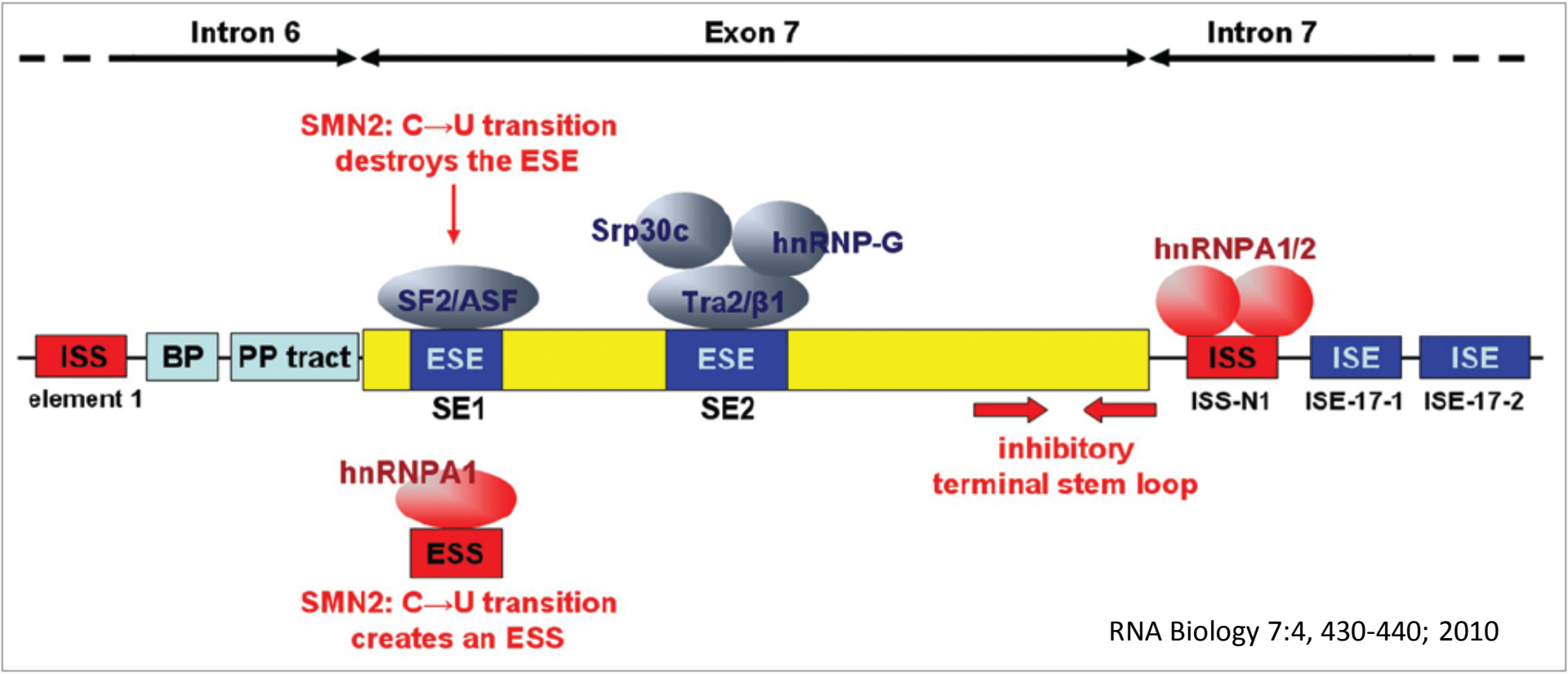
Human Molecular Genetics, 2010, Vol. 19

Schematic of the human SMN locus. The human SMN genes, SMN1 and SMN2, are located in close proximity on chromosome 5. The SMN2 locus is likely derived from a recent duplication event of a genomic region spanning 500 kb which contains additional genes and microsatellite markers. The SMN genes comprise nine exons and eight introns and encode an identical protein product. A silent C–T transition in exon 7 of SMN2 alters a critical exonic splice enhancer and results in a strong reduction of exon 7 inclusion during splicing. Consequently, 85% of the mature mRNA lacks exon 7 (D7), highlighted by the RT–PCR in the bottom panel. The truncated protein is defective in SMN self-association and is degraded rapidly.

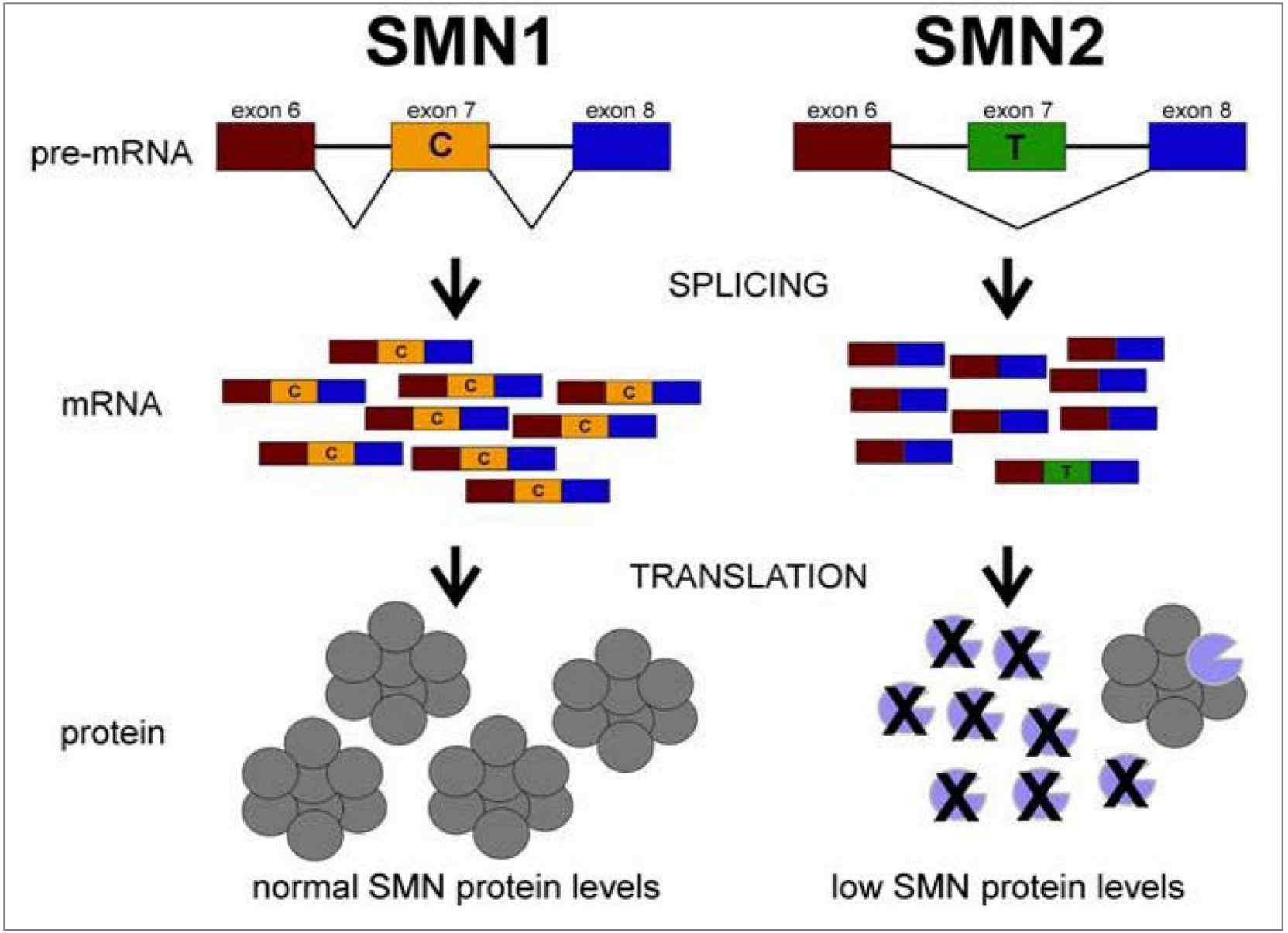
Human Molecular Genetics, 2010, Vol. 19, Review Issue 1, R111–R118

- No phenotype-genotype correlations were initially observed - the homozygous deletion of *SMN1* is present in 95% of SMA patients, independent of the SMA type (I-IV).
- **The *SMN2* copy number modifies the severity of disease.**
- The majority of patients with SMA I have one or two copies of *SMN2*; most patients with SMA II have three *SMN2* copies; and most patients with SMA III have three or four *SMN2* copies.





Splicing architecture of exon 7 of the human *SMN1* and *SMN2* genes.
The diagram represents exon 7 (yellow box) and its flanking intronic regions (lines). Elements inhibiting exon 7 inclusion are shown in red, whereas the positive elements are represented in dark blue. The branch point (BP) and polypyrimidine tract (PP tract) are indicated in light blue. **SF2/ASF and Tra2/β1 bind to the exonic splicing enhancers SE1 and SE2, respectively. The recognition of SE1 by SF2/ASF is prevented in *SMN2*, due to the C→U transition. This sequence alteration also creates a heterogeneous nuclear RNPA1-dependent splicing silencer. Exon 7 is extremely short (only 54 bp).**



SMN2 - the loss of amino acids that are encoded by exon 7 results in the production of SMN protein with severely decreased oligomerization efficiency and stability. The SMN monomers are rapidly degraded. Thus, loss of *SMN1* results in reduction of SMN levels in most tissues.

Nat Rev Neurosci. 2009 August ; 10(8): 597–609.

SMN1 and SMN2 genes: structure and splicing

The SMN1 and SMN2 have identical gene structure and are 99.9% identical at the sequence level. The essential difference between the two genes is a single nucleotide change in exon 7 (C or T as indicated). This single nucleotide change affects the splicing of the gene. Thus the majority of SMN transcripts from SMN2 lack exon 7 whereas those from SMN1 contain exon 7. However, because SMN2 does produce some full-length SMN it can be viewed as a gene with reduced function but not loss of function. The loss of amino acids that

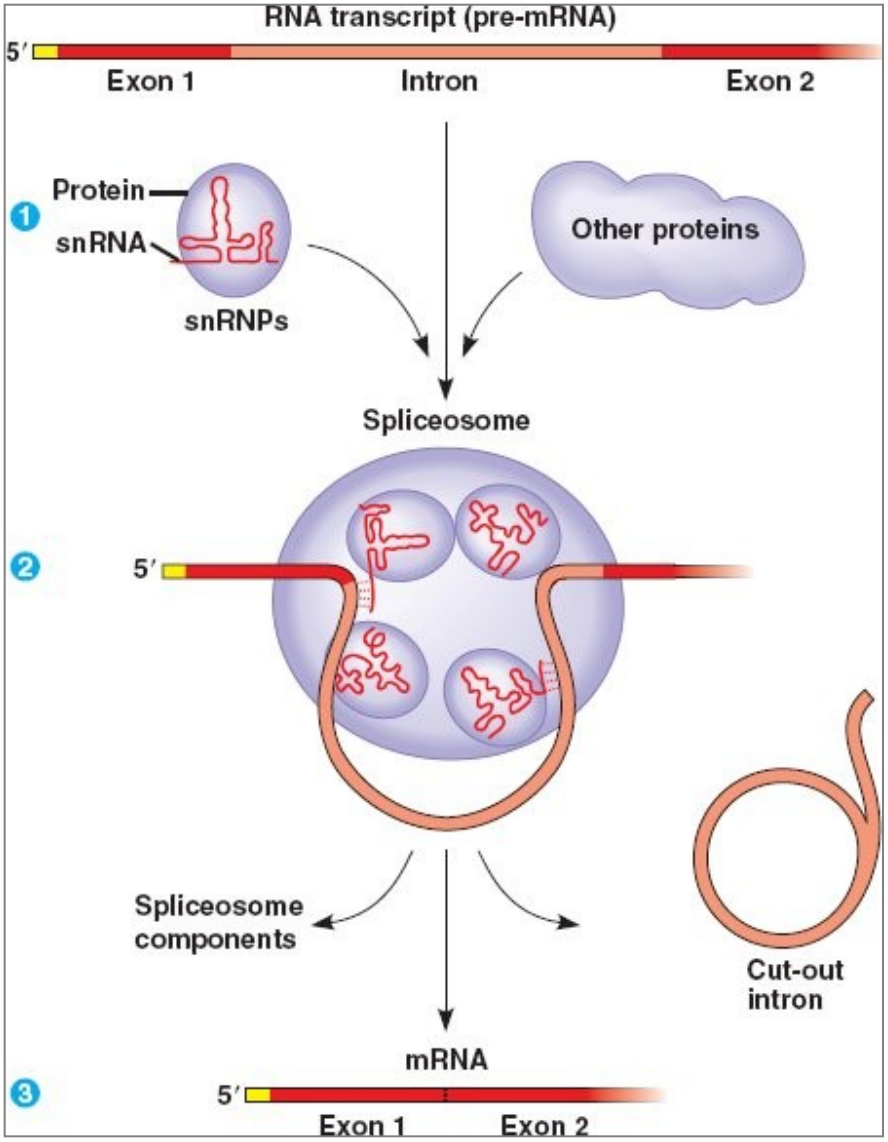
are encoded by exon 7 results in the production of SMN protein with severely decreased oligomerization efficiency and stability. The SMN monomers are rapidly degraded. Thus, loss of SMN1 results in reduction of SMN levels in most tissues. The SMN oligomer is represented as an octomer based on gel filtration of SMN complexes formed in vitro.

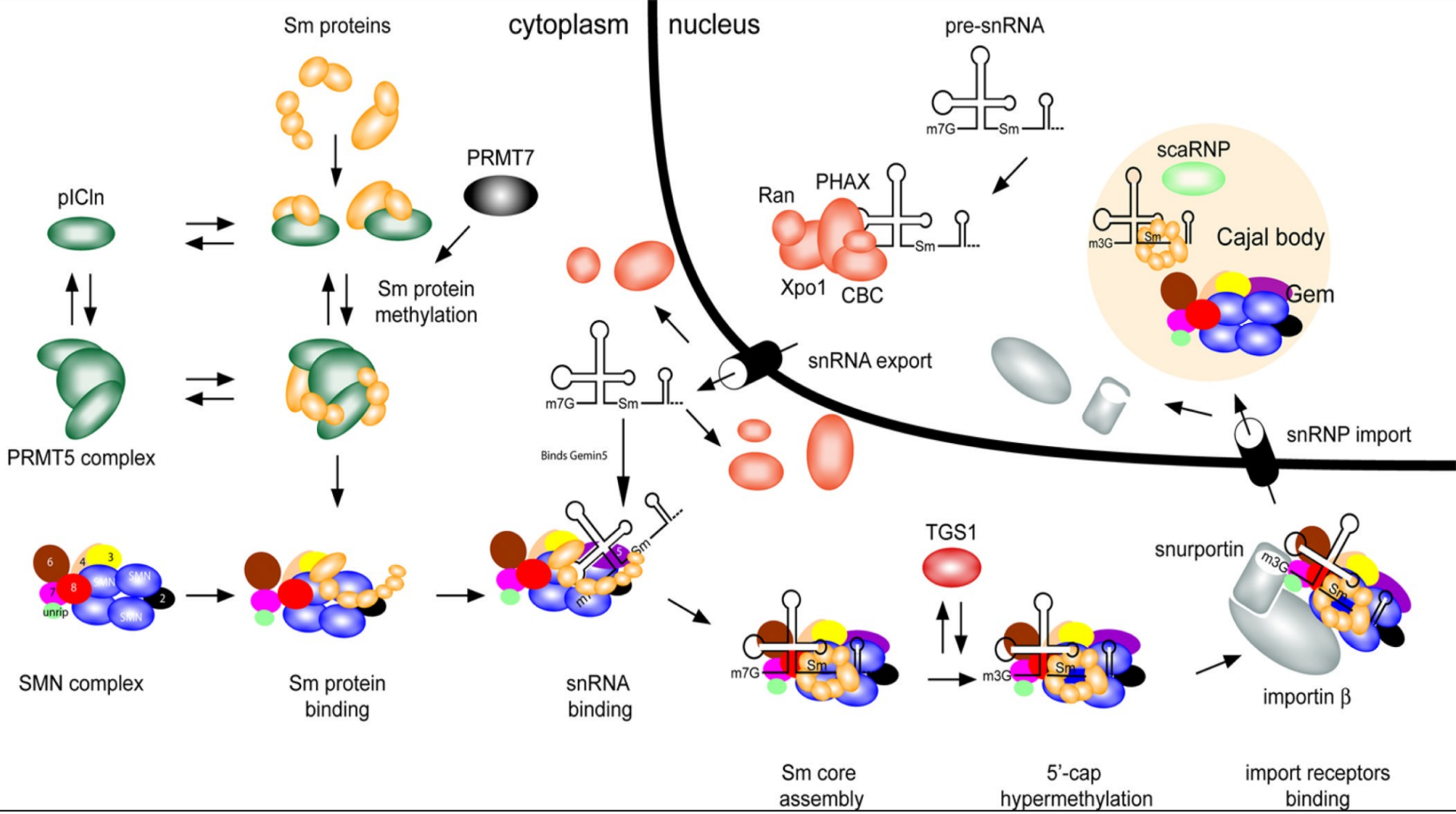
Nat Rev Neurosci. 2009 August ; 10(8): 597–609.

- **The SMN protein has an essential function involving production of small nuclear ribonucleoproteins (snRNPs).**

- Each snRNP particle is composed of small nuclear RNA (snRNA) of approximately 150 nucleotides, several Sm proteins and a number of specific proteins that are unique for each snRNP.

- snRNPs are active in recognizing and removing introns from *pre*-mRNA in the nucleus.





SMN function in snRNP assembly: In the cytoplasm: Sm proteins bind to the SMN complex. The SMN complex is composed of SMN, Gemins 2-8 and Unrip. The snRNA is transcribed in the nucleus and then transported to the cytoplasm. The SMN complex places the Sm proteins onto the snRNA. The 5 cap of the snRNA is hypermethylated, allowing the SMN complex with the snRNA to bind snurportin and importin, which mediates transport of the SMN complex with an assembled snRNP into the nucleus. In the nucleus, the SMN complex and snRNPs localize to the Cajal body and snRNPs undergo further maturation.

Function of SMN in snRNP assembly

Small nuclear ribonucleoproteins (snRNPs) are active in recognizing and removing introns from pre-mRNA in the nucleus. Each snRNP particle is composed of small nuclear RNA (snRNA) of approximately 150 nucleotides, several Sm proteins and a number of specific proteins that are unique for each snRNP. Survival motor neuron (SMN) functions in the cytoplasm to assemble Sm proteins onto the snRNAs to produce an active snRNP.

A) In the cytoplasm the 7 Sm proteins bind to the chloride conductance regulatory protein (pICln). In vitro studies reveal that pICIn first binds the Sm proteins as two separate complexes: SmB, SmD3, and SmD1, SmD2. The latter subsequently binds SmE, SmF and SmG44 The protein arginine methyltrasferase (PRMT5 complex) and PRMT7 methylate the Sm proteins SmB, SmD1 and SmD3. Sm proteins are released from pICln-PRMT5 complex and bind the SMN complex.

B1) The SMN complex is composed of SMN, Gemins2-8 and unrip. SMN is shown in the figure as an oligomer as it has been shown to self-associate and it has been suggested that oligomerization is critical for SMN function. The exact numbers of SMN monomers in a SMN complex is unknown (it has been suggested to be an octomer). The Gemins are shown as single units for simplicity as the exact stoichiometry of the SMN complex has not been determined.

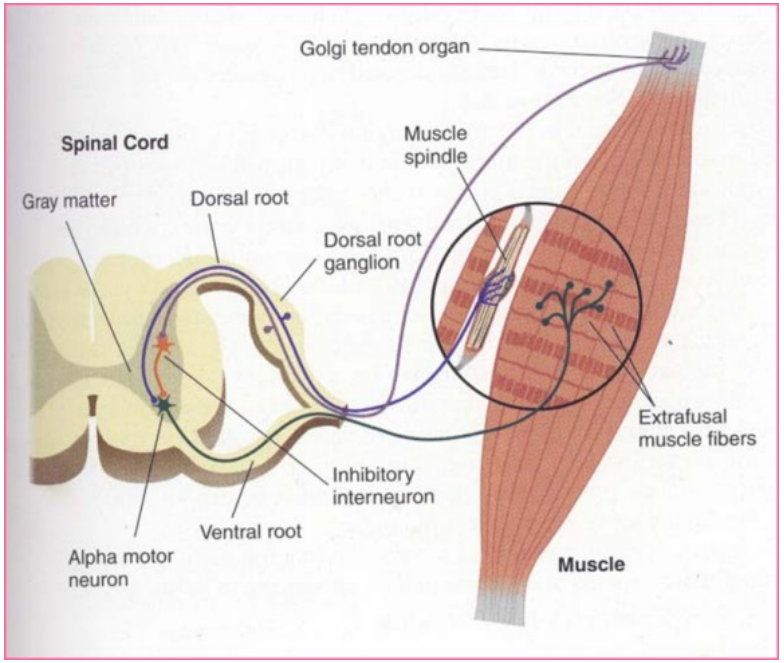
B2) snRNA is transcribed in the nucleus and then binds the export proteins phosphorylated adaptor for RNA export (PHAX), Cap-binding complex (CBC), exportin (Xpo1) and rasrelated nuclear protein GTP (Ran), which transport it to the cytoplasm. In vertebrates, the snRNA is brought into the Sm protein-bound SMN complex by binding to Gemin5.

C) The SMN complex places the Sm proteins onto the snRNA. The m7G cap of the snRNA is hypermethylated by trimethylguanosine sythetase 1 (TGS), allowing the SMN complex with the snRNA to bind snurportin and importin, which mediates transport of the SMN complex with an assembled snRNP into the nucleus.

D) In the nucleus the SMN complex and snRNPs localize to the Cajal body and snRNPs undergo further maturation. Depending on the cell type and developmental stage, SMN can localize as a separate body adjacent to the Cajal body.

Nat Rev Neurosci. 2009 August ; 10(8): 597–609.

SMA is characterised by the degeneration of alpha motor neurons of the anterior horns of the spinal cord - <u>not all cell types are equally affected in SMA patients.</u>

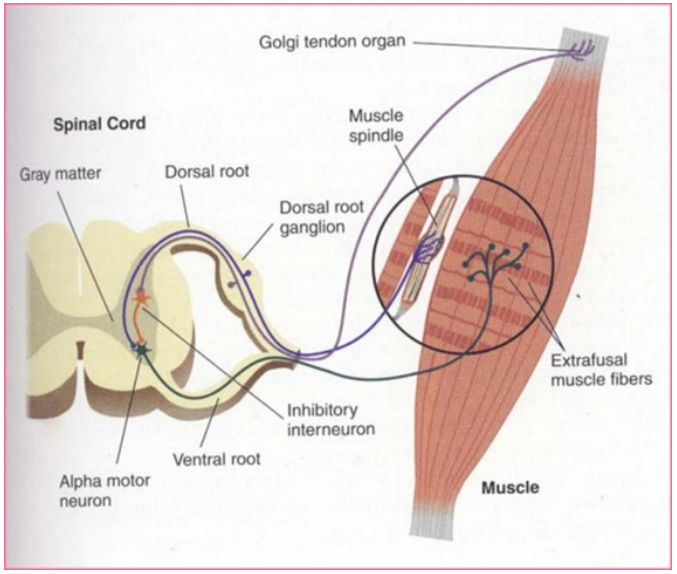


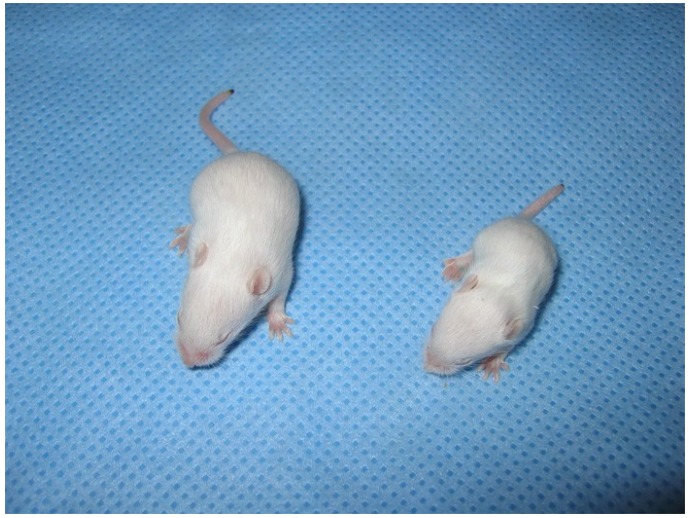
- **The vast majority of functional SMN protein is produced by *SMN1* and low level by *SMN2* (poor inclusion of exon 7).**

- **SMA (homozygous deletion of *SMN1*): *SMN2* is the only source of functional SMN protein, the degree of exon 7 inclusion becomes critical.**

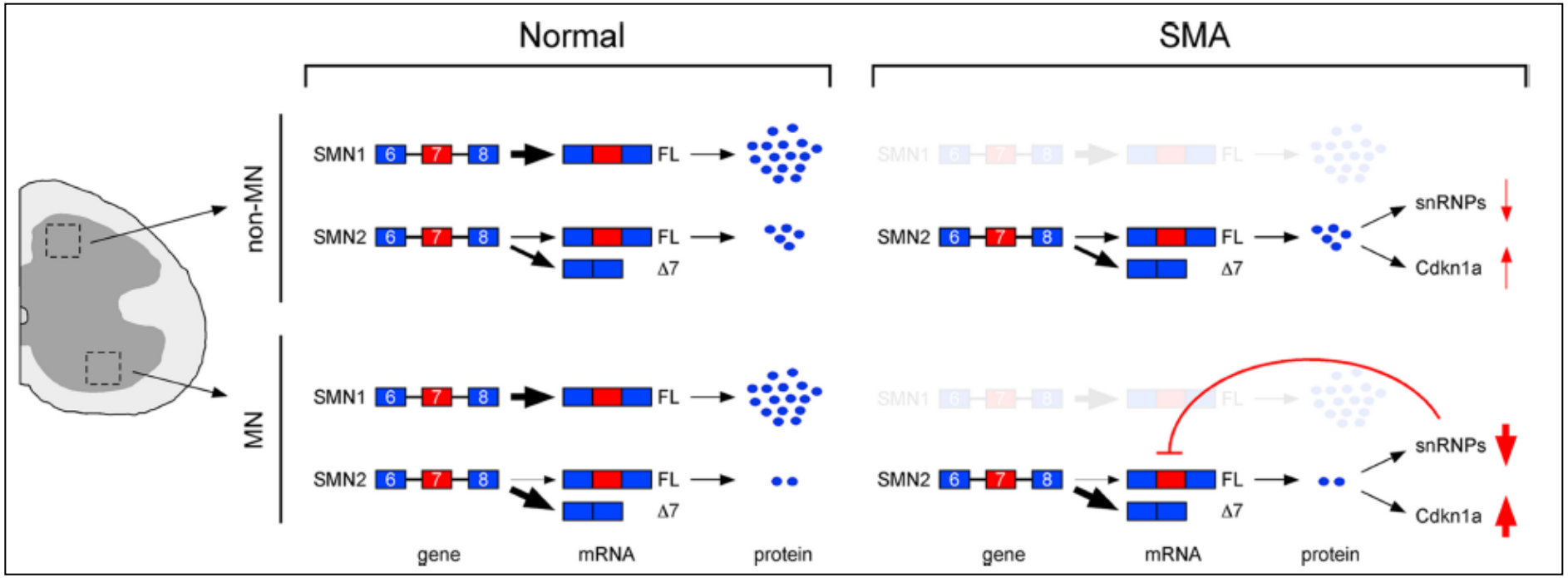
Mouse models:

- **Motor neurons express 4-fold less full-length *SMN2* mRNA than dorsal horn cells from the same spinal segment.** This difference is due to changes not in *SMN2* gene transcription but in **the efficiency of exon 7 inclusion in the *SMN2* mRNA**.
- **SMA (the loss of *SMN1*): 4-fold difference in the levels of full length *SMN2* mRNA is critical for motor neuron function. Studies showed that <u>two copies of the *SMN2* gene produce enough protein for normal function of most cells except motor neurons</u>.**





Model for the selective vulnerability of motor neurons in SMA.
Under normal conditions, the *SMN2* gene produces smaller amounts of the full-length *SMN* mRNA and protein in motor neurons (MN) than in non-motor neurons (non-MN).
Upon loss of the *SMN1* gene (SMA), reduced snRNP levels trigger a negative feedback loop affecting exon 7 splicing that might contribute to further decreasing SMN expression and enhancing downstream defects specifically in SMA MNs.

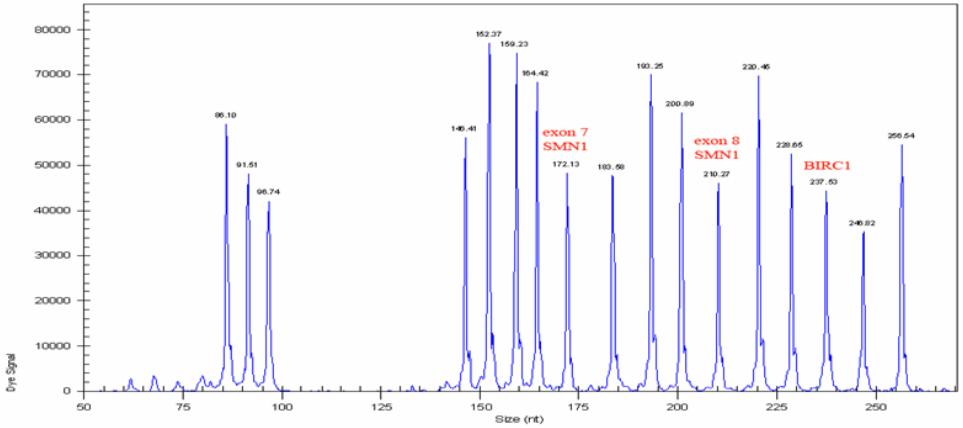


Ruggiu M., Molecular and Cellular Biology 2012

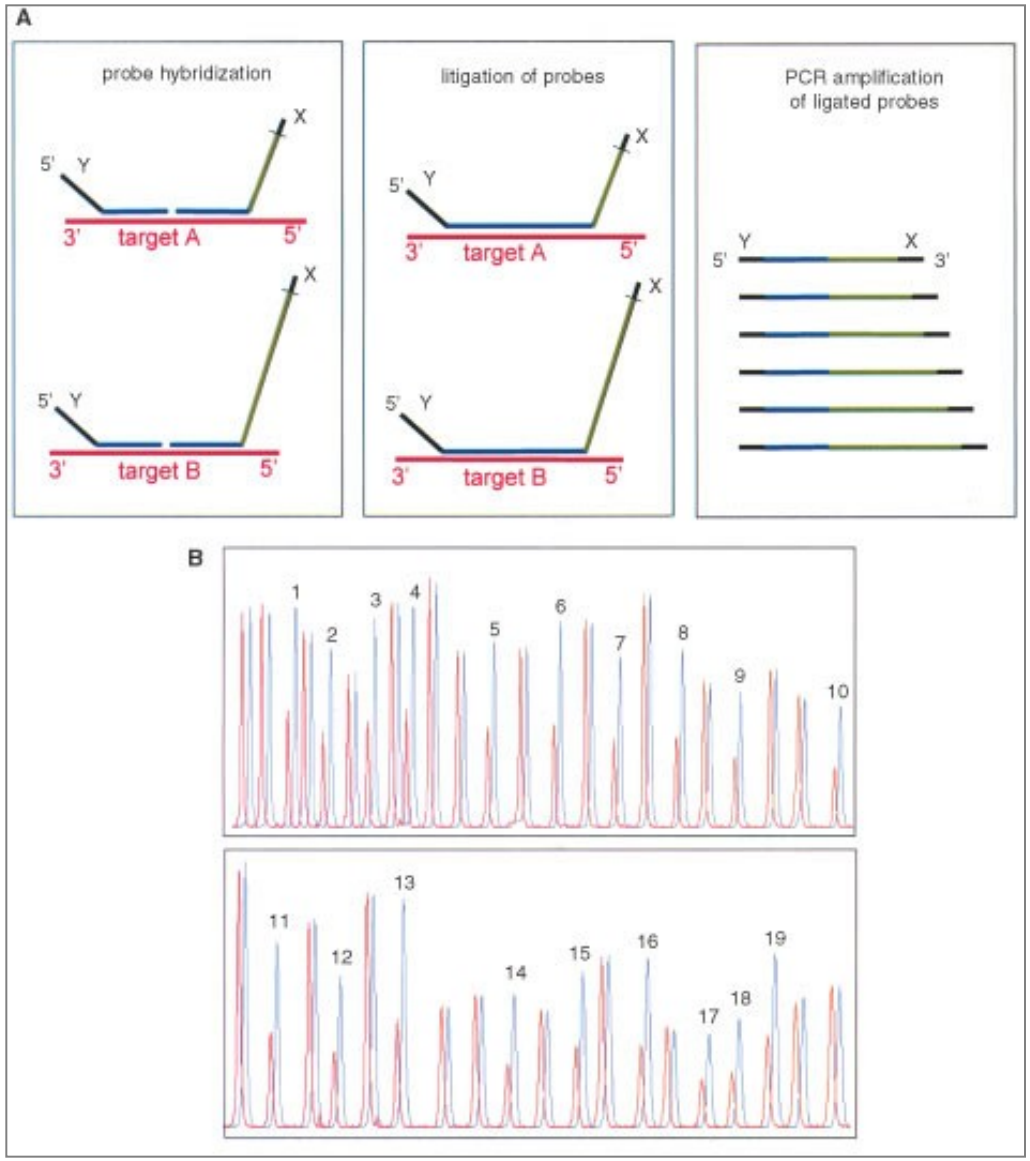
- 95% způsobeno homozygotní delecí *SMN1* genu
- 5% způsobeno delecí *SMN1* genu na jednom chromozomu a bodovou mutací na druhém.

MLPA (Multiple ligation dependent probe amplification)

- stanovení počtu kopií *SMN1* genu na genom,
- stanovení přenašečství SMA v rodinách s výskytem SMA (1 kopie - přenašeč, 2 kopie - zdravý),
- vytipování pacientů u kterých se bude provádět sekvenční analýza *SMN1* genu (1 kopie *SMN1*).



Multiplex Ligation-dependent Probe Amplification (MLPA)



Denatured genomic DNA is hybridised with a mixture of probes. Each MLPA probe consists of two oligonucleotides. The two parts of each probe hybridise to adjacent target sequences and are ligated by a thermostable ligase. All probe ligation products are amplified simultaneously by PCR using a single primer pair labeled with a fluorescence mark. The amplification product of each probe has a unique length. Amplification products are separated by capillary electrophoresis. Relative amounts of probe amplification products reflect the relative copy number of target sequences.

The SMA probe mix contains 18 different control probes as well as exon 7 and 8 probes specific for SMN1 and SMN2.