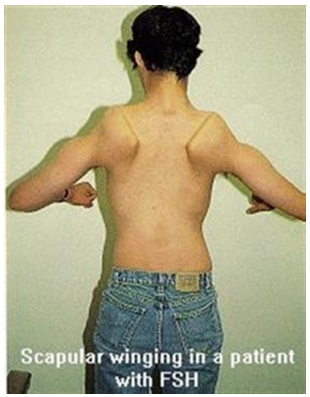
University Hospital Brno

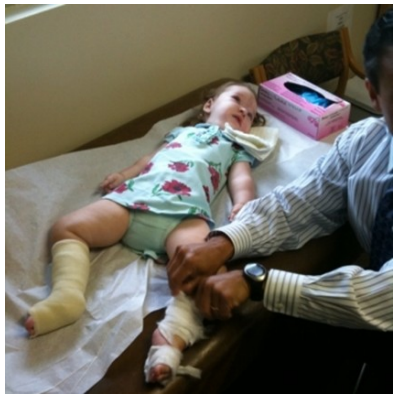
Centre of molecular biology and genetics

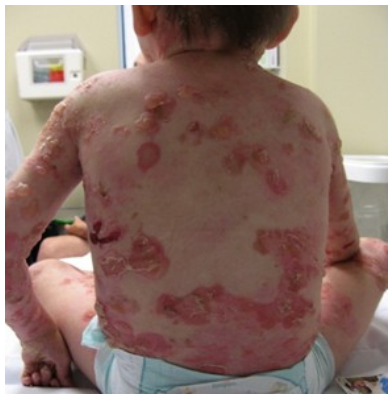
Department of inherited genetic disorders

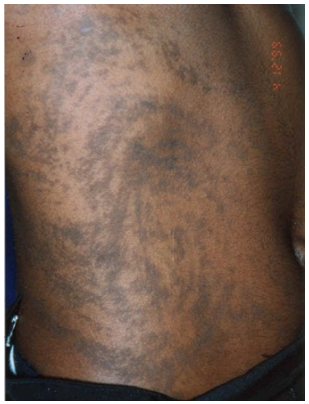
Molecular genetic diagnostics of monogenic diseases

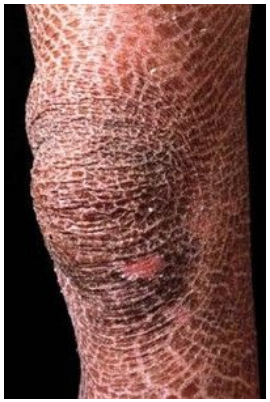
- Neuromuscular diseases
- Epilepsies
- Skin diseases
- Connective tissue diseases
- Metabolic diseases





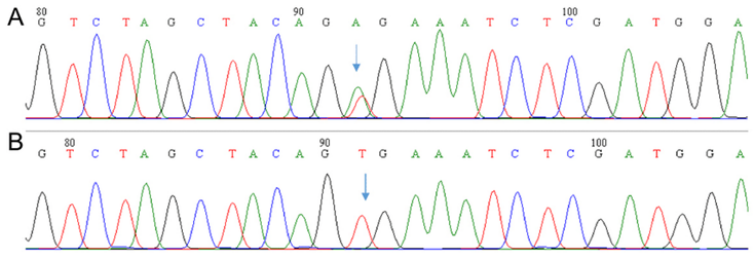


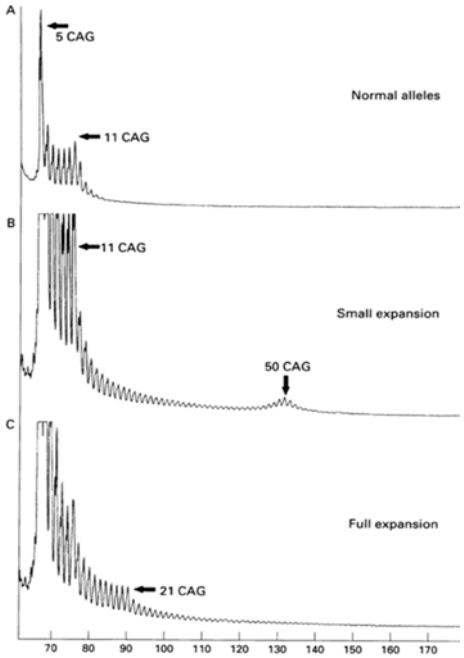


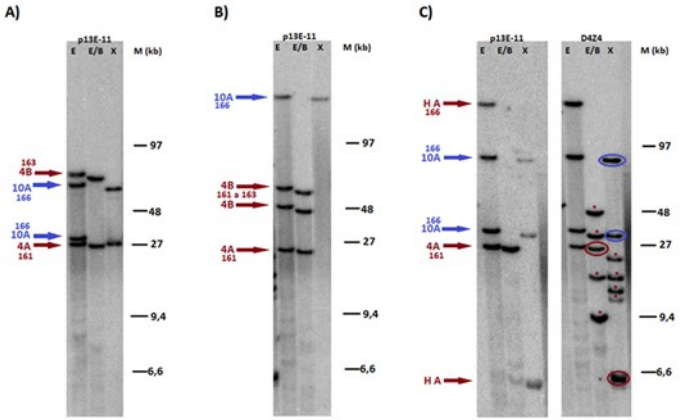


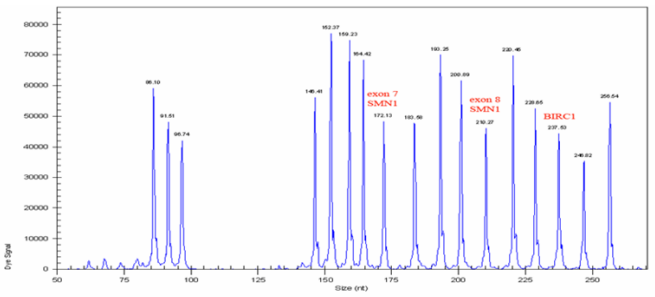
Methodological approaches used

- **Classic sequencing** (identification of small scale variants)
- **Next generation sequencig** (identification of small scale variants, large deletions/duplications)
- **Multiple ligation dependent probe amplification** (identification of large deletions/duplications)
- **Repeat-primed PCR** (identification of repeat expansions)
- **Southern blot and hybridisation** (identification of repeat expansions/deletions)



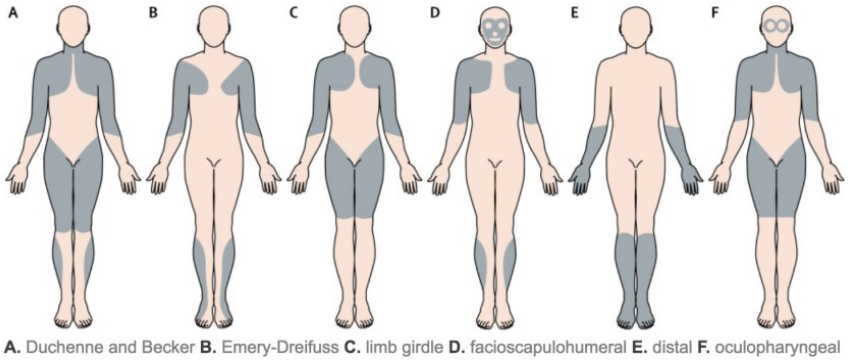






Two monogenic diseases for presentation of different approaches in molecular genetic diagnostics

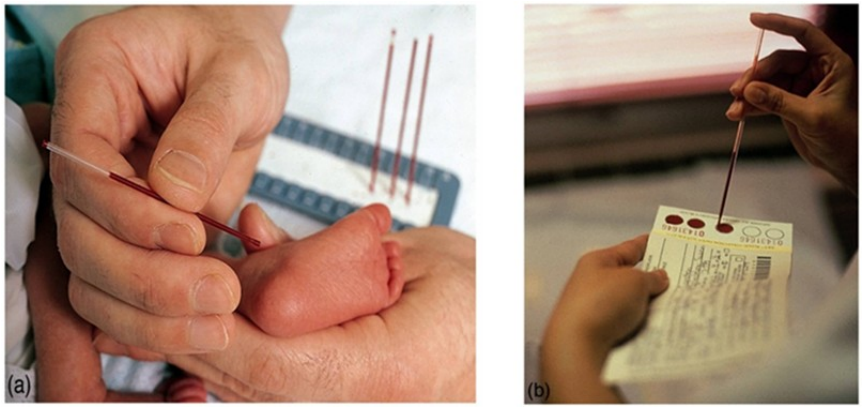
- **Phenylketonuria** – the disease in which the gene for molecular genetic diagnostics is determined on the basis of biochemical findings
- **Muscular dystrophy** – the disease encompasing different types of muscular disorders; biochemical and other findings are mostly not specific enough for detemination of a specific types and so a gene for molecular genetic diagnostics



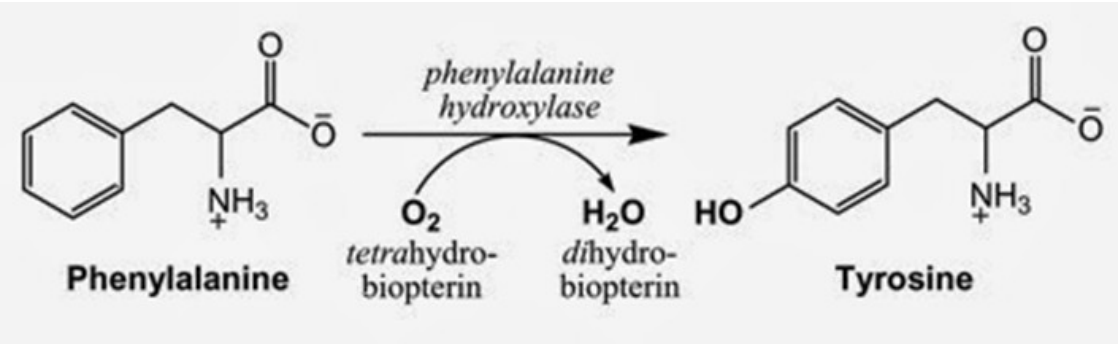
A. Duchenne and Becker **B.** Emery-Dreifuss **C.** limb girdle **D.** facioscapulohumeral **E.** distal **F.** oculopharyngeal

Phenylketonuria (PKU)

- Diagnosed by newborn screening, on the basis of increased phenylalanine and the ratio phenylalanine/tyrosine



- Increased Phe and Phe/Tyr is the indicator for DNA analysis of **the *PAH* gene** encoding **the phenylalaninhydroxylase**



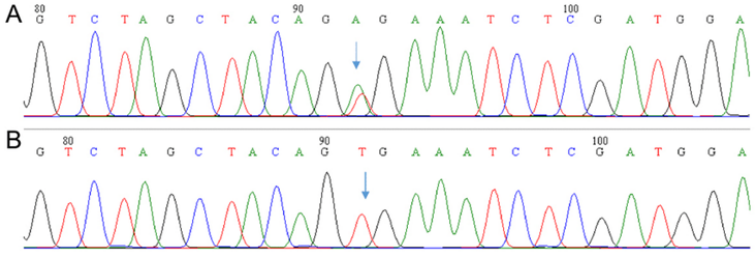
- In 98% of cases, two pathogenic variants in the *PAH* gene are identified.

Muscular dystrophies and myopathies

- To date, 162 genes associated with clinical manifestation of muscular dystrophy/myopathy have been described.
- Clinical, biochemical, pathological,…. findings are mostly not specific enough for selection of a gene for molecular genetic analysis
- **Which gene to analyse?**

How was the question solved before – genes were analysed sequentially by **classical DNA sequencing**, starting with a gene with the most likely mutation occurence, in case of a negative result another gene was selected ….. TIME AND FINANCIALLY CONSUMING and only a certain number of selected genes were analysed.

How is the question solved now – all genes associated with the disease are analysed at the same time (in parallel) by **next generation sequencing (NGS)** … FAST AND RELATIVELY CHEAP

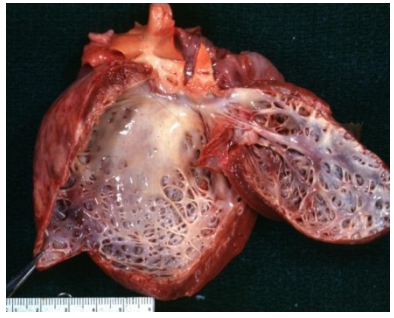


Muscular dystrophies/myopaties are characterised by **clinical heterogeneity**

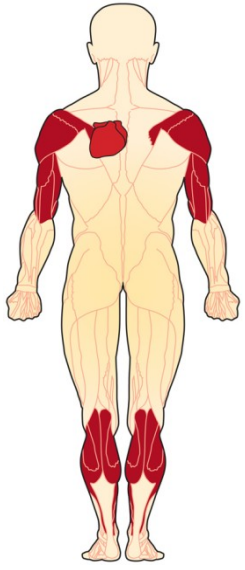
- **Pathogenic sequence variants in 1 gene cause several types of muscle diseases; example: *LMNA*, lamin A/C (protein localised in nuclear envelope)**

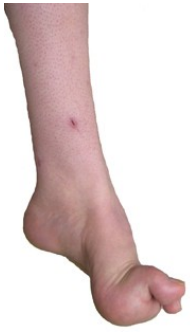
1. Cardiomyopathy, dilated, 1A
2. Charcot-Marie-Tooth disease, type 2B1
3. Emery-Dreifuss muscular dystrophy 2, AD
4. Emery-Dreifuss muscular dystrophy 3, AR
5. Hutchinson-Gilford progeria
6. Heart-hand syndrome, Slovenian type
7. Lipodystrophy, familial partial, 2
8. Malouf syndrome
9. Mandibuloacral dysplasia
10. Muscular dystrophy, congenital
11. Muscular dystrophy, limb-girdle, type 1B
12. Restrictive dermopathy, lethal

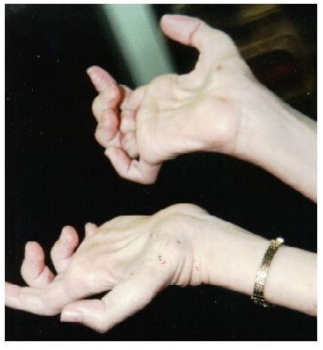
http://www.omim.org/



www.wikidoc.org/index.php/Dilated_cardiomyopathy







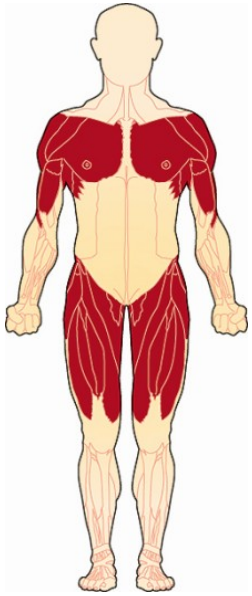
http://dxline.info/diseases/charcot-marie-tooth-disease

http://neuromuscular.wustl.edu/time/hmsn.html

http://scientia1.files.wordpress.com/2013/01/progeria.jpg

Muscular dystrophies/myopaties are characterised by **genetic heterogeneity**

- **One type of disease is caused by pathogenic variants in 1 from several possible genes**



Example: LGMD – 29 genes associated with LGMD have been described so far.

Table 1
Comparison of the previous LGMD nomenclature to the proposed classification system after the definition has been applied to the current list of LGMD. Conditions which are no longer considered LGMDs are highlighted in grey with a reason for their exclusion given.

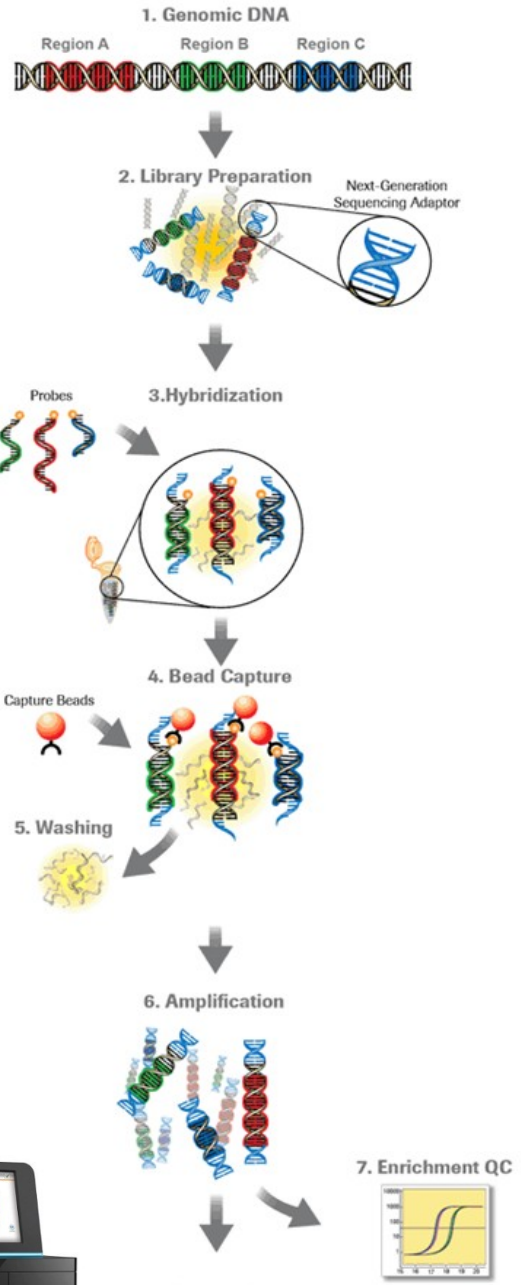
Old name	Gene	Proposed new nomenclature	Reason for exclusion
LGMD 1A	*Myot*	Myofibrillar myopathy	Distal weakness
LGMD 1B	*LMNA*	Emery–Dreifuss muscular dystrophy (EDMD)	High risk of cardiac arrhythmias; EDMD phenotype
LGMD 1C	*CAV3*	Rippling muscle disease	Main clinical features rippling muscle disease and myalgia
LGMD 1D	*DNAJB6*	LGMD D1 DNAJB6-related	
LGMD 1E	*DES*	Myofibrillar myopathy	Primarily false linkage; distal weakness and cardiomyopathy
LGMD 1F	*TNP03*	LGMD D2 TNP03-related	
LGMD 1G	*HNRNPDL*	LGMD D3 HNRNPDL-related	
LGMD 1H	?	Not confirmed	False linkage
LGMD 1I	*CAPN*	LGMD D4 calpain3-related	
LGMD 2A	*CAPN*	LGMD R1 calpain3-related	
LGMD 2B	*DYSF*	LGMD R2 dysferlin-related	
LGMD 2C	*SGCG*	LGMD R5 γ-sarcoglycan-related[a]	
LGMD 2D	*SGCA*	LGMD R3 α-sarcoglycan-related	
LGMD 2E	*SGCB*	LGMD R4 β-sarcoglycan-related	
LGMD 2F	*SGCD*	LGMD R6 δ-sarcoglycan-related	
LGMD 2G	*TCAP*	LGMD R7 telethonin-related	
LGMD 2H	*TRIM32*	LGMD R8 TRIM 32-related	
LGMD 2I	*FKRP*	LGMD R9 FKRP-related	
LGMD 2J	*TTN*	LGMD R10 titin-related	
LGMD 2K	*POMT1*	LGMD R11 POMT1-related	
LGMD 2L	*ANO5*	LGMD R12 anoctamin5-related	
LGMD 2M	*FKTN*	LGMD R13 Fukutin-related	
LGMD 2N	*POMT2*	LGMD R14 POMT2-related	
LGMD 2O	*POMGnT1*	LGMD R15 POMGnT1-related	
LGMD 2P	*DAG1*	LGMD R16 α-dystroglycan-related	
LGMD 2Q	*PLEC*	LGMD R17 plectin-related	
LGMD 2R	*DES*	myofibrillar myopathy	Distal weakness
LGMD 2S	*TRAPPC11*	LGMD R18 TRAPPC11-related	
LGMD 2T	*GMPPB*	LGMD R19 GMPPB-related	
LGMD 2U	*ISPD*	LGMD R20 ISPD-related	
LGMD 2V	*GAA*	Pompe disease	Known disease entity, histological changes
LGMD 2W	*PINCH2*	PINCH-2 related myopathy	Reported in one family
LGMD 2X	*BVES*	BVES related myopathy	Reported in one family
LGMD 2Y	*TOR1AIP1*	TOR1AIP1 related myopathy	Reported in one family
LGMD 2Z	*POGLUT1*	LGMD R21 POGLUT1-related	
Bethlem myopathy recessive	*COL6A1, COL6A2, COL6A3*	LGMD R22 collagen 6-related	
Bethlem myopathy dominant	*COL6A1, COL6A2, COL6A3*	LGMD D5 collagen 6-related	
Laminin α2-related muscular dystrophy	*LAMA2*	LGMD R23 laminin α2-related	
POMGNT2-related muscular dystrophy	*POMGNT2*	LGMD R24 POMGNT2-related	

[a] Sarcoglycan-related LGMDs rationalised based on order of gene discovery.

Molecular genetic diagnostic of muscular dytrophies /myopathies: **targeted NGS** is a rapid and cost-effective way to detect variants in selected sets of genes or gene regions.

Principle:

- DNA samples are converted into sequencing libraries – DNA is randomly sheared into smaller fragments by mechanical or enzymatic methods, adapters for sequencing and multiplexing are added to DNA ends (multiplexing enables analysis of more DNA samples in one sequencing run)
- Regions of interest within the library are captured using biotinylated oligonucleotide probes. These probes are designed to hybridize to regions of interest.
- After hybridization of probes to fragmented DNA, streptavidin-bound magnetic beads are used to separate the probe-targeted fragment complex from other fragments that are not bound to probes.
- Amplification of targeted regions
- NGS of targeted regions



NGS generates a large number of sequence variants → these variants need to be interpreted → each identified variant should be assigned to one of five classes.

1. **Benign sequence variants**
2. **Likely benign sequence varints**
3. **Sequence variants of uncertain significance**
4. **Likely pathogenic sequence variants**
5. **Pathogenic sequence variants**

Genet Med. 2015 May ; 17(5): 405–424. doi:10.1038/gim.2015.30.

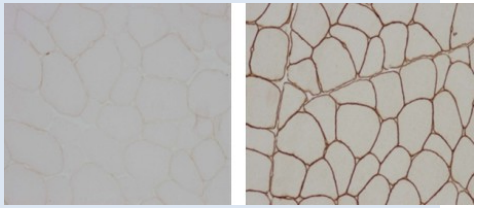
Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

Sue Richards [Chair, ACMG],

Case study: the need to combine DNA and mRNA diagnostics

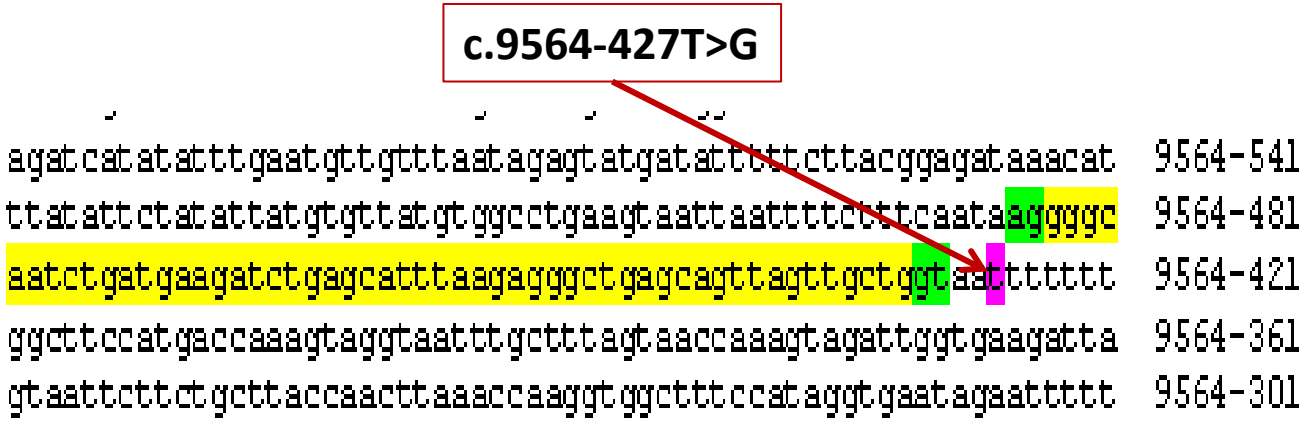
Patient with suspected Duchenne muscular dystrophy (DMD gene, Xp21)

Pathological analysis: dystrophin deficiency in muscle tissue



Molecular genetic analysis:

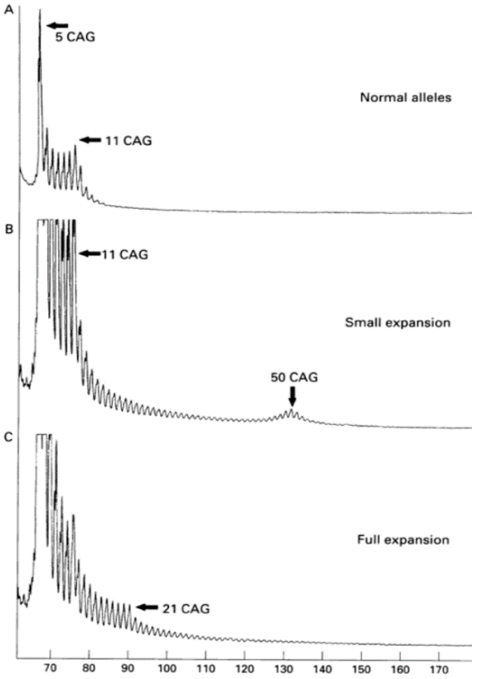
1. **DNA analysis: no pathogenic variants identified**
2. **mRNA analysis of the *DMD* gene: a pathogenic variant identified (the insertion of 53 nt. from intron 65 into mRNA)**
3. **On the basis of mRNA results, DNA analysis of intron 65 performed: pathogenic variant deep inside of the intron 65 identified; this results in insertion of a part of intron sequence into the dystrophin mRNA**

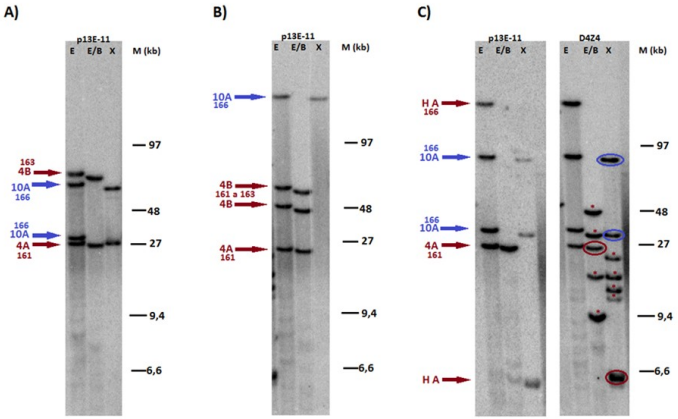


Muscular dystrophies/myopathies – molecular genetic diagnostics

Overall 162 genes

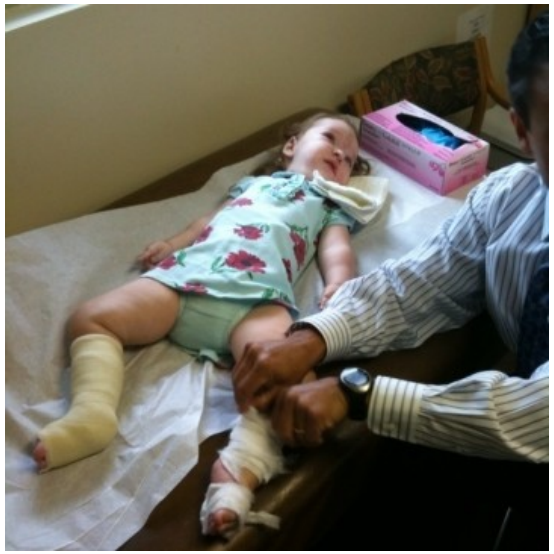
159 genes	2 genes	1 gene
	MD1, MD2	FSHD
➢ **NGS** • Identification of small scale pathogenic variants, large gene deletions / duplications	➢ **Repeat primed-PCR** • Identification of repeat expansions	➢ **Souther blot + hybridisation** • Identification of repeat deletions

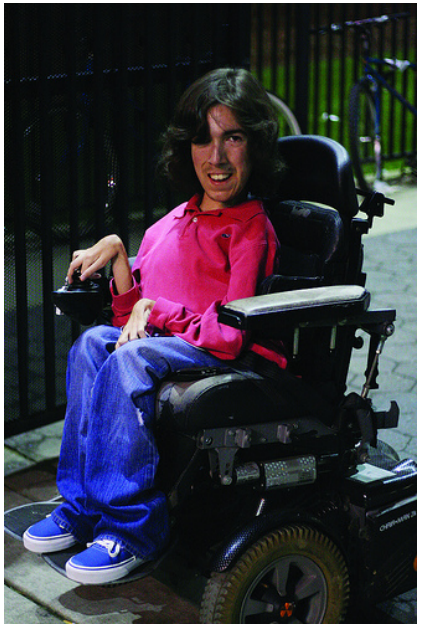


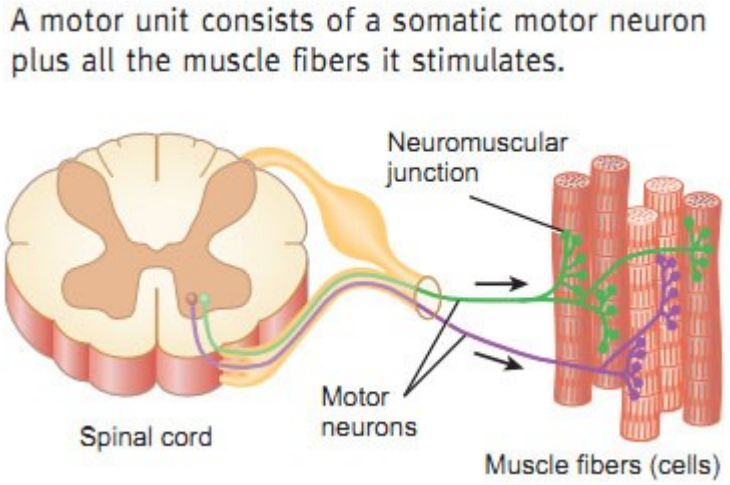


Spinal muscular atrophy (SMA)

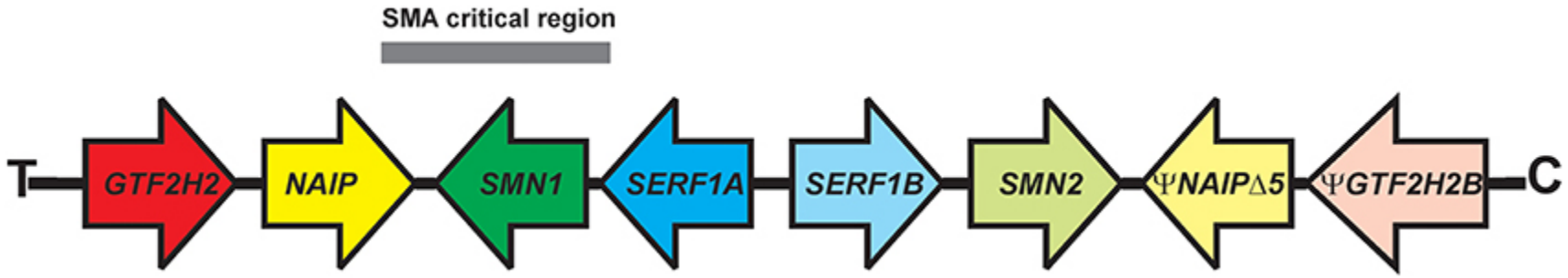
- Autosomal recesive disease
- Incidence: 1 in 6,000 - 10,000 live births
- Carrier frequency: 1 in 40 - 60
- The second most frequent fatal disease with autosomal recesive inheritance (after cystic fibrosis)
- Characterised by degeneration of alpha motor neurons







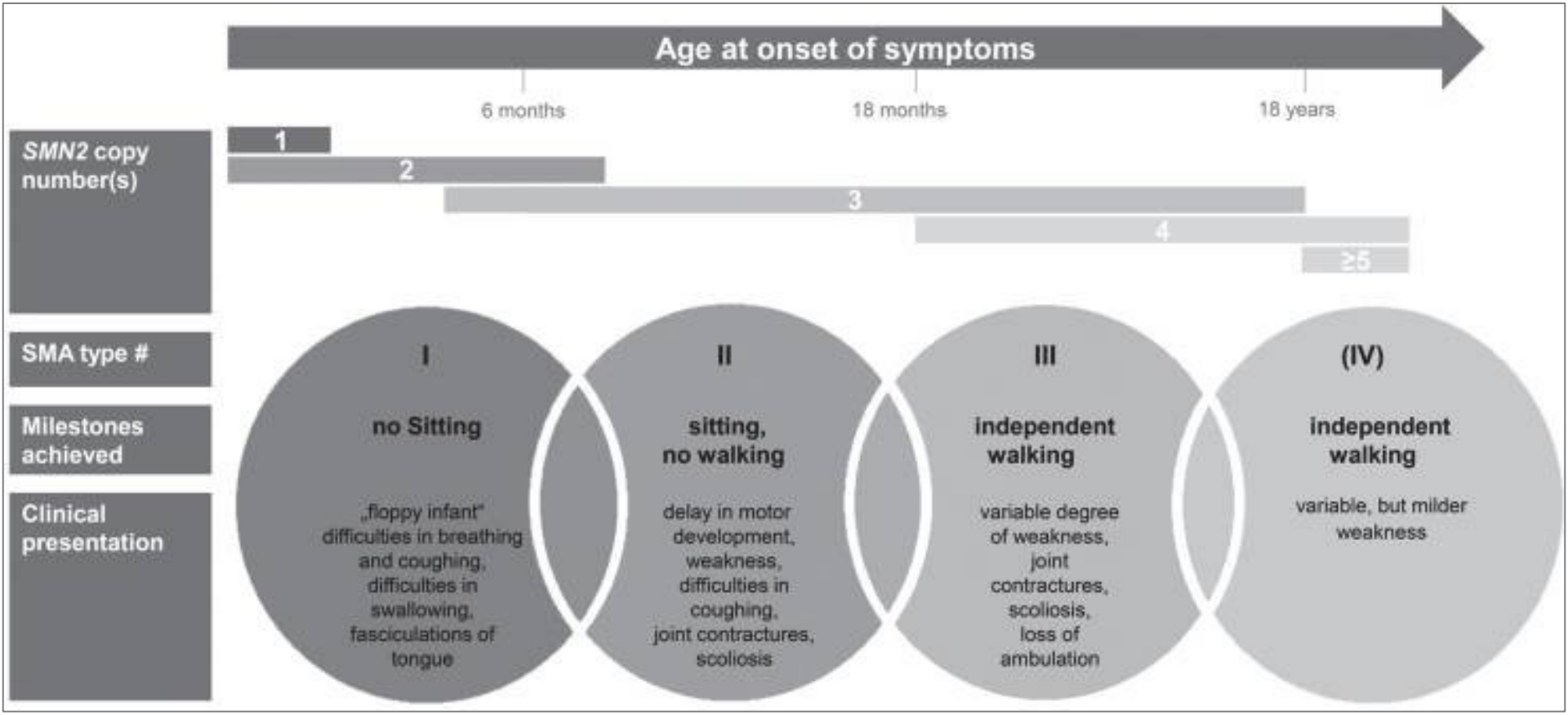
- **95% of SMA is caused by homozygous deletion of the *SMN1* gene (Survival of Motor Neuron 1).**
- *SMN1* is located on the chromosome region 5q12-5q13.38 containing 500-kb inverted duplication→ *SMN1* has its almost identical copy – the *SMN2* gene.
- *SMN1* and *SMN2* are homologous to except for few nucleotides.
- The is copy number variation of *SMN1* and *SMN2* in human genom.



Based on the age of onset and clinical course, 4 clinical types of SMA are distinguished:

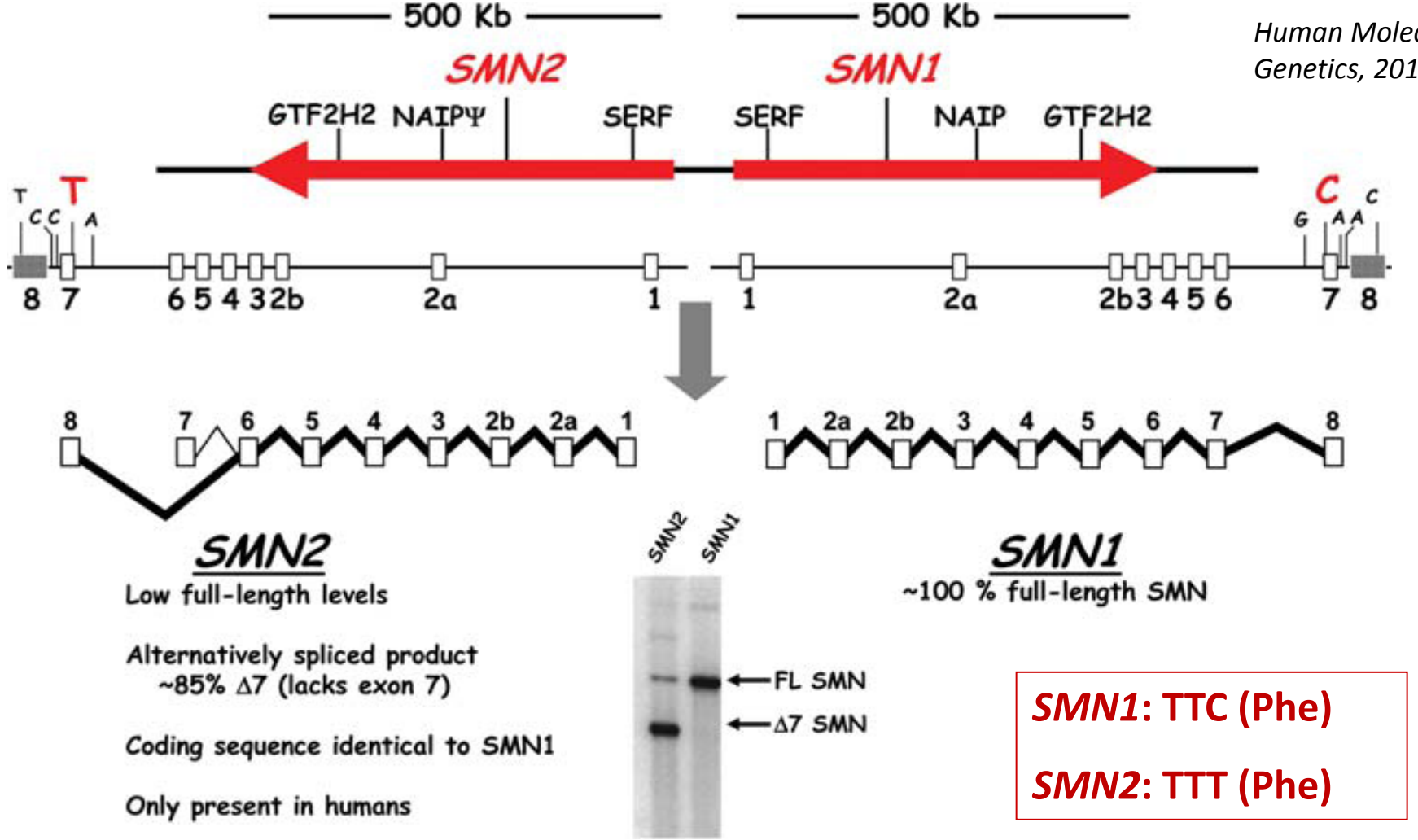
- **Type I** – characterised by severe muscle weakness and hypotonia at birth or within the first 6 months; death from respiratory failure occurs usually within the first 2 years.
- **Type II** – first symptoms begin 6-18 months after birth and life expectancy is 2-30 years; patients are able to sit but unable to walk independently.
- **Type III** - first symptoms are typically observed after 2 years of life; patients are able to walk but often are wheelchair-bound within or after adolescence.
- **Type IV** - symptoms appear in adulthood; patients have mild motor impairment.

- **95% of SMA (regardless of the type) is caused by the homozygous deletion of the *SMN1* gene.**
- **Clinical severity is modified by copy number the *SMN2* gene.**



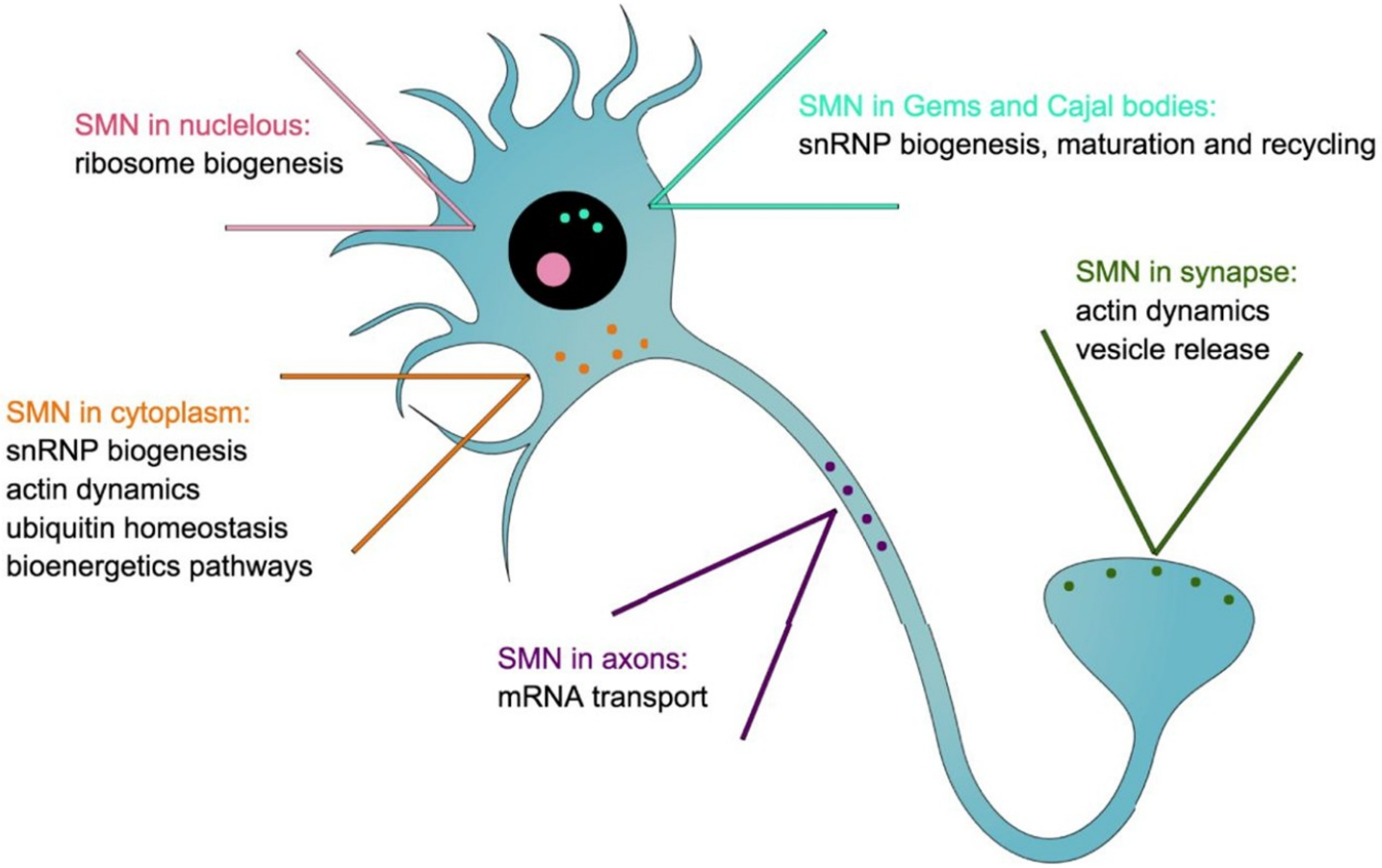
Clinical classification of SMA types according to onset, milestones achieved, and clinical presentation. Typically associated SMN2 copy numbers are displayed.

Schorling DC, J Neuromuscul Dis. 2020; 7(1): 1–13.



Human Molecular Genetics, 2010

SMN1 and *SMN2* are located in close proximity on chromosome 5. Both *SMN* genes have 9 exons and encode an identical protein. **Fundamental difference between *SMN1* and *SMN2* is a silent C→T substitution in exon 7 of *SMN2*, which results in a strong reduction of exon 7 inclusion during mRNA splicing.** Consequently, 85% of the mature mRNA from *SMN2* lacks exon 7 and encodes truncated unstable protein.



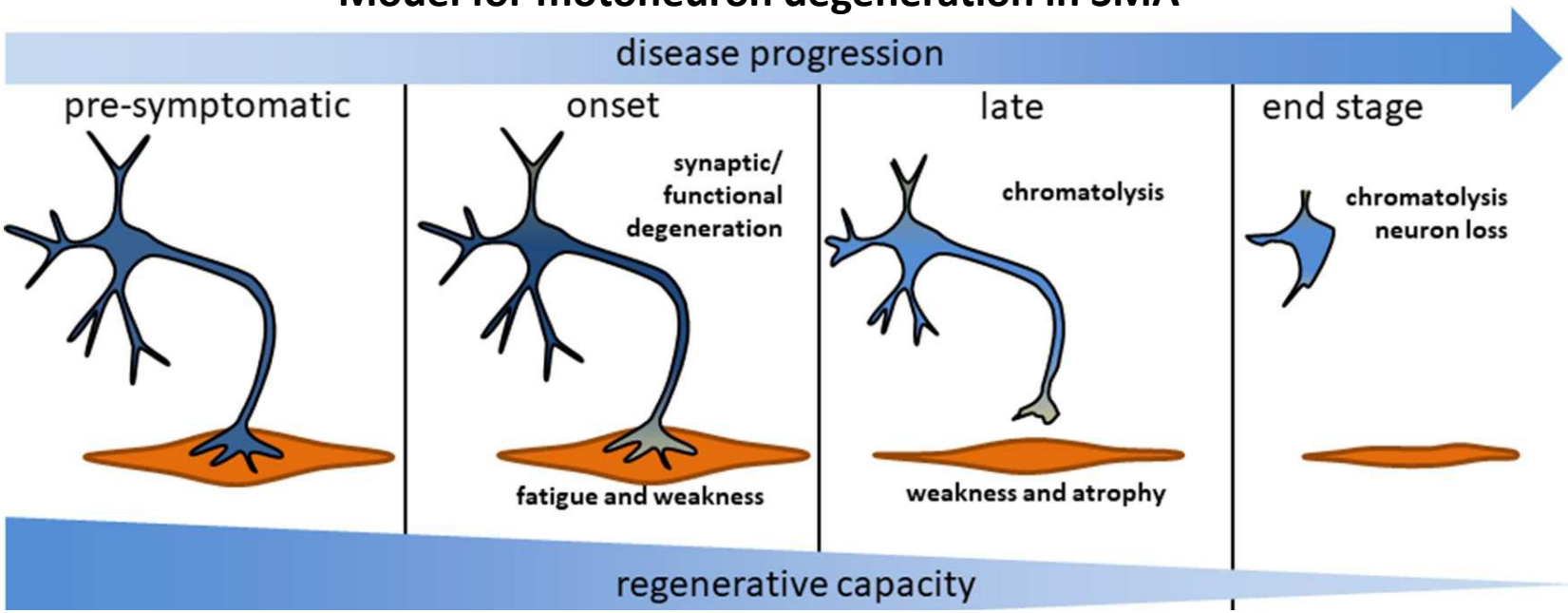
The SMN protein is expressed ubiquitously.

In motor neurons, SMN has several functions:

- in cytoplasm - production of small nuclear ribonucleoproteins (snRNPs are active in recognizing and removing introns from *pre*-mRNA) and actin dynamics,
- in nucleus - snRNP biogenesis,
- in nucleolus - ribosome biogenesis,
- in axons - mRNA transport and local translation,
- in the synpase - actin dynamics and vesicle release.

Bowerman M. Dis. Model. Mech. 2017;10:943-954

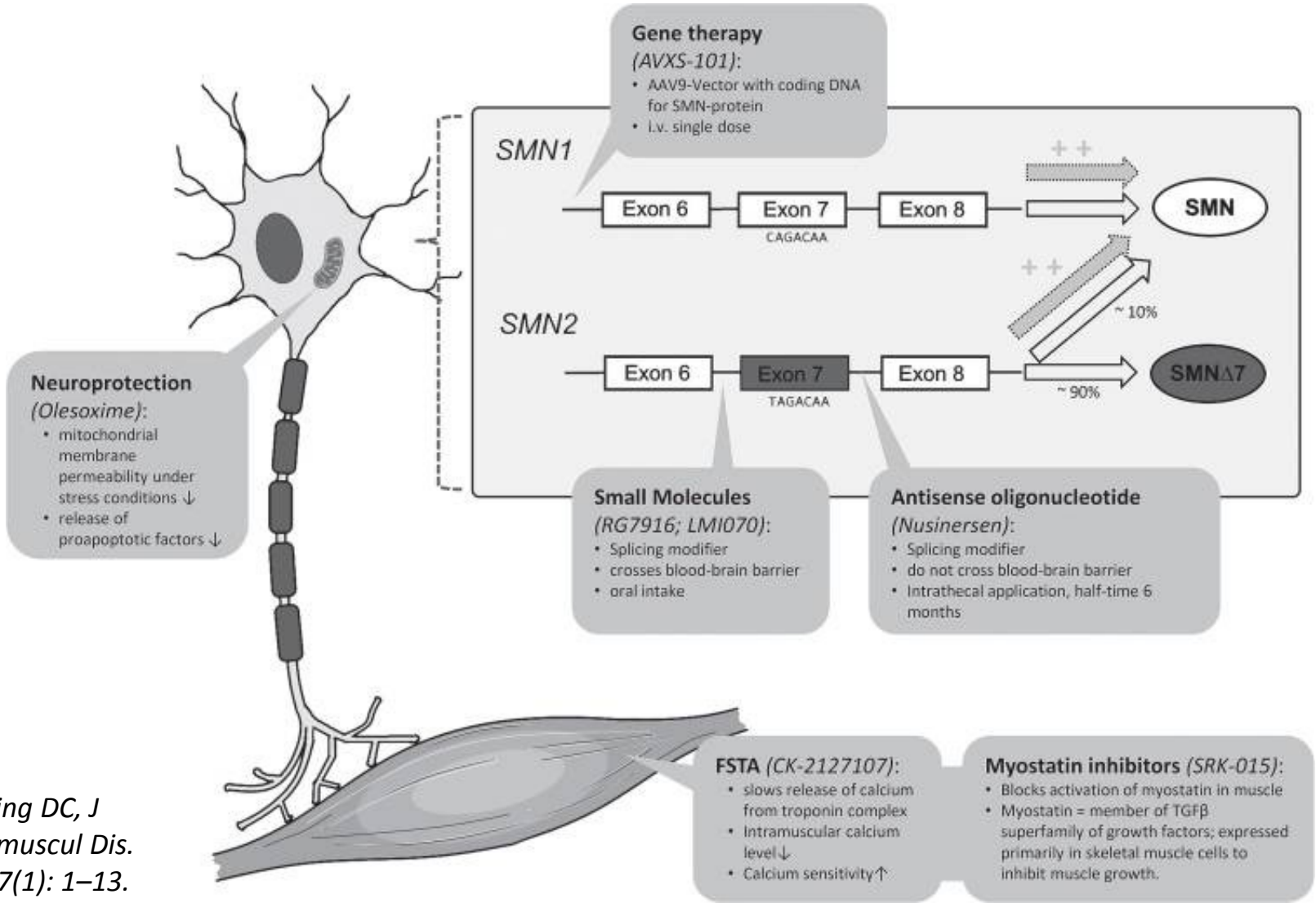
Model for motoneuron degeneration in SMA



- Motoneuron loss observed in post mortem analyses is preceded by functional degeneration of central synapses and neuromuscular junctions and subsequent axonal damage.
- During disease progression those processes become less reversible. **The complete loss of motoneuron is a irreversible change.**
- **The beneficial effects of SMA therapies are dependent on disease duration at the time of intervention. Disease duration before treatment is critical and a delayed intervention leads to a less efficient rescue. The effect of SMA therapies is strongest in pre-symptomatic patients.**

Hensel N. Frontiers in Neurology 2020 Vol 11

Figure summarizes SMA therapeutic approaches and illustrates the respective molecular mechanisms of action. Actual SMA therapeutic developments can be subdivided into therapies **(1) modifying splicing of *SMN2* (production of more amount of full length mRNA), (2) replacing the *SMN1* gene, and (3) upregulating muscle growth.**



Schorling DC, J Neuromuscul Dis. 2020; 7(1): 1–13.

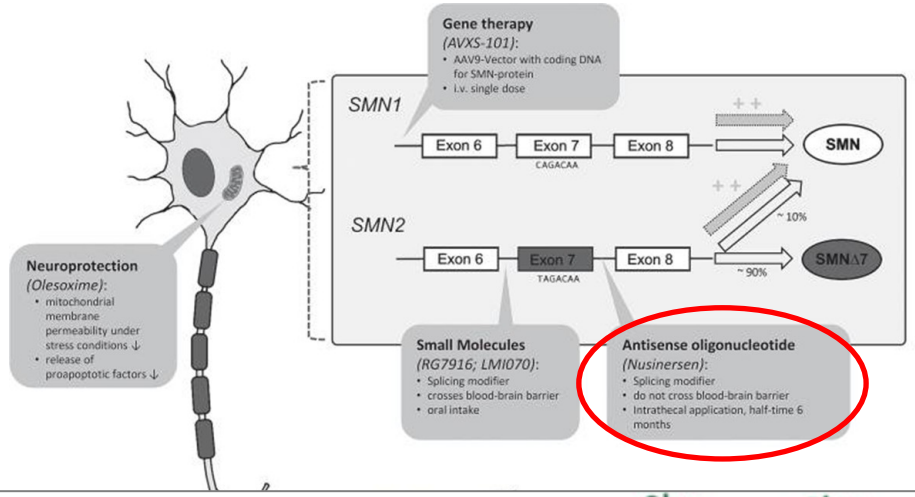
The first drug approved for SMA treatment was Spinraza (Nusinersen), an antisense-oligonucleotide (ASO) that enhances the inclusion of exon 7 in mRNA transcripts of *SMN2*.

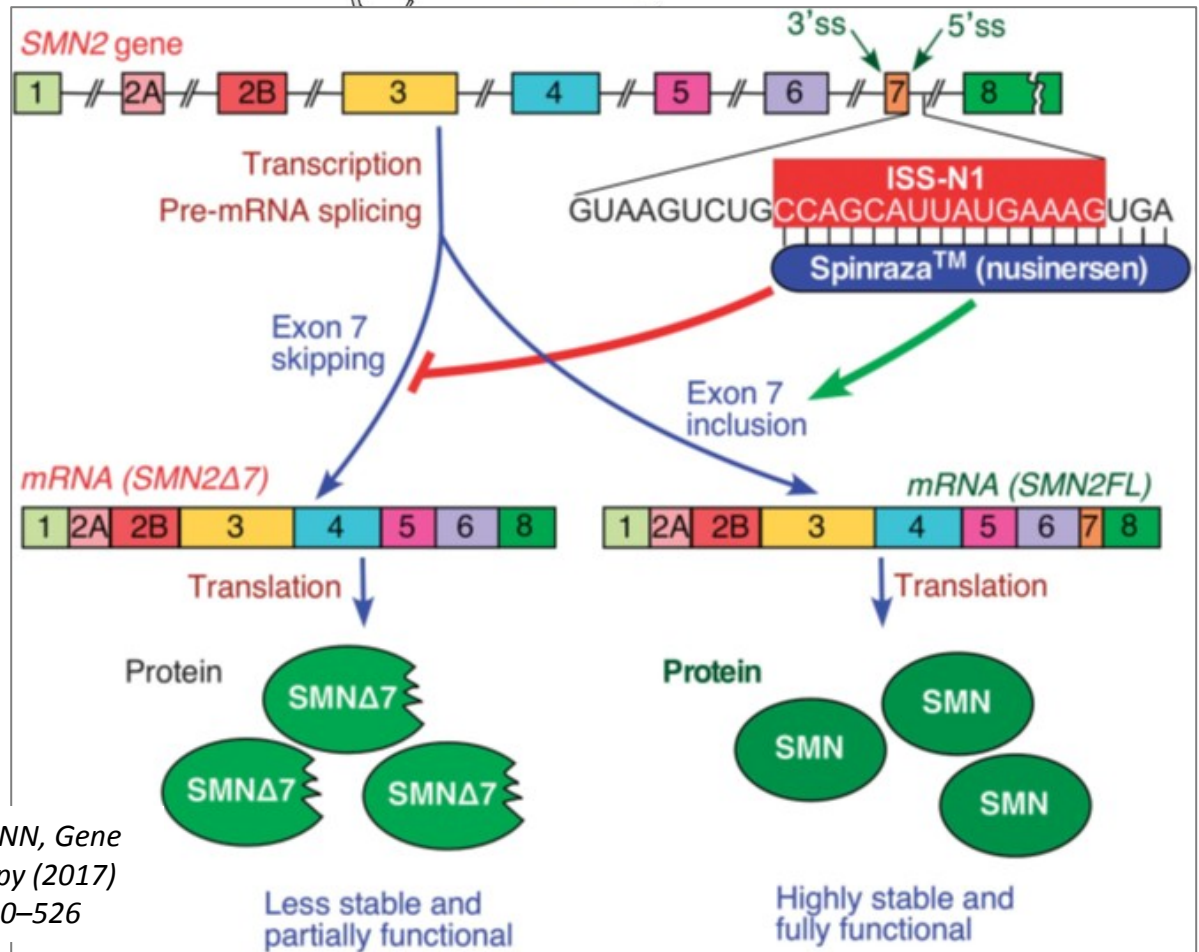
Spinraza binds to an intronic splice-silencing-site in intron 7 of *SMN2* and thereby suppresses the binding of other splice-factors.

This results in an increased proportion of *SMN2*-mRNA with included exon 7 and consecutively more functional full-length SMN protein.

Spinraza was approved by the Federal Drug Agency (FDA) in 2016.

Spinraza is administered intrathecally, does not cross blood-brain barrier.



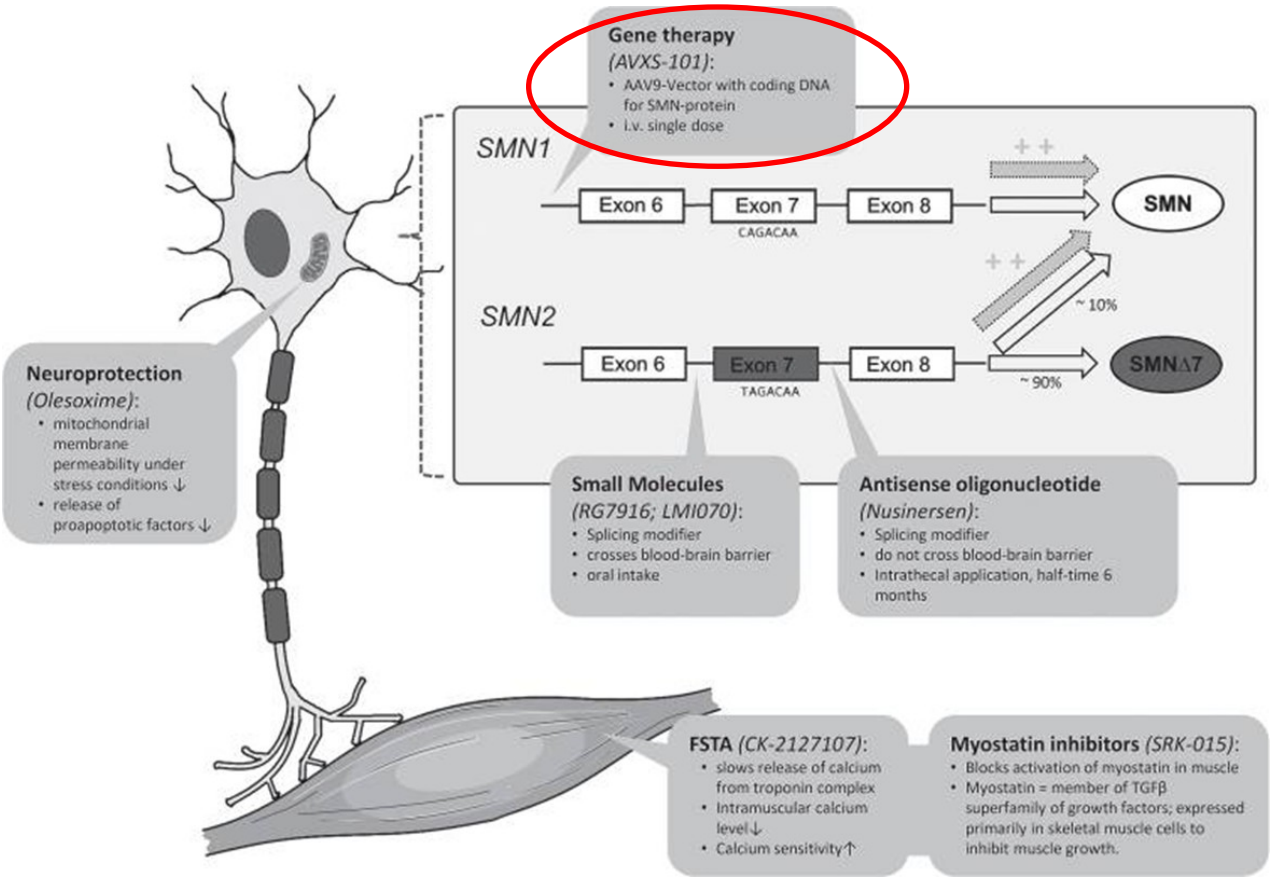


Singh NN, Gene Therapy (2017) 24, 520–526

The most advanced approach in the treatment of SMA is gene therapy, which directly targets the dysfunctional *SMN1* gene.

In 2019, the FDA approved Zolgensma (Onasemnogene Abeparvovec), an adeno-associated virus 9 (AAV9) delivering a cDNA which codes the full length SMN protein, as a gene replacement therapy.

Zolgensma (AVXS-101) was approved for intravenous application in patients with SMA under 2 years of age; intrathecal applications will be studied in children aged 2-6 years. The AAV9 crosses the blood-brain barrier which induces SMN expression in the CNS and in peripheral organs.

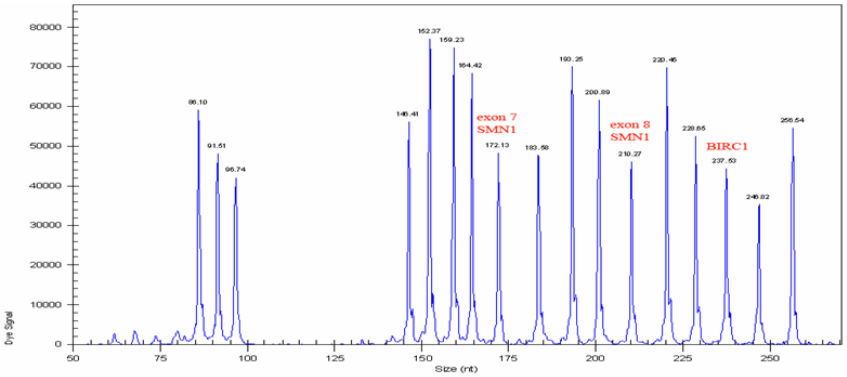


Schorling DC, J Neuromuscul Dis. 2020; 7(1): 1–13.

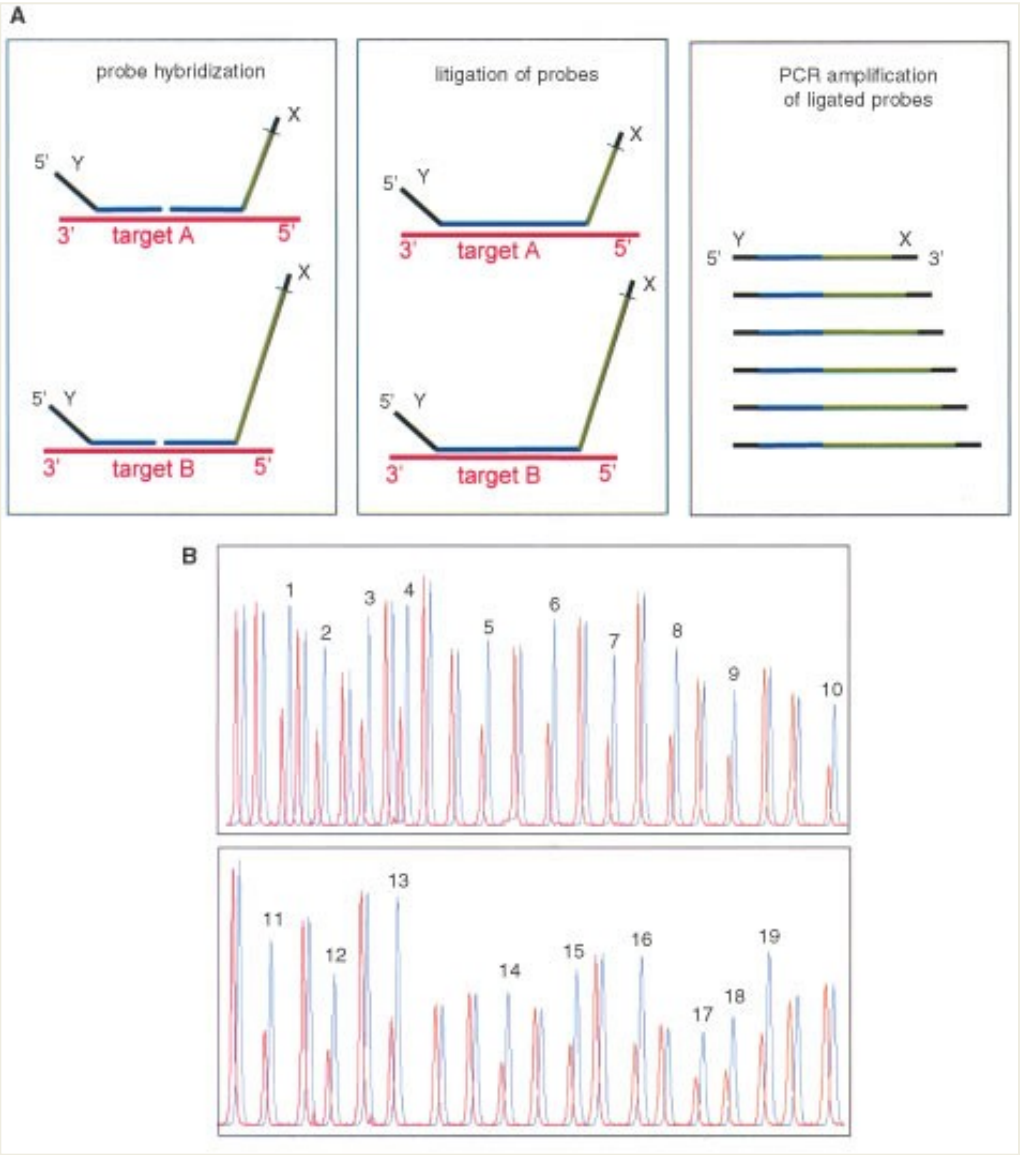
- 95% of SMA patients have homozygous deletion of the *SMN1* gene.
- 5% of patients have deletion of the *SMN1* gene on one chromosome and a small scale pathogenic variant on the second one.

The basic methodical approach in SMA molecular genetic diagnostics is MLPA (multiple ligation dependent probe amplification), which determines the copy number of *SMN1* and *SMN2* for identification:

- patients with homozygous *SMN1* deletion,
- patients with heterozygous *SMN*1 deletion; subsequently *SMN1* is sequenced for identification of second pathogenic variant,
- heterozygous carriers of *SMN1* deletion,
- *SMN2* copy number for determination of presumed SMA type and suitability of treatment.



Multiplex Ligation-dependent Probe Amplification (MLPA)

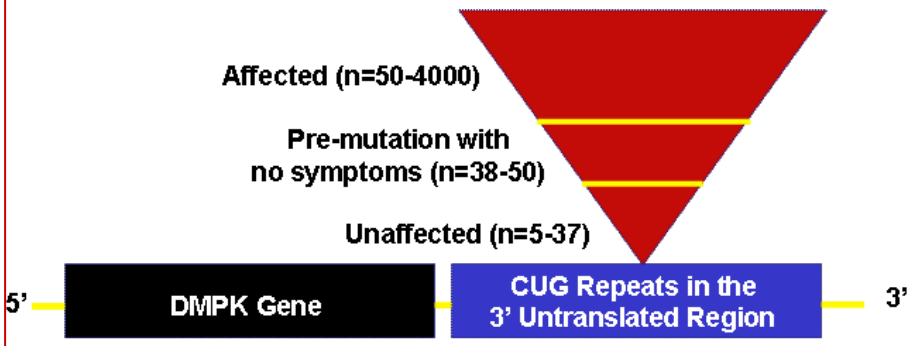


- Denatured genomic DNA is hybridised with a mixture of probes.
- Each MLPA probe consists of two oligonucleotides. The two parts of each probe hybridise to adjacent target sequences and are ligated by a ligase.
- All probe ligation products are amplified simultaneously by PCR using a single primer pair 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.

Myotonic dystrophy type 1; MD1

MD1 is caused by expansion of the CTG repeat in the 3'UTR of the dystrophia myotonica protein kinase gene (*DMPK*, 19q13.3),

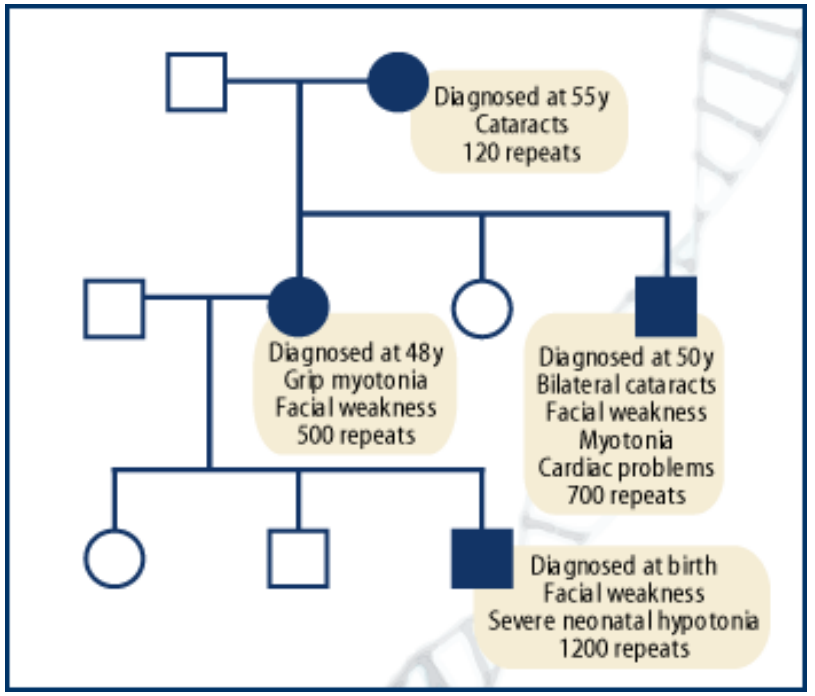
- **AD inheritance**



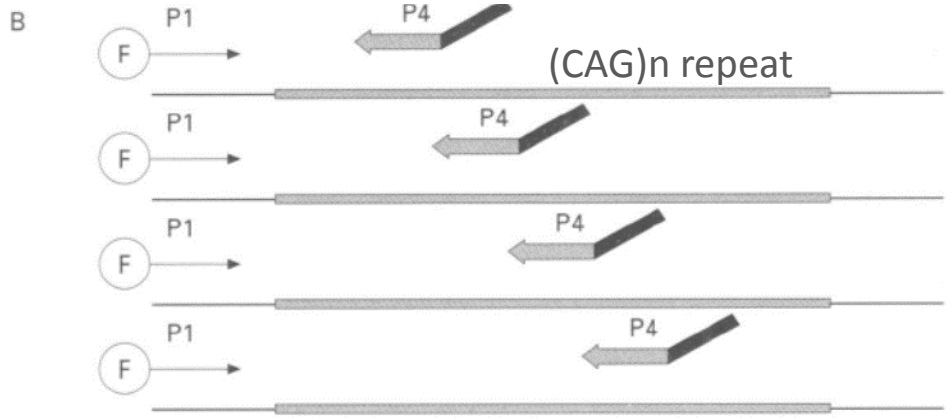
- Individuals with 5 to 37 repeats are unaffected.
- Individuals with 38-50 repeats carry the premutation. These individuals are asymptomatic. However, these repeats are unstable and can expand during meiosis. As a result, such individuals are at risk of having affected children.
- ~ 50 to 150 repeats are consistent with the mild adult-onset form of MD1,
  ~ 100 to 1000 repeats are consistent with the classic adult or childhood onset form of MD1,
  > 1000 repeats are consistent with the congenital form of MD1 and often result in severe neonatal complications.

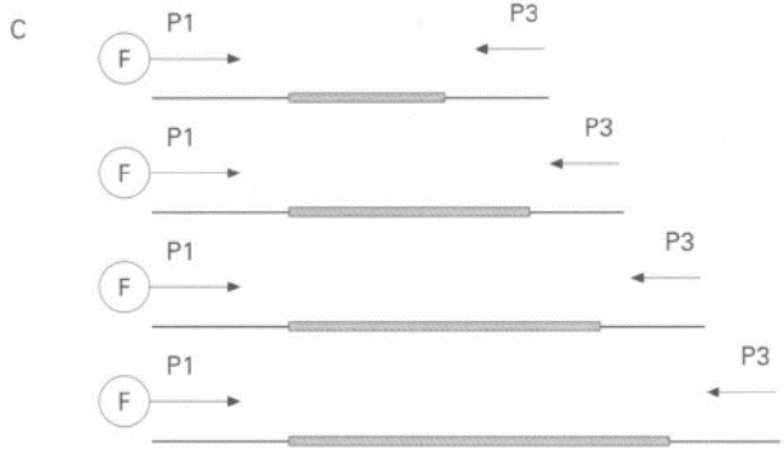
- The expanded CTG repeat - **dynamic mutation** - the number of repeats tends to increase in size over generations.

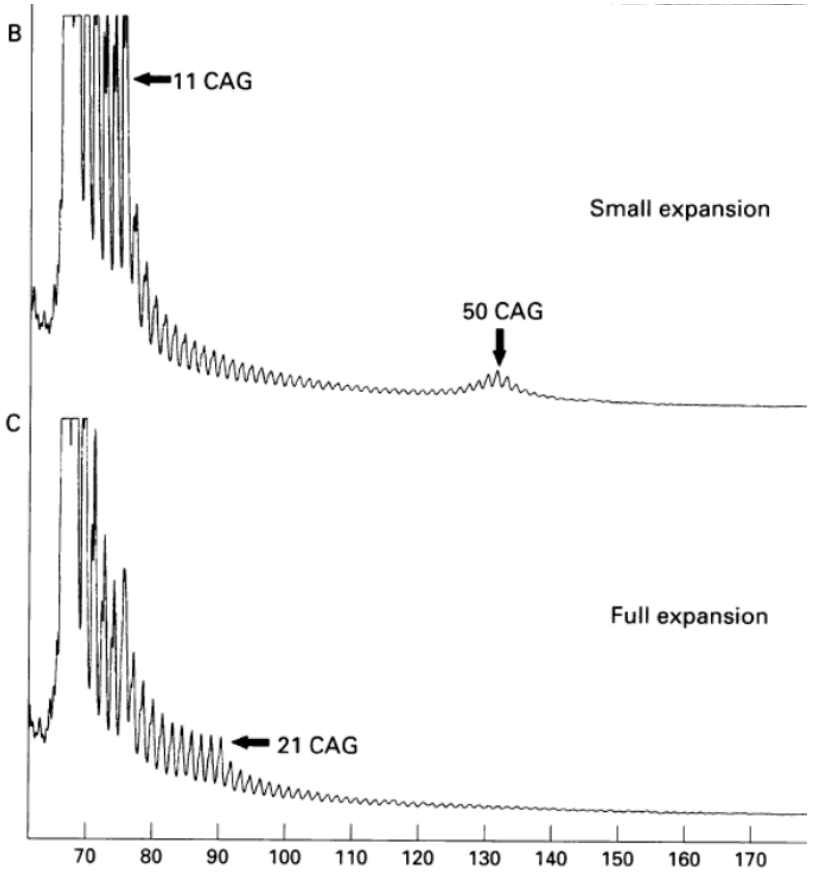
- Expansion of the CTG repeats commonly occurs during meiosis. As a result, children of affected individuals tend to have severe symptoms and earlier onset than their parents.



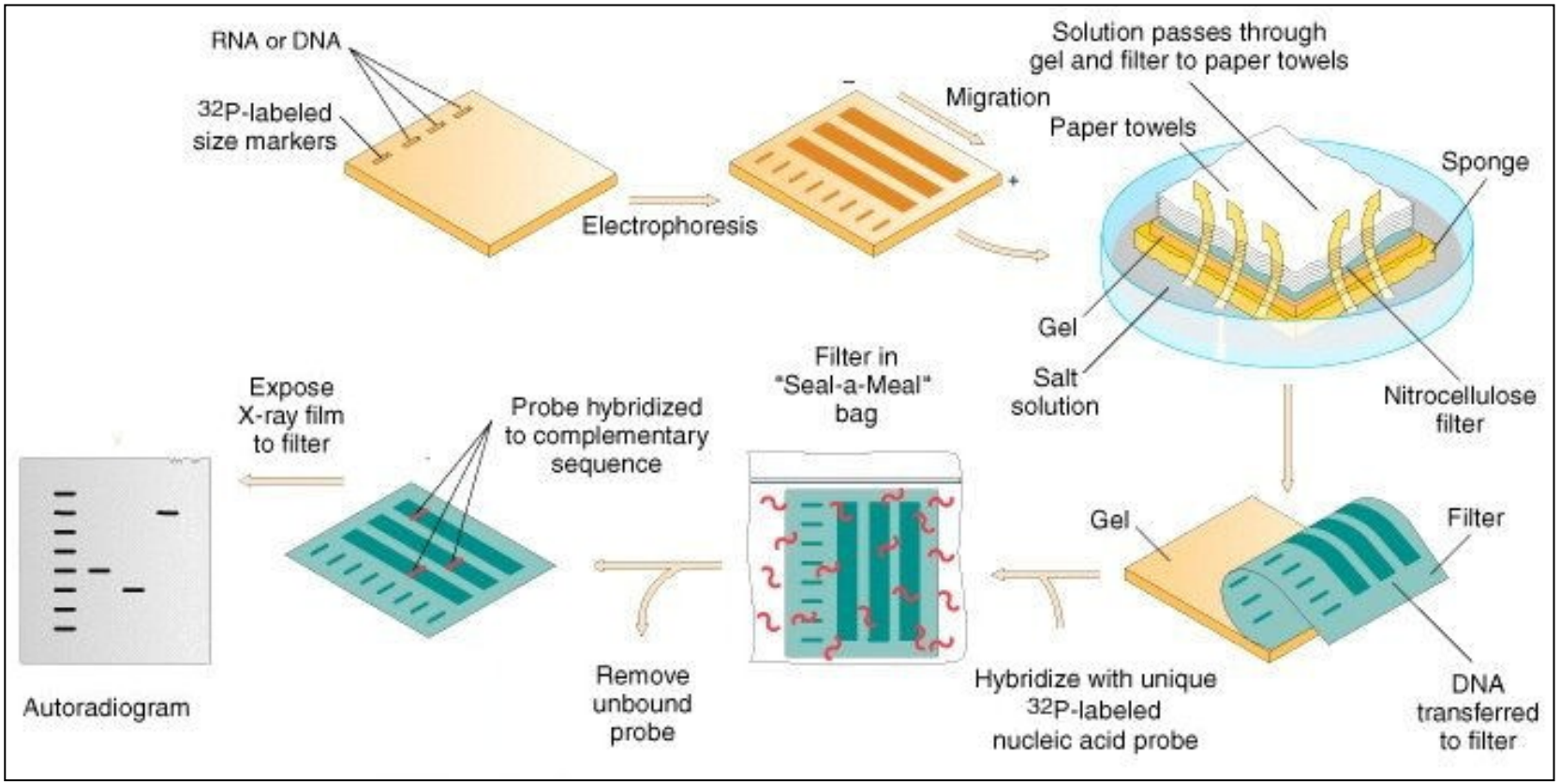
The repeat primed PCR (RP-PCR) uses three primers: one that is fluoresceinated and flanks the repeat region (P1-F), a second that is complementary to the repeat but carries a nonspecific tail sequence (P4), and a third that is complementary to the nonspecific tail sequence (P3). RP-PCR produces a characteristic profile of amplicons of increasing length, which differ by the length of a repeat unit (3 bp), but of diminishing yield. This profile enables the rapid identification of a pathogenic repeat.





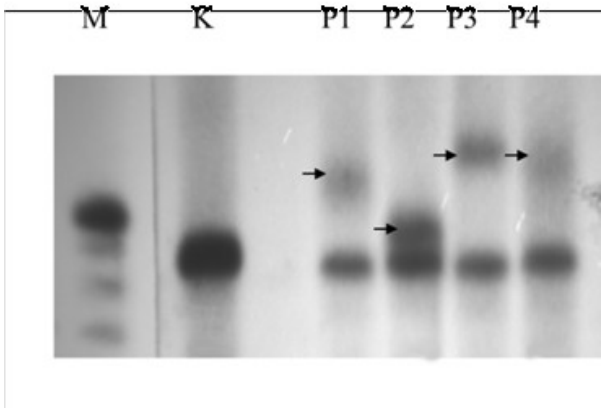


- **RP-PCR enables identification of a pathogenic repeat but not the size of the expansion.**
- **The size can be determined by Southern blot a hybridizace.**



DNA is

- cleaved by a restriction endonuclease,
- electrophoresed,
- transfered to membrane (bounding of negative charged DNA to positive charged membrane is mostly used),
- hybridised to a radioactive labeled probe,
- after removing of a unbound probe autoradiography is performed



Molecular genetic diagnostics

- the results must be interpreted with knowledge of the molecular nature of the disease and knowledge of the structure and function of encoded protein.
- the results must be interpreted in relation to the patient 's phenotype and results of other patient examinations (biochemistry, pathology, NMR, EMG, etc.).
- it is necessary to return to the results of already examined patients with an unconfirmed genetic diagnosis and test them with new techniques and perform new interpretations of the identified sequence variants.
- it is necessary to participate in international quality control of DNA diagnostics for individual diseases.