Molecular genetic diagnostics of monogenic diseases

Jana Zídková
zidkova.jana@fnbrno.cz

University Hospital Brno

Centre of molecular biology and genetics

Department of inherited genetic disorders

Molecular genetic diagnostics of rare diseases:

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

Why are we actually finding out?

1. Confirmation of clinical diagnosis

- psychological support
- prediction of the course of the disease
- specific treatment - certain disease, certain mutation (example at the end of the lecture)

2. Segregation of variants/disease in family members

- early treatment (in preclinical phase)
- genetic counseling – testing of partner, preimplantation diagnostics, prenatal diagnostics

Molecular genetic diagnostics of monogenic diseases

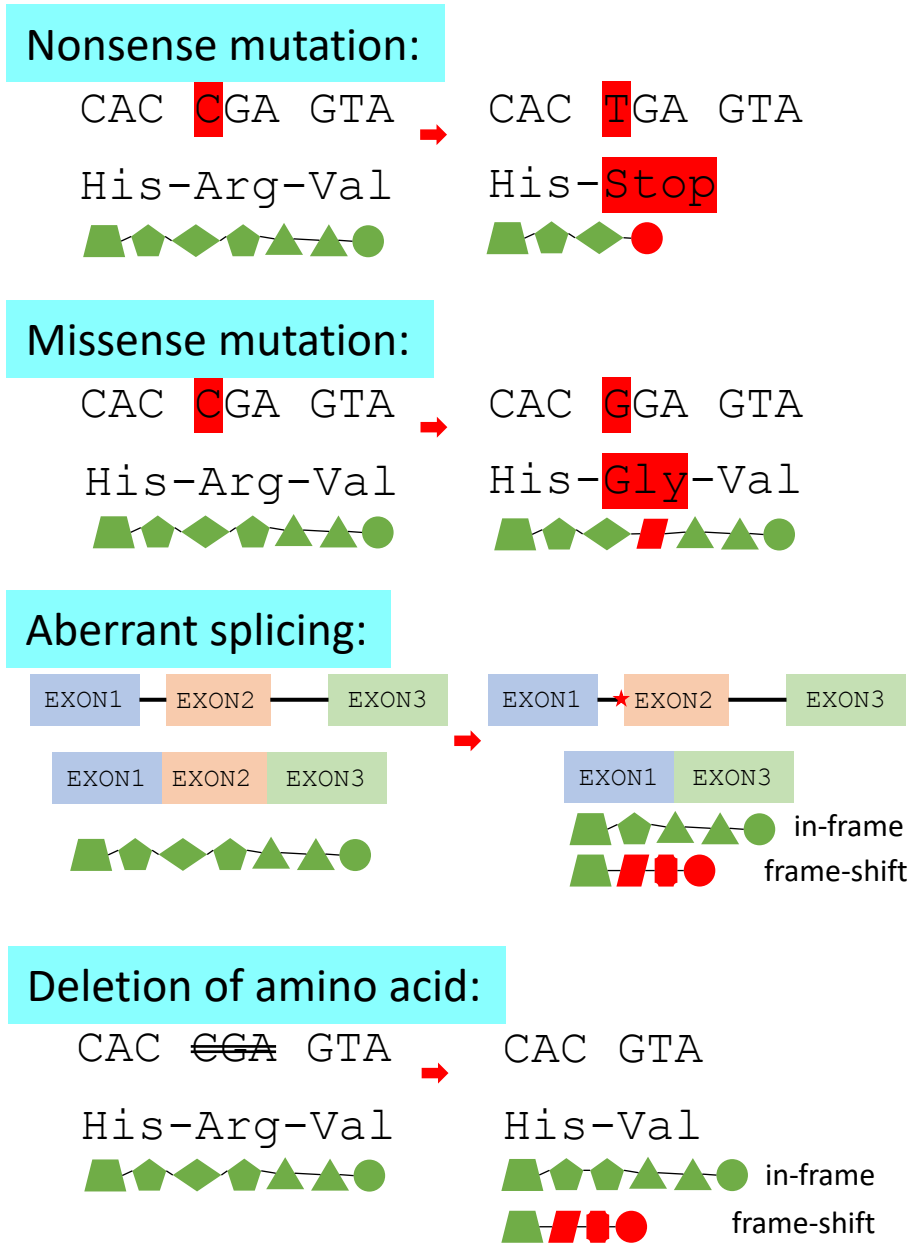
Germline mutation/mutations in **one gene**, not large deletions/insertions containing several genes:

- **identification of small scale variants:**

nucleotide substitutions → nonsense mutation (creating premature stop codon)
→ missense mutation (change of amino acid)
→ affecting the splice site (aberrant splicing)

small deletions/insertions → change in amino acid chain (without/with creating premature stop codon)

- whole exon deletions / duplications (copy number variations, CNV)

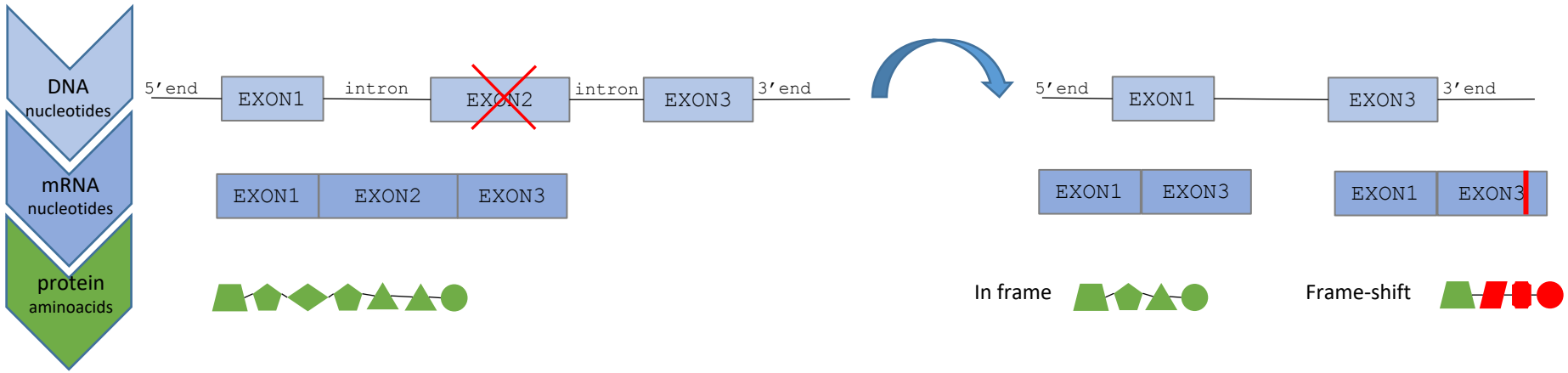


Molecular genetic diagnostics of monogenic diseases

Germline mutation/mutations in one gene, not large deletions/insertions containing several genes:

- identification of small scale variants: nucleotide substitutions, small deletions / insertions
- whole exon deletions/duplications (copy number variations, CNV)
  - → **in-frame** - change in amino acid chain without creating premature stop codon
  - → **frame-shift** - change in amino acid chain with creating premature stop codon

Deletion of one exon:



Molecular genetic diagnostics of monogenic diseases

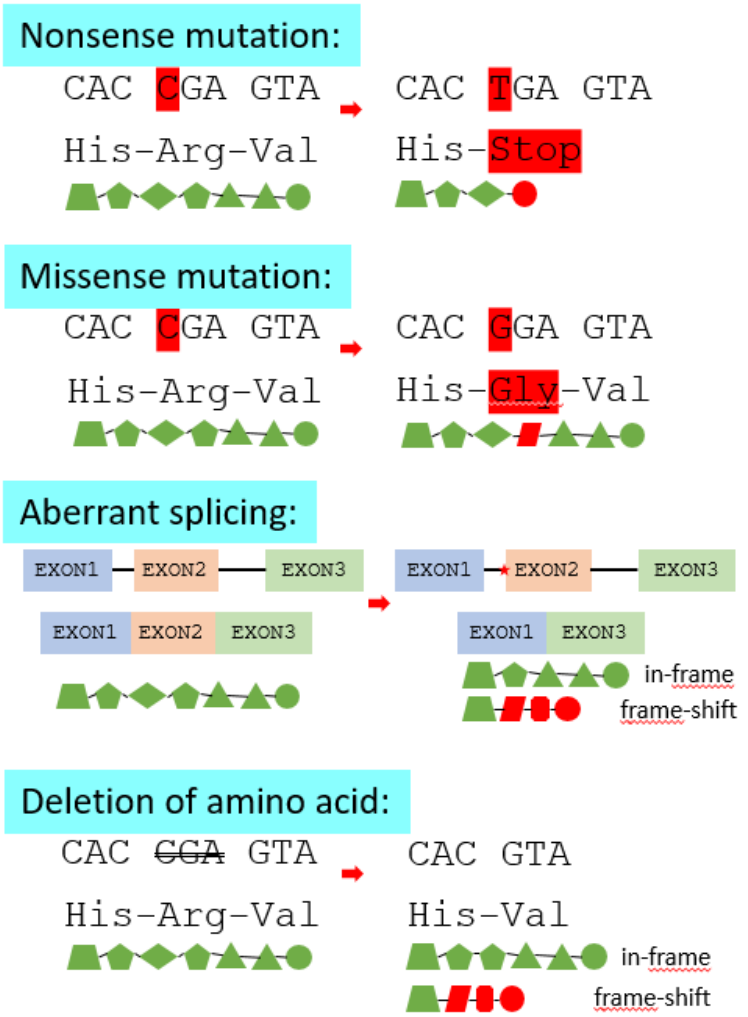
Material: DNA isolated from the whole blood

1. Classic Sanger sequencing
2. Next generation sequencing
3. MLPA – CNV detection
4. RP-PCR - detection of repeat expansions
5. Southern blot and hybridization - detection of repeat expansions / deletions

1. Classic Sanger sequencing

- identification of small scale variants: nucleotide substitutions, small deletions / insertions

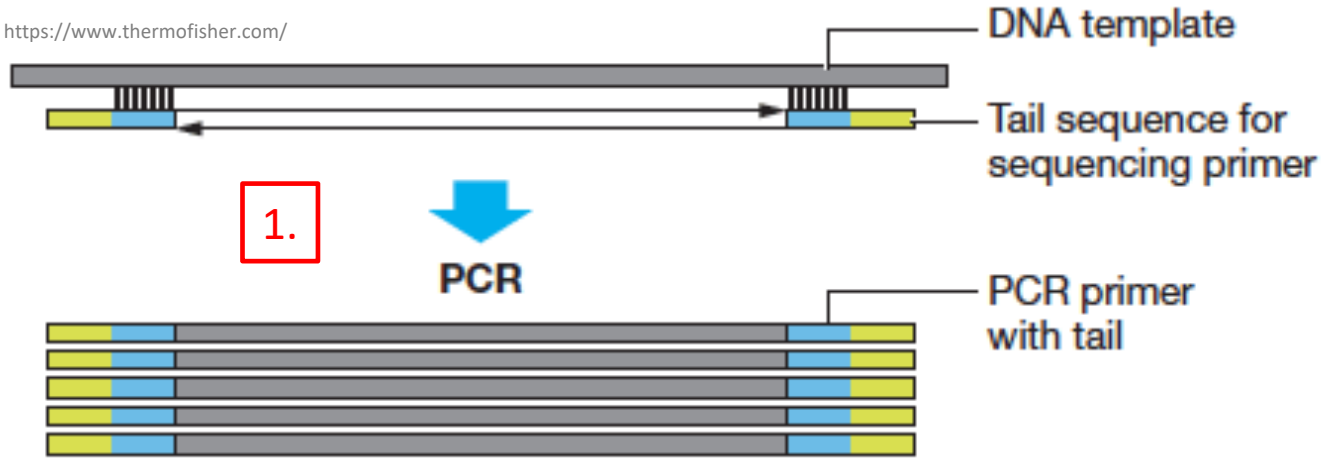
- ***Method description:***
  1. PCR (polymerase chain reaction, amplification of known target sequence)
  2. sequencing
  3. result



1. Classic Sanger sequencing

- identification of small scale variants: nucleotide substitutions, small deletions / insertions
- ***Method description:***
  1. PCR (polymerase chain reaction, amplification of known target sequence) > 2. sequencing > 3. result

https://www.thermofisher.com/



1. PCR

- PCR amplifies a specific region of a DNA
- specific primers - complementary to the target region
- reagents: DNA polymerase, primers, deoxynucleoside triphosphates (dNTPs), buffer solution, bivalent cations (typically magnesium)
- volume of 10–100 µL in small reaction tubes
- thermal cycler - heats and cools the reaction tubes to achieve the temperatures required

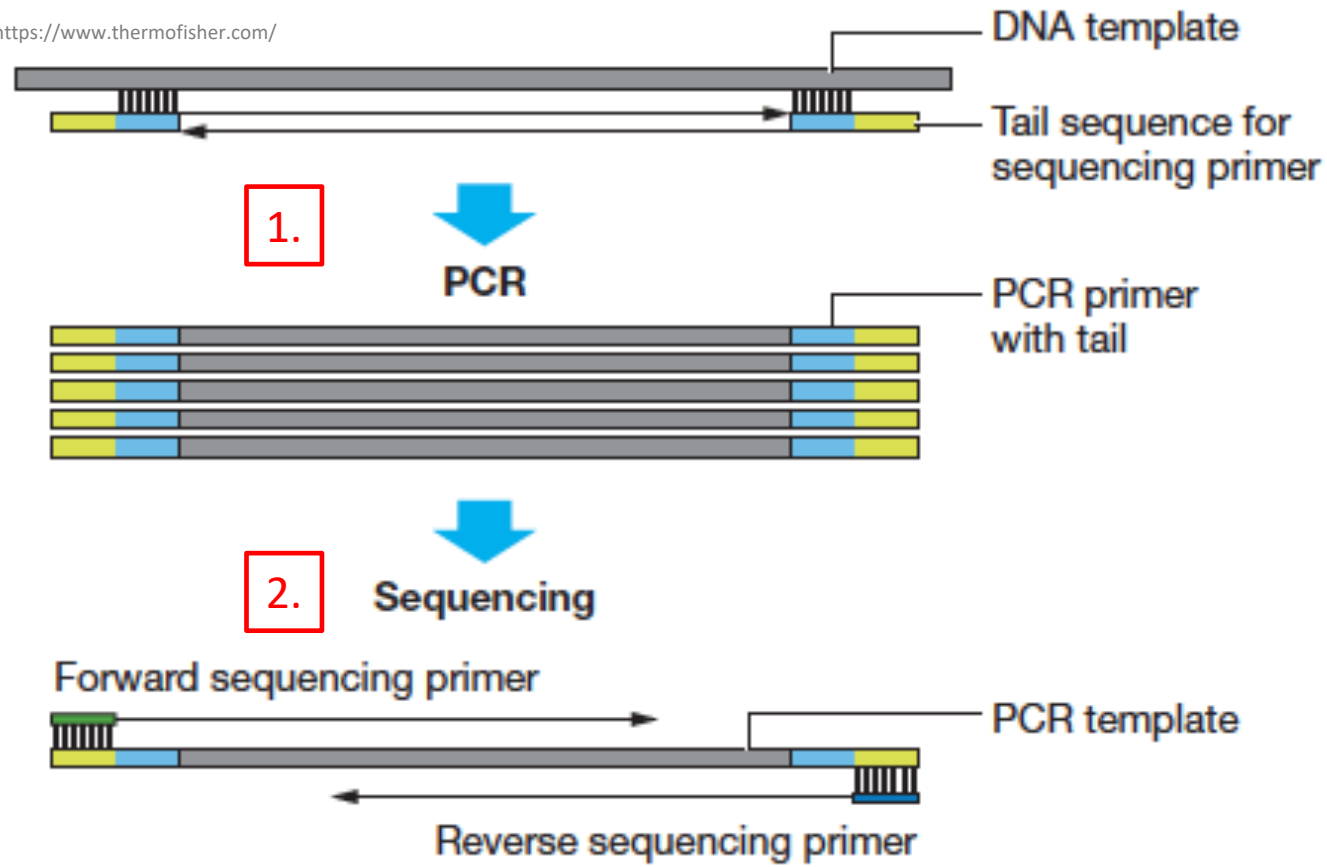
PCR procedure

1. start denaturation	94°C/6min	
2. denaturation	94°C/1min	cycling steps2-4; 20-40x
3. annealing	60°C/1min	
4. elongation	72°C/1min	
5. final elongation	72°C/6min	

1. Classic Sanger sequencing

- identification of small scale variants: nucleotide substitutions, small deletions / insertions
- ***Method description:***
  1. PCR (polymerase chain reaction, amplification of known target sequence) > 2. sequencing > 3. result

https://www.thermofisher.com/



2. sequencing

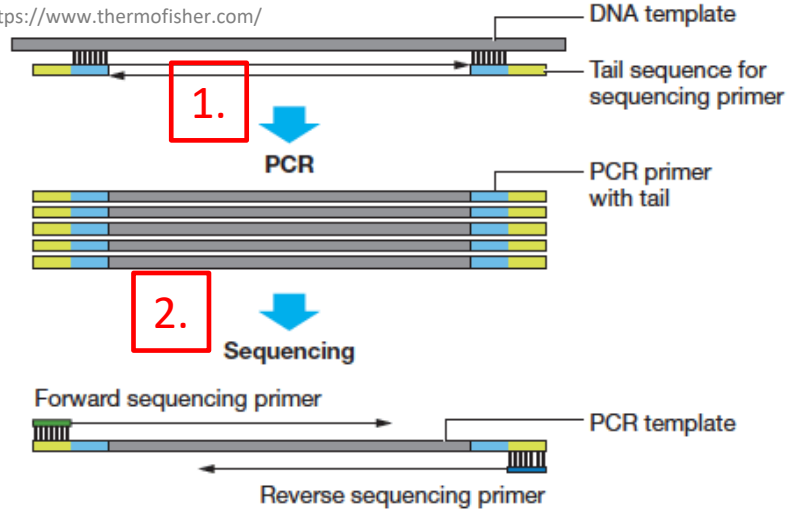
- determining of nucleotide order
- chain termination method
- one primer
- sequencing reaction in cycler > capillary electrophoresis

1. Classic Sanger sequencing

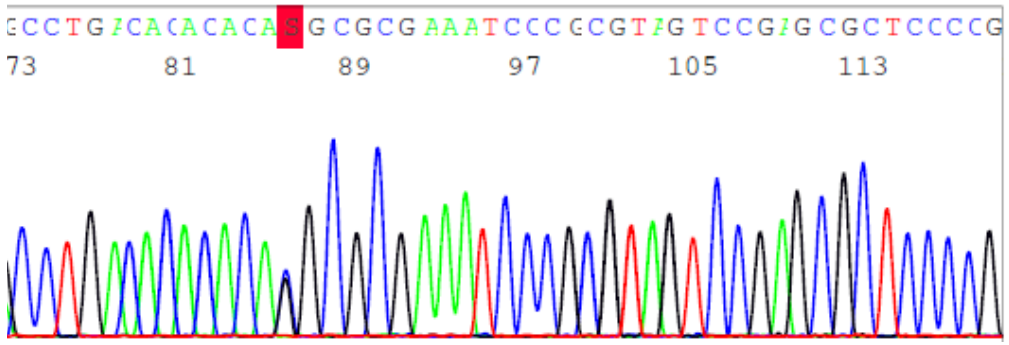
- identification of small scale variants: nucleotide substitutions, small deletions / insertions
- ***Method description:***

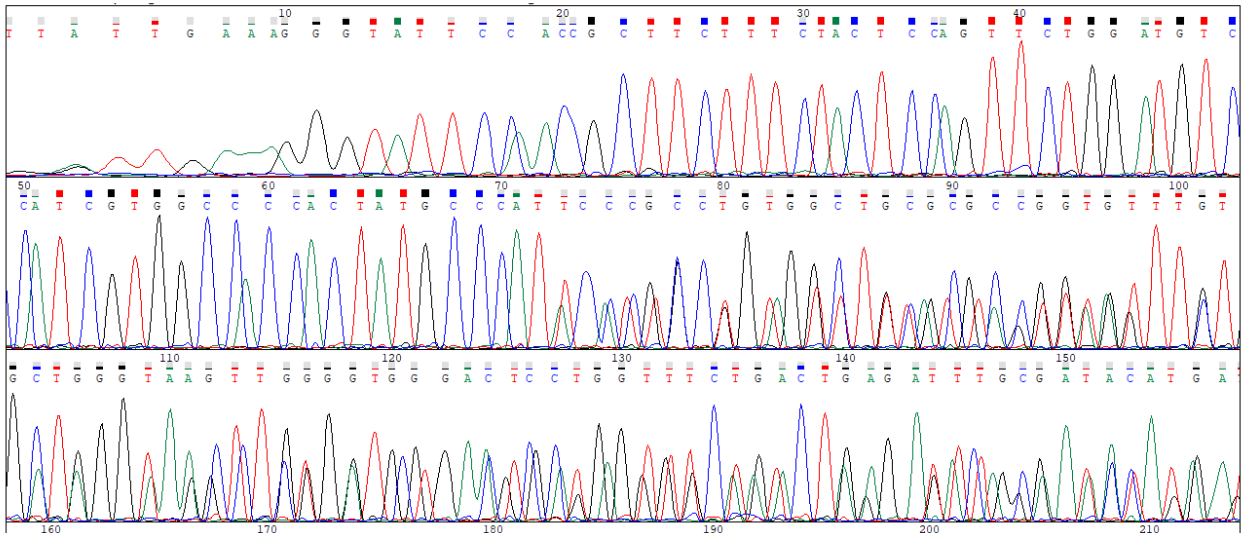
  1. PCR (polymerase chain reaction, amplification of known target sequence) > 2. sequencing > 3. result

https://www.thermofisher.com/



3. result





- nucleotide order
- change of one/several nucleotides
- heterozygote / homozygote

1. Classic Sanger sequencing

A. Sequencing of the certain part of the gene including the position of pathogenic variant
- segregation of variant in family members

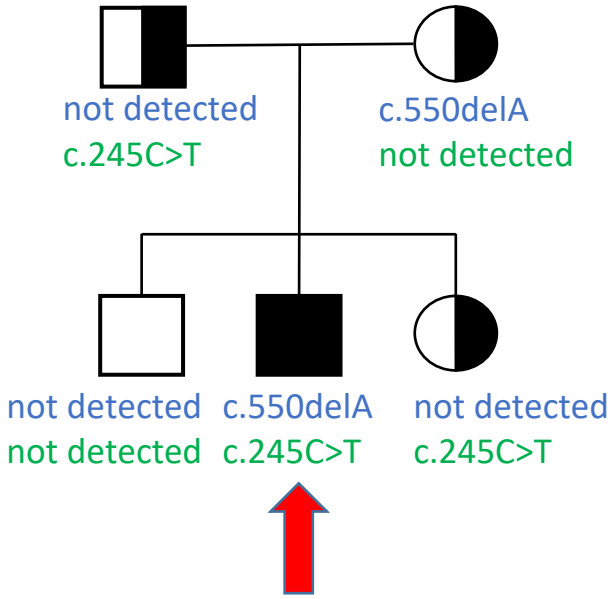
B. Sequencing of the whole gene by several PCR reactions:
- in past: gene by gene approach (time-consuming and costly)
- gene with clear clinical-genetic relationship, not a very long gene (example: phenylketonuria)

1. Classic Sanger sequencing

A. Sequencing of the certain part of the gene including the position of pathogenic variant
   - segregation of variant in family members

Patient with **autosomal recessive** limb girdle muscular dystrophy, 2 pathogenic variants in *CAPN3* gene
CAPN3: c.245C>T and c.550delA

Presence of variants in:
- mother
- father
- brother
- sister



not detected
not detected

unaffected – risk of disease the same as in the population

not detected
c.245C>T

c.550delA
not detected

unaffected but carrier of one pathogenic variant:
- parents: 25% probability of child with disease (preimplantation diagnostics, prenatal diagnostics)
- sister: testing of partner

1. Classic Sanger sequencing

A. Sequencing of the certain part of the gene including the position of pathogenic variant
- segregation of variant in family members

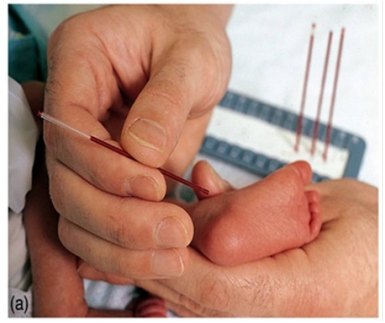
B. Sequencing of the whole gene by several PCR reactions:
- in past: gene by gene approach (time-consuming and costly)
- gene with clear clinical-genetic relationship, not a very long gene (example: phenylketonuria)

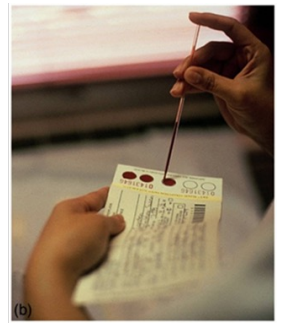
1. Classic Sanger sequencing

B. Sequencing of the whole gene: gene with clear clinical-genetic relationship, not a very long gene
example: phenylketonuria

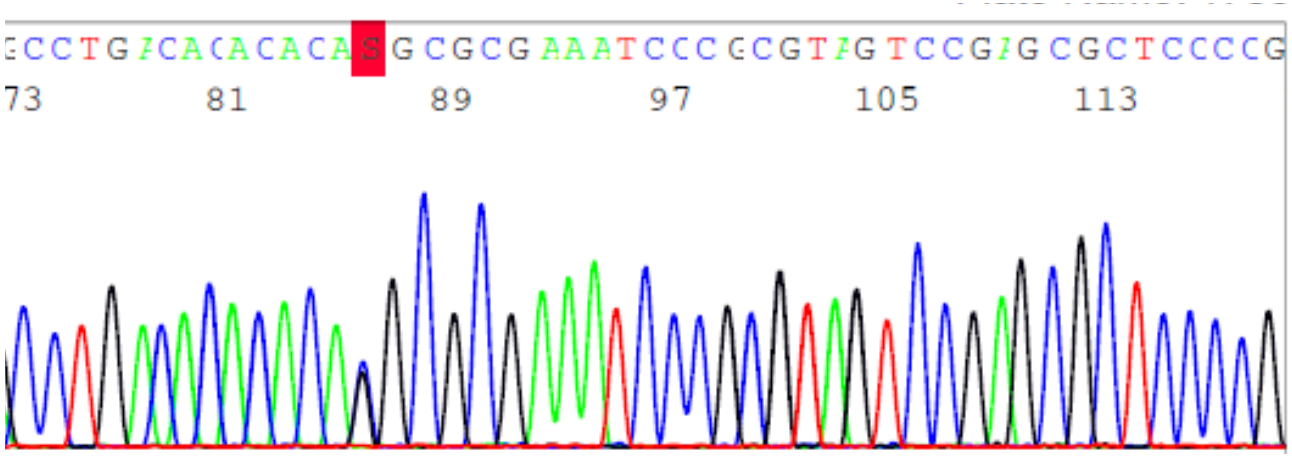
Phenylketonuria (PKU)

autosomal recessive metabolic disease (deficiency of phenylalanine hydroxylase); gene *PAH* (12q23.2)





- 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 phenylalanine hydroxylase**



- in 98% of cases, two pathogenic variants in the *PAH* gene are identified = clinical diagnosis of PKU is confirmed

Molecular genetic diagnostics of monogenic diseases

1. Classic Sanger sequencing
2. Next generation sequencing
3. MLPA – CNV detection
4. RP-PCR - detection of repeat expansions
5. Southern blot and hybridization - detection of repeat expansions / deletions

Material: DNA isolated from the whole blood

2. Next generation sequencing (NGS)

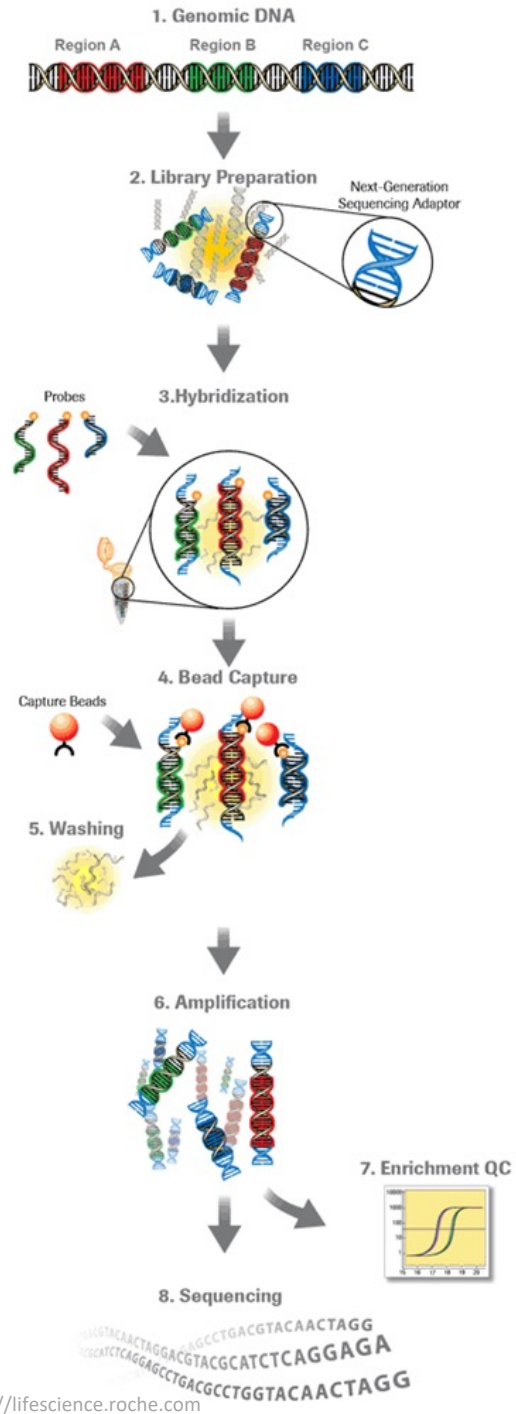
➢ *1. DNA samples are converted into sequencing libraries*

- DNA is randomly sheared into smaller fragments by mechanical or enzymatic methods
- adapters for sequencing and multiplexing are added to DNA ends
- regions of interest within the library are captured using oligonucleotide probes (hybridization)
- probe-targeted fragment complex is separated from other fragments that are not bound to probes
- amplification of targeted regions

oligonucleotide probes:
- targeted panel sequencing - selected sets of genes or gene regions
- whole exome sequencing (WES)
- whole genome sequencing (WGS)

➢ 2. sequencing

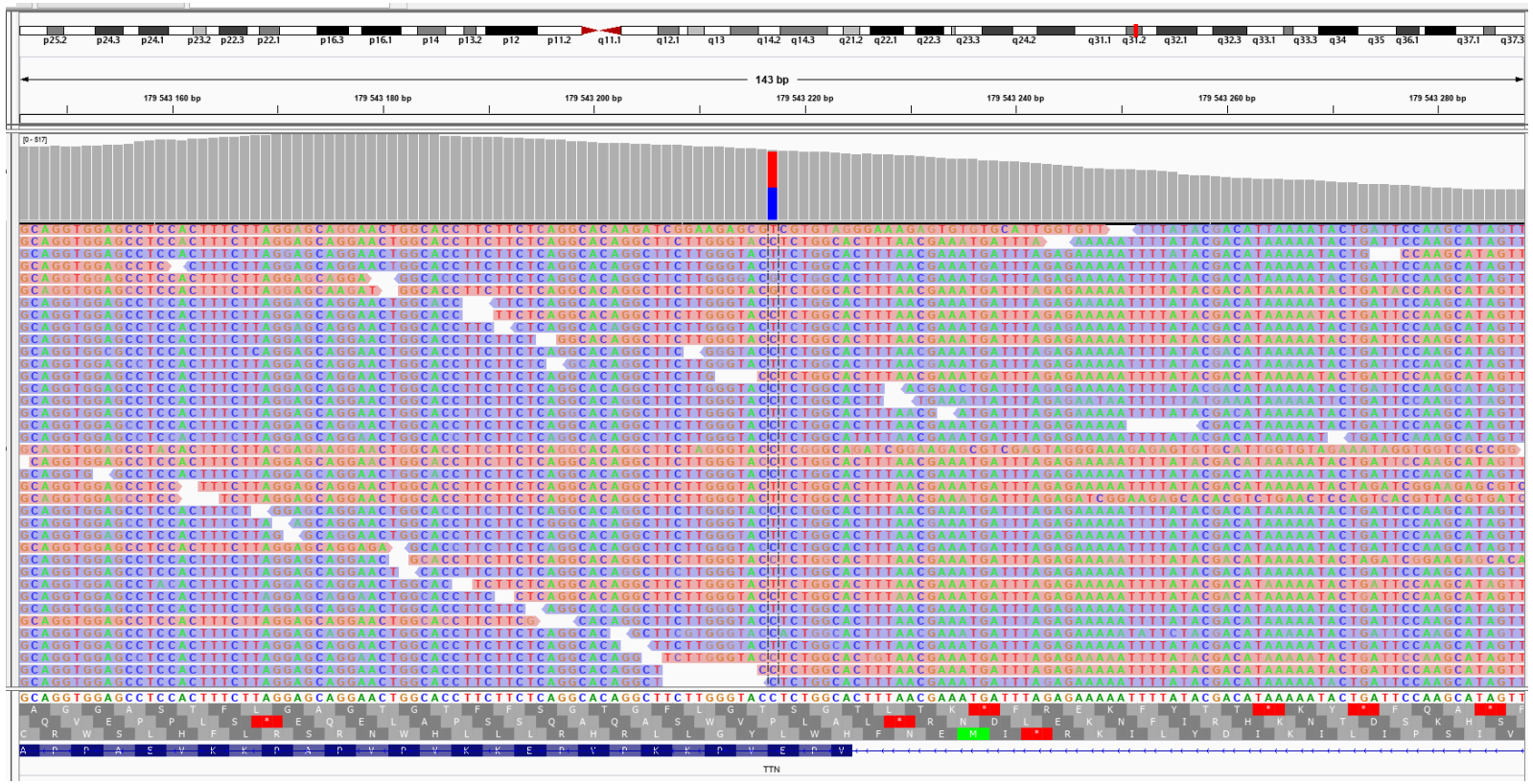
- NGS of targeted regions
- sequences millions of fragments in a massively parallel fashion
- improving speed and accuracy while reducing the cost of sequencing



2. Next generation sequencing (NGS)

➢ 3. data analysis

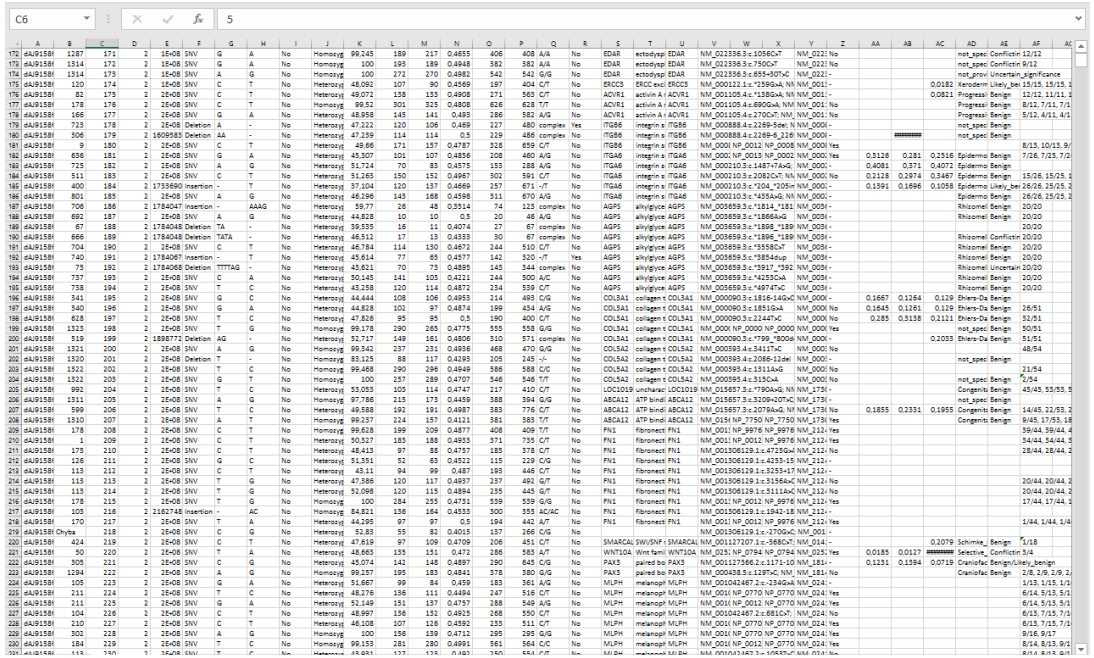
- the instrument software identifies nucleotides (a process called base calling) and the predicted accuracy of those base calls
- by commercial software or bioinformatics pipelines
- includes alignment of NGS reads
- identification and annotation of sequence variants



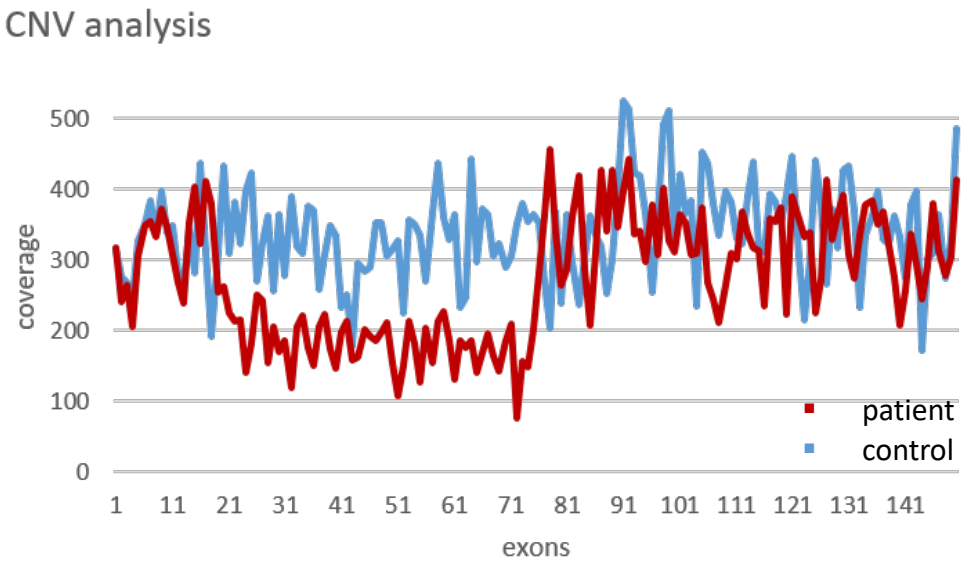
2. Next generation sequencing (NGS)

Result:
- identification of small scale variants: nucleotide substitutions, small deletions / insertions
- whole exon deletions / duplications (copy number variations, CNV)

Identification of a large number of sequence variants



Identification of whole exon deletions / duplications: by comparing the number of reads for individual exons



2. Next generation sequencing (NGS)

Example:

Patient's case:

Patient has complained of difficulties in running a climbing stairs since his early teens. Patient has elevated creatine kinase levels. Proximal weakness in the upper and lower limbs has been progressive and he displays wasting of trunk muscles and slight hyperlordosis.

Neurologist requests analysis of genes associated with muscular dystrophy.

2. Next generation sequencing (NGS)

Example:

Muscular dystrophies and myopathies

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

In past before NGS:

- genes analysed sequentially by **classical DNA sequencing**
- starting with a gene with the most likely mutation occurrence > negative result > another gene
- TIME AND FINANCIALLY CONSUMING
- only a certain number of genes analysed

NGS era:

- all genes associated with the disease analysed at the same time (in parallel) = **targeted panel**
- FAST AND RELATIVELY CHEAP

2. Next generation sequencing (NGS)

Example:

Muscular dystrophies and myopathies

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

In past before NGS:

- genes analysed sequentially by **classical DNA sequencing**
- starting with a gene with the most likely mutation occurrence > negative result > another gene
- TIME AND FINANCIALLY CONSUMING
- only a certain number of genes analysed

NGS era:

- all genes associated with the disease analysed at the same time (in parallel) = **targeted panel**
- FAST AND RELATIVELY CHEAP

We use panel including all 162 genes

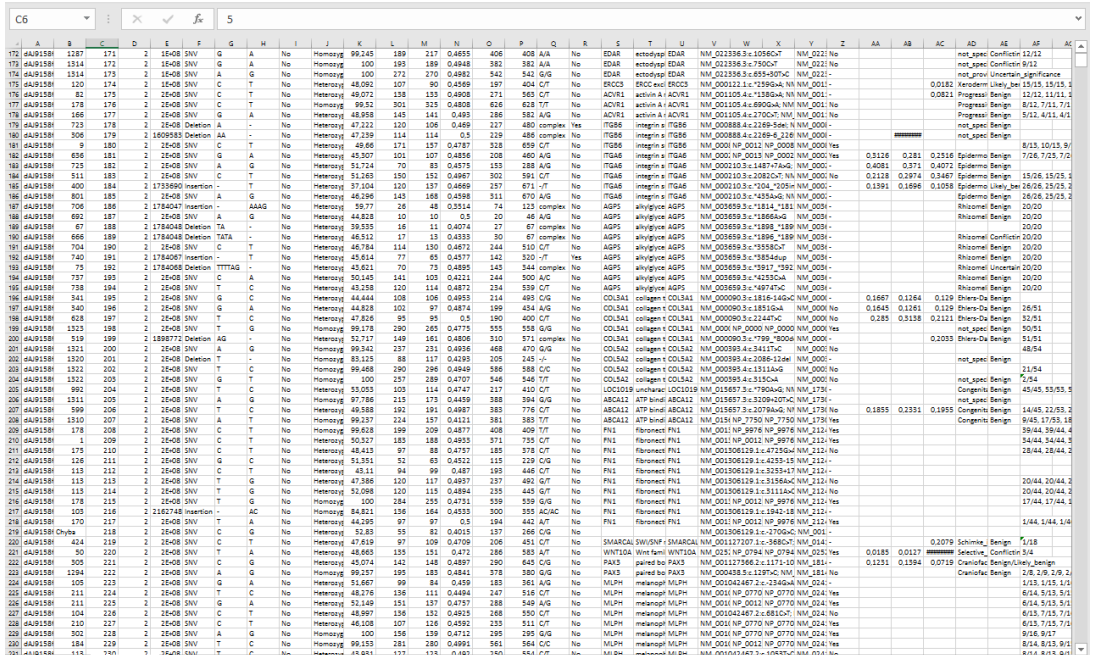
2. Next generation sequencing (NGS)

Example:

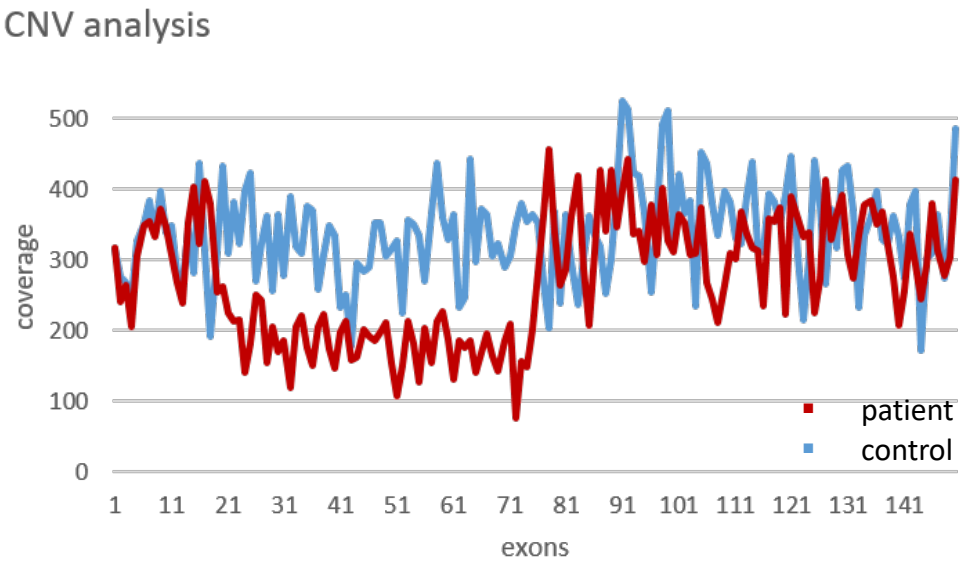
Result:

- identification of small scale variants: nucleotide substitutions, small deletions / insertions
- whole exon deletions / duplications (copy number variations, CNV)

Identification of a large number of sequence variants



Identification of whole exon deletions / duplications: by comparing the number of reads for individual exons



Interpretation of causality

2. Next generation sequencing (NGS)

Interpretation of 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],

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

2. Next generation sequencing (NGS)

Interpretation of sequence variants

Patient's case:
Patient has complained of difficulties in running a climbing stairs since his early teens. Patient has elevated creatine kinase levels. Proximal weakness in the upper and lower limbs has been progressive and he displays wasting of trunk muscles and slight hyperlordosis.

Result:

A) Identification of pathogenic variant/variants

e.g. two pathogenic variants in *CAPN3* > confirmed diagnosis of limb girdle muscular dystrophy

CAPN3 (NM_000070.3): c.245C>T p.(Pro82Leu) / c.550delA p.(Thr184Argfs*36)

B) Identification of variants of uncertain significance: e.g. variant in *DYSF*

- genetic-clinical correlation
- segregation of variant in family
- type of inheritance

C) Only benign variants identified - diagnosis was not confirmed

- pathogenic variant in unanalyzed gene, in noncoding region
- WES, WGS

2. Next generation sequencing (NGS)

Interpretation of sequence variants

Limitations of NGS:

- panel + WES: analysis of coding regions
  sequencing about 95-98% of selected regions
- occurrence of pseudogene, regions with high similarity: difficult non-specific mapping
- is not a suitable method for diseases associated with the expansion / deletion of repetitive sequences

Molecular genetic diagnostics of monogenic diseases

1. Classic Sanger sequencing
2. Next generation sequencing
3. MLPA – CNV detection
4. RP-PCR - detection of repeat expansions
5. Southern blot and hybridization - detection of repeat expansions / deletions

Material: DNA isolated from the whole blood

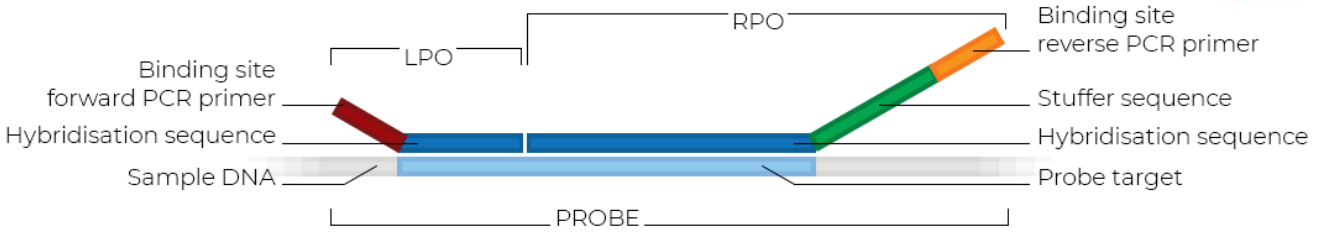
3. MLPA (Multiplex Ligation-dependent Probe Amplification)

- gold standard for CNV detection (whole exons deletions/duplications)
- targeted analysis of a specific gene/genes
- available for certain genes

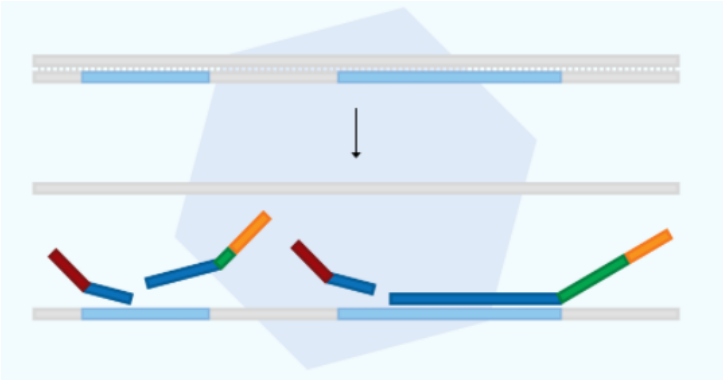
3. MLPA (Multiplex Ligation-dependent Probe Amplification)

Method description:

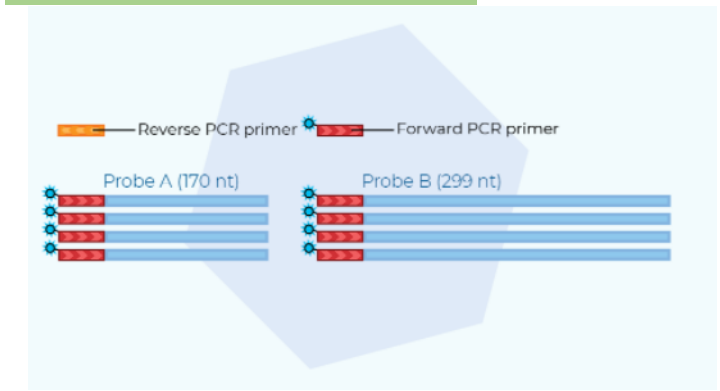
MLPA probe



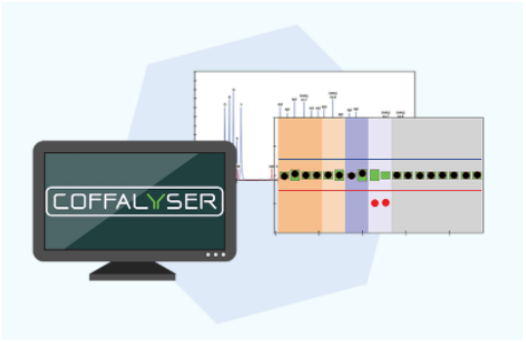
1. Sample denaturation and probe hybridisation



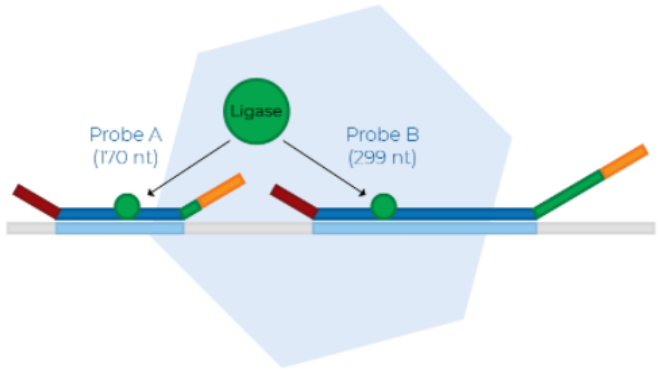
3. Probe amplification



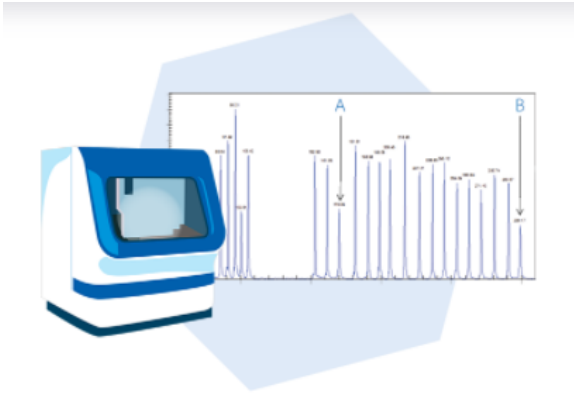
5. Data analysis

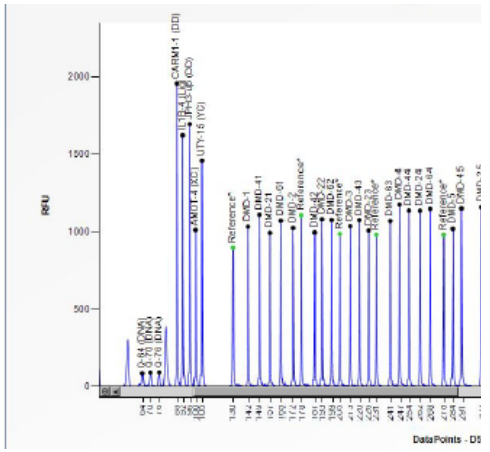


2. Probe ligation



4. Fragment separation





3. MLPA (Multiplex Ligation-dependent Probe Amplification)

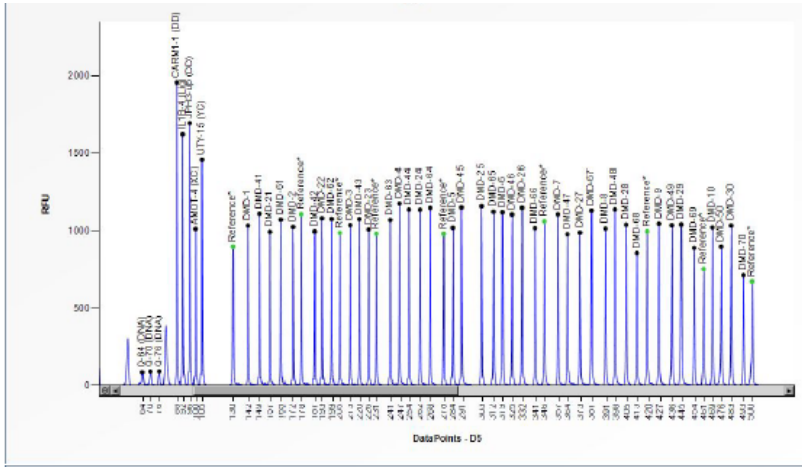
First example:

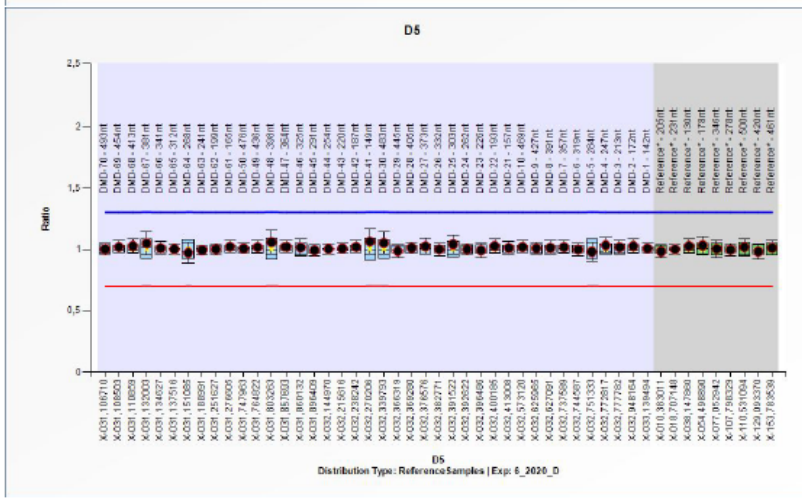
Boy has delayed motor function acquisition and proximal muscle weakness. Due to severely elevated creatine kinase levels (15,000 U/L), neurologist suspects a clinical diagnosis of Duchenne muscular dystrophy.

Duchenne muscular dystrophy, gene *DMD* (chromosome X):

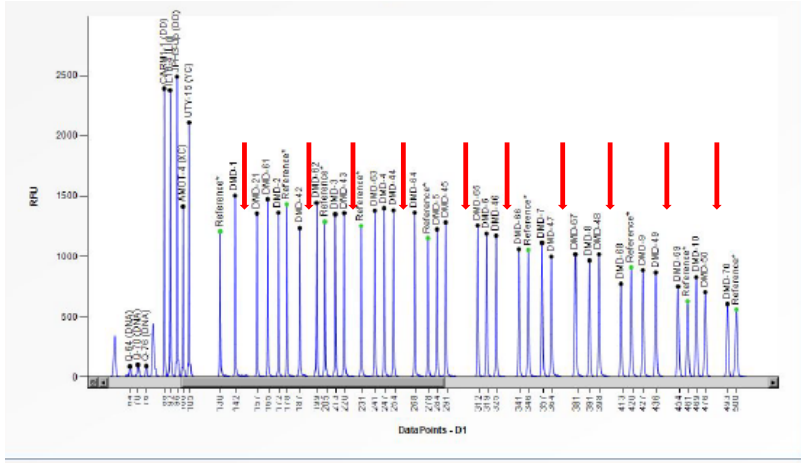
- whole exon deletions (68%) and duplications (10%)
- MLPA is the first choice method
- man = affected
  woman = carrier

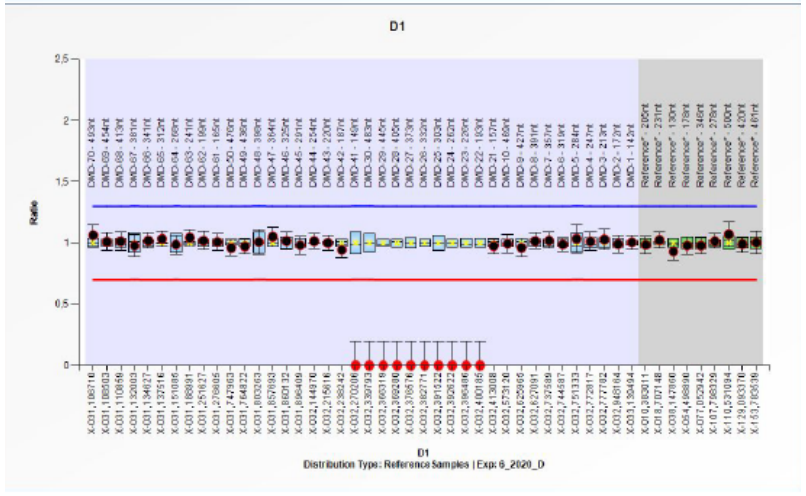
control





10 exons deletion, hemizygous, man





Result:

due to the large DMD size, the MLPA has two parts, the second part is not shown

- deletion of exons 22-41 hemizygous
- out-of-frame deletion leading to creation premature stop codon
- confirmed diagnosis of Duchenne muscular dystrophy

3. MLPA (Multiplex Ligation-dependent Probe Amplification)

First example:

Duchenne muscular dystrophy, gene *DMD* (chromosome X):

- whole exon deletions (68%) and duplications (10%)
- MLPA is the first choice method
- man = affected
  woman = carrier

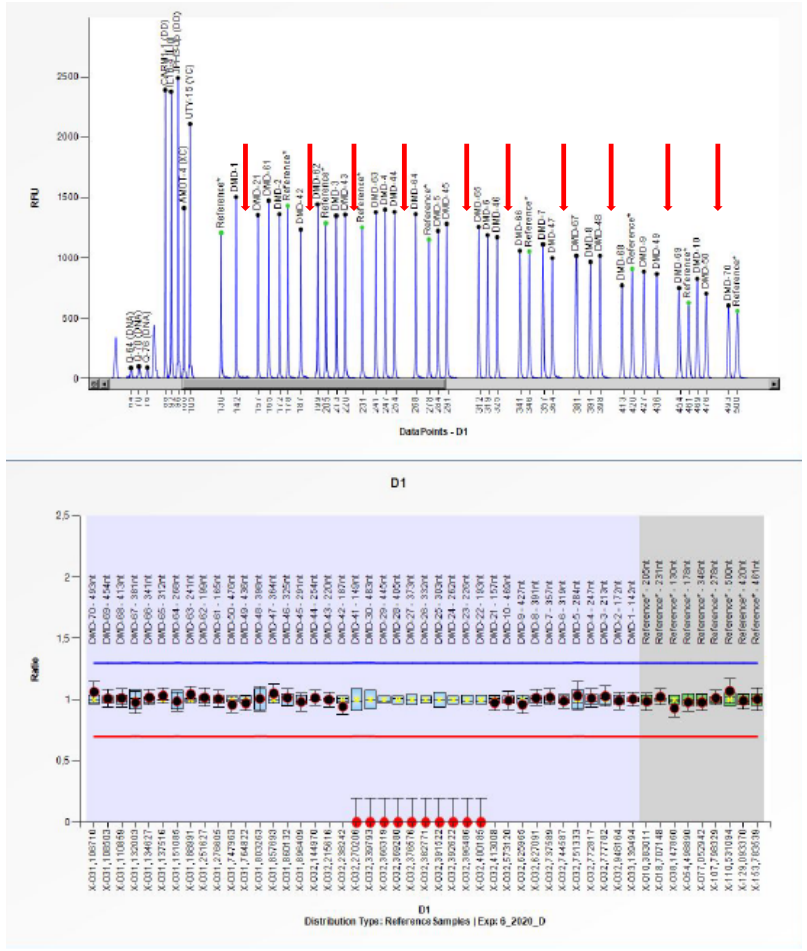
Boy's mother

Result:

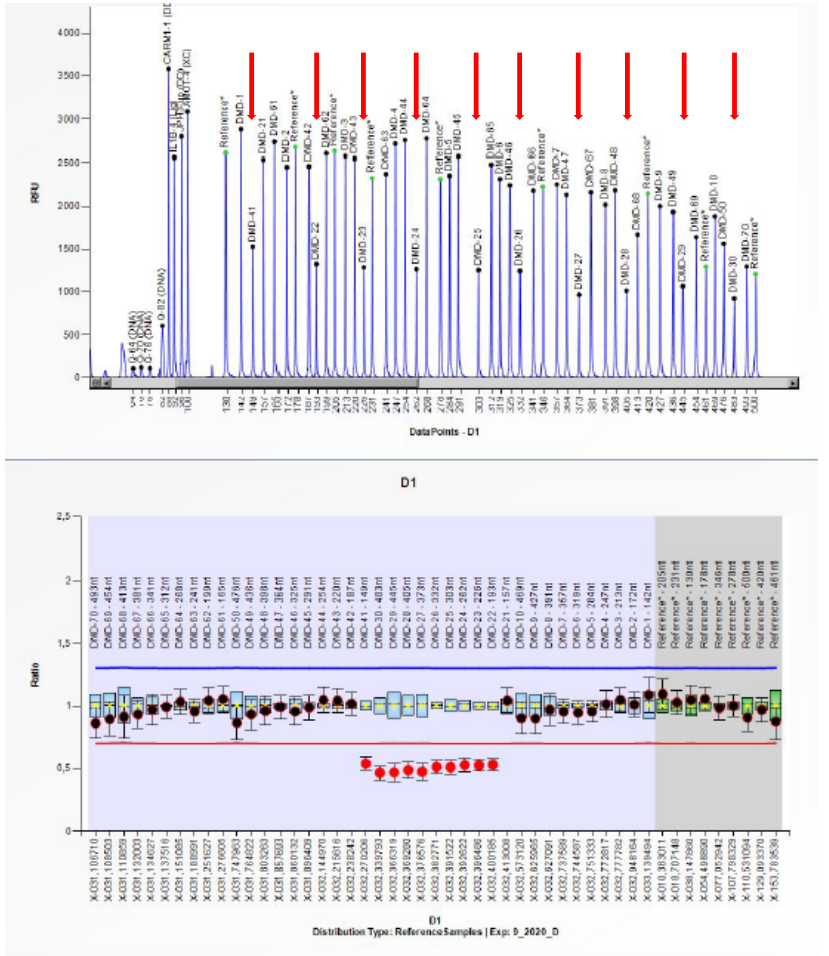
due to the large DMD size, the MLPA has two parts, the second part is not shown

- deletion of exons 22-41 heterozygous
- out-of-frame deletion leading to creation premature stop codon
- mother is carrier of Duchenne muscular dystrophy

10 exons deletion, hemizygous, man



10 exons deletion, heterozygous, woman

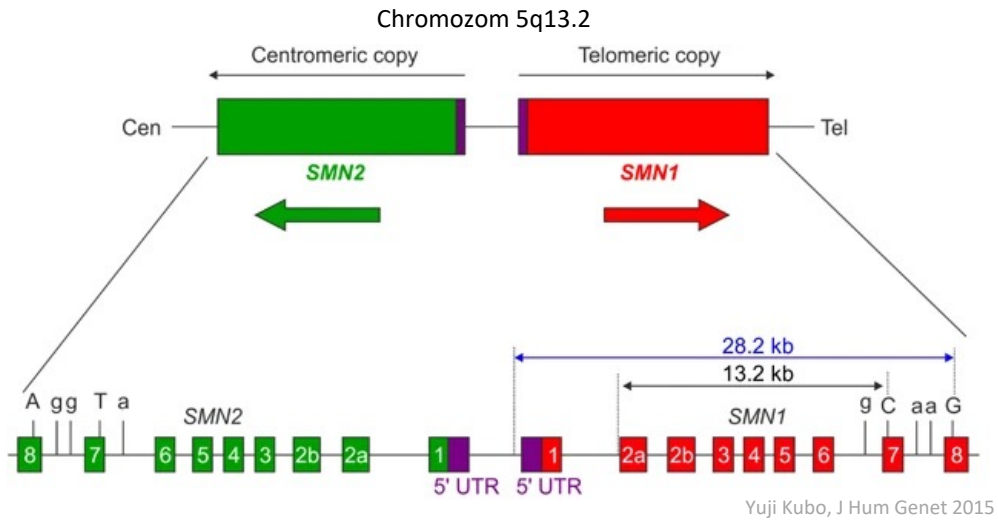


3. MLPA (Multiplex Ligation-dependent Probe Amplification)

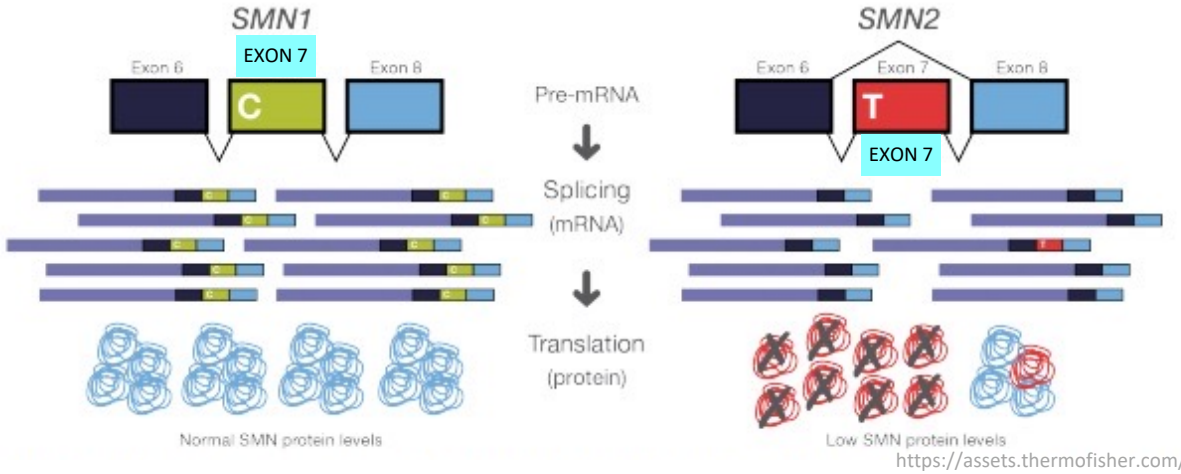
Second example: **Spinal muscular atrophy (SMA)**, gene *SMN1*

- autosomal recessive disease
- incidence: 1 in 6,000 - 10,000 live births
- second most frequent fatal disease with autosomal recessive inheritance (after cystic fibrosis)
- characterized by degeneration of alpha motor neurons
- newborn screening

- 95% caused by homozygous deletion of the *SMN1* gene
- *SMN1* has its almost identical copy – *SMN2* gene (*SMN1* and *SMN2* are homologous to except for few nucleotides)
- copy number variation of *SMN1* and *SMN2* in human genome
- clinical severity is modified by copy number the SMN2 gene



Yuji Kubo, J Hum Genet 2015



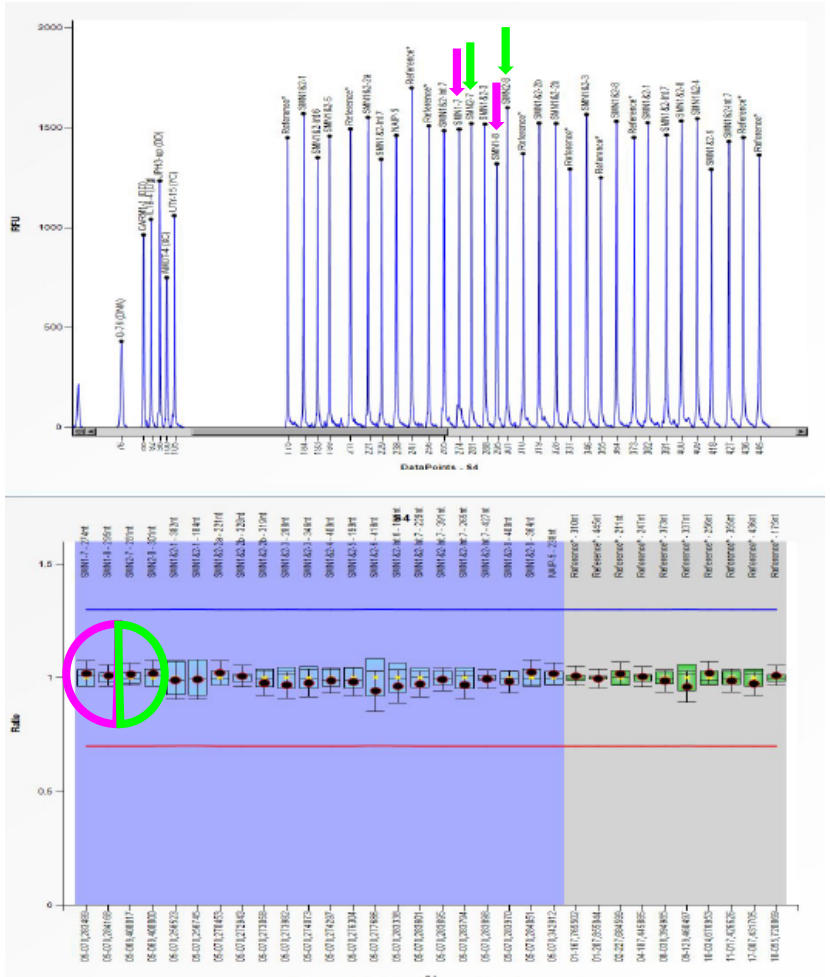
https://assets.thermofisher.com/

3. MLPA (Multiplex Ligation-dependent Probe Amplification)

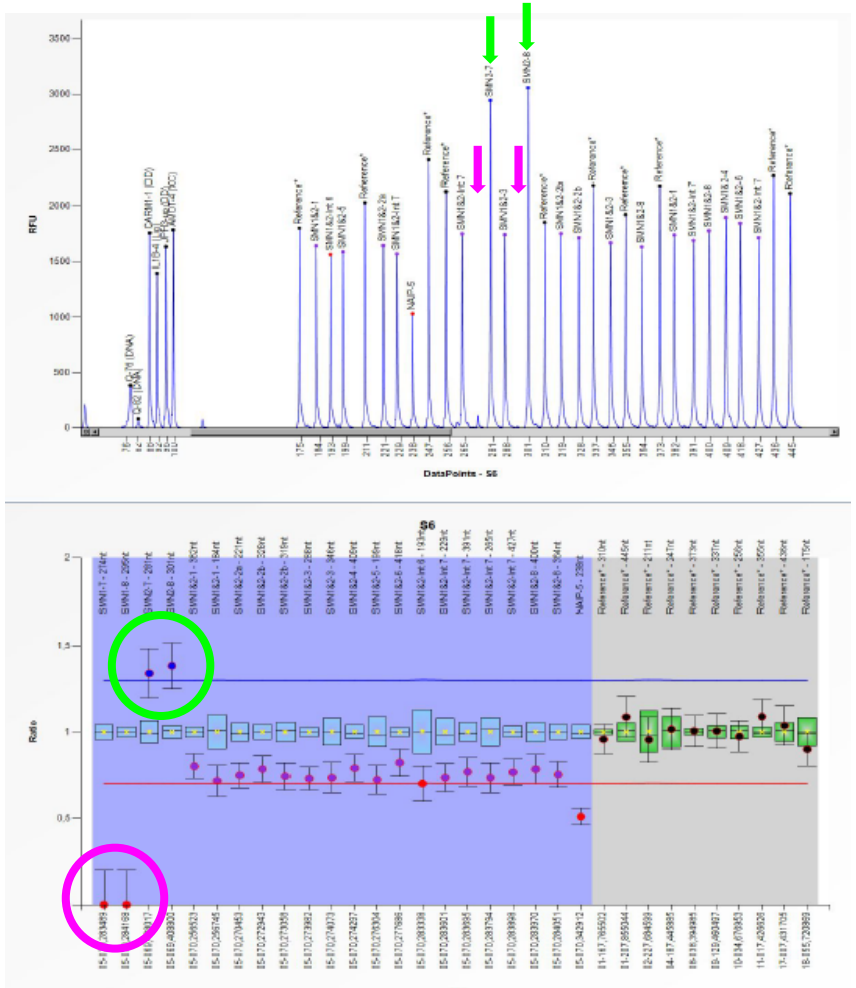
Second example: Spinal muscular atrophy (SMA), *SMN1*

- patients with homozygous *SMN1* deletion
- heterozygous carriers of *SMN1* deletion
- *SMN2* copy number

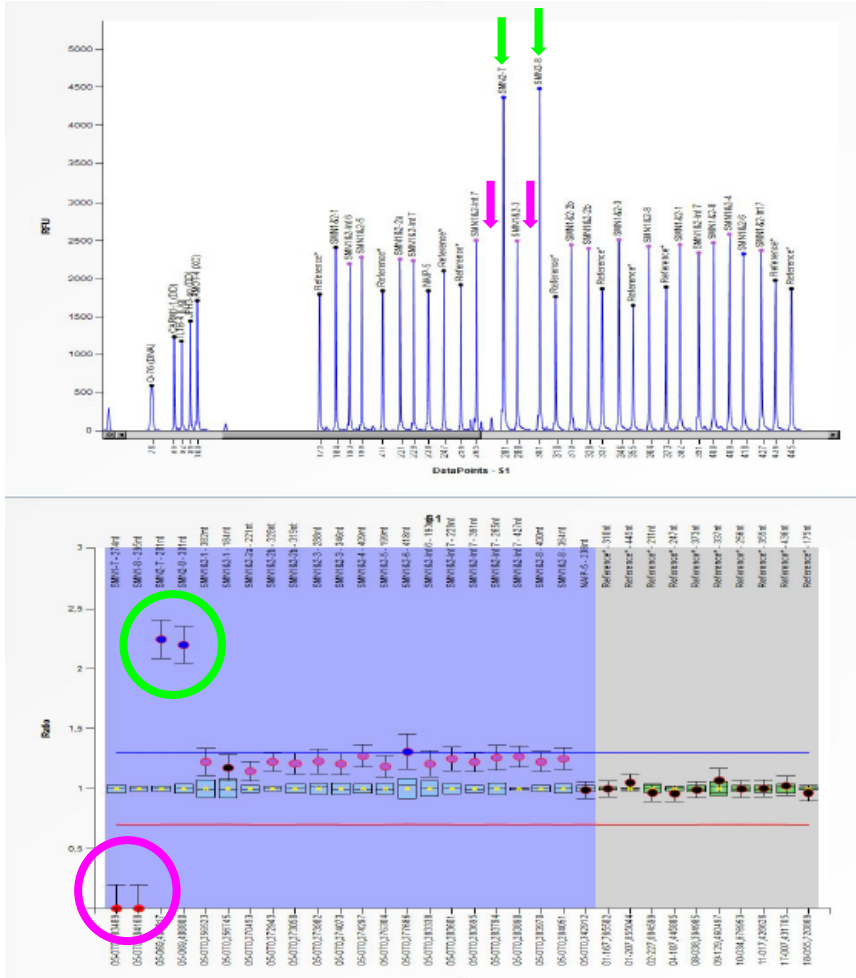
Control: 2x SMN1, 2x SMN2



SMA: 0x SMN1, 3x SMN2



SMA: 0x SMN1, 5x SMN2



Molecular genetic diagnostics of monogenic diseases

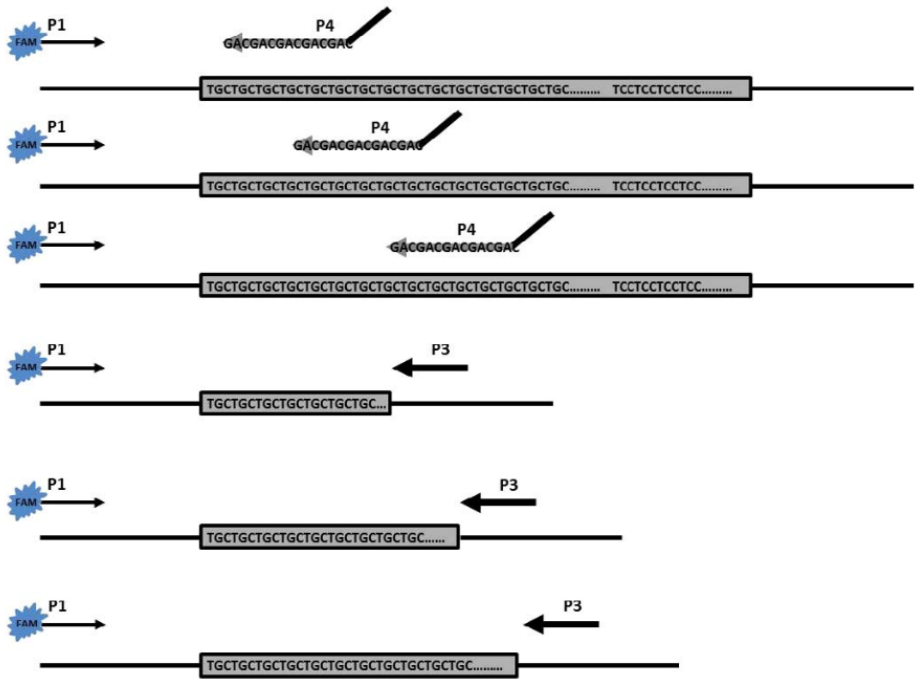
1. Classic Sanger sequencing
2. Next generation sequencing
3. MLPA – CNV detection
4. RP-PCR - detection of repeat expansions
5. Southern blot and hybridization - detection of repeat expansions / deletions

Material: DNA isolated from the whole blood

4. RepeatPrimed-PCR (RP-PCR)

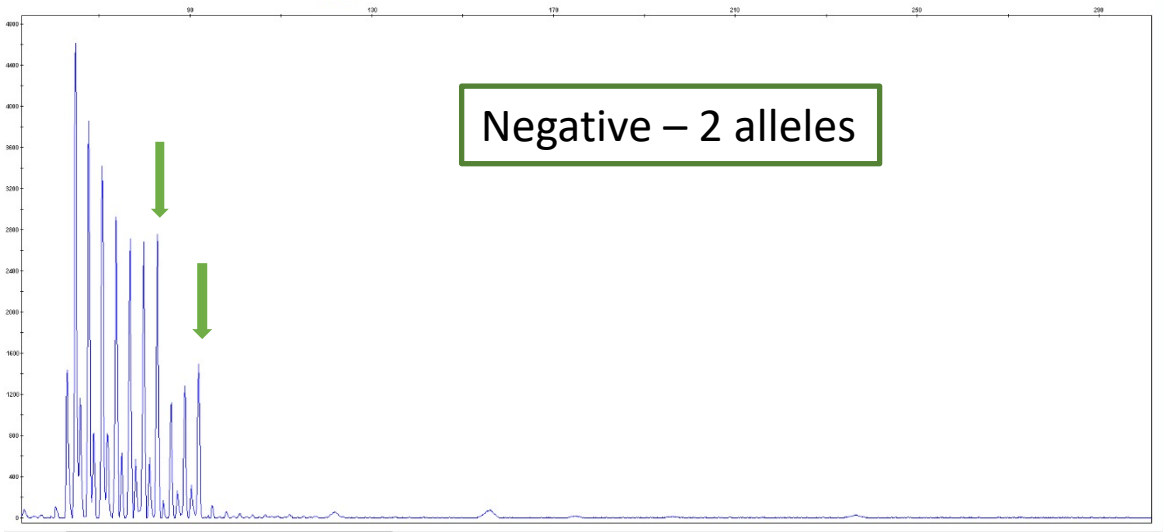
- detection of repeat expansions (usually three nucleotides)
- presence / absence of expansion
- does not determine length of expansion (number of repeats)

- ***Method description:***
  - PCR with three primers
  - characteristic profile of amplicons of increasing length, which differ by the length of a repeat unit (3 bp)

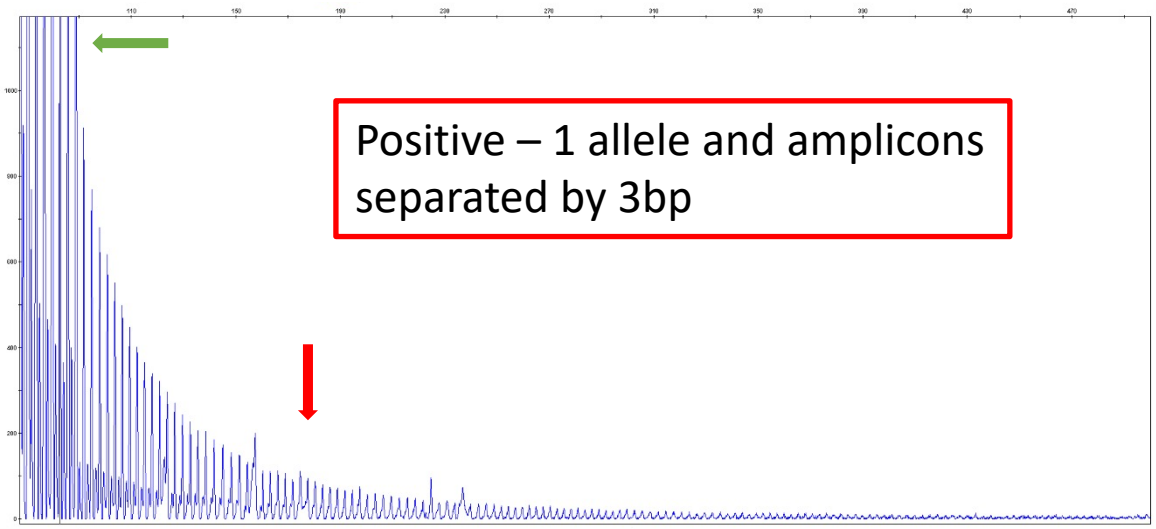


V. Mootha, Inv. ophthal. & vis. science 2014

RP-PCR and fragment analysis:



Negative – 2 alleles



Positive – 1 allele and amplicons separated by 3bp

4. RepeatPrimed-PCR (RP-PCR)

Example:

Patient is displaying subtle signs of myotonia as well as ptosis and weak eye lid closure. Neurologist requested testing for myotonic dystrophy 1.

Myotonic dystrophy type 1

- expansion of CTG repeat in 3’UTR of *DMPK* (9q13.32)
- autosomal dominant inheritance
- correlation between number of repeats and severity of the phenotype:
  - 5-37 repeats - unaffected
  - 38–50 repeats - premutation, asymptomatic
  - 51–149 repeats - mild adult-onset form
  - 150–1000 repeats - classic MD1
  - >1000 repeats - congenital form MD1

20x CTG

```
tccgcggccg gcgaacgggg ctcgaagggt ccttgtagcc gggaatgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgggg ggatcacaga
ccatttcttt ctttcggcca ggctgaggcc ctgacgtgga tgggcaaact gcaggcctgg
```

140x CTG

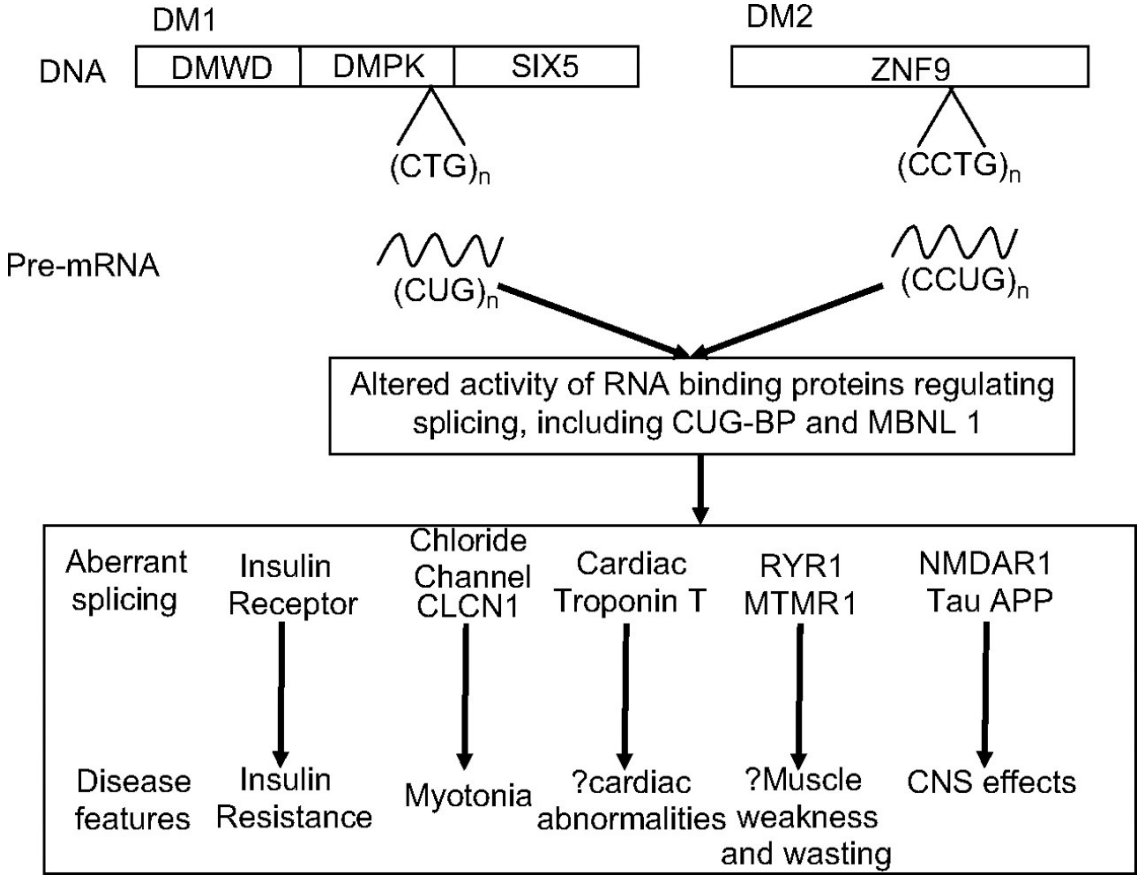
```
tccgcggccg gcgaacgggg ctcgaagggt ccttgtagcc gggaatgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgctg ctgctgctgc
tgctgctgct gctgctgctg ctgctgctgc tgctgctgct gctgctgggg ggatcacaga
ccatttcttt ctttcggcca ggctgaggcc ctgacgtgga tgggcaaact gcaggcctgg
```

- anticipation - 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.

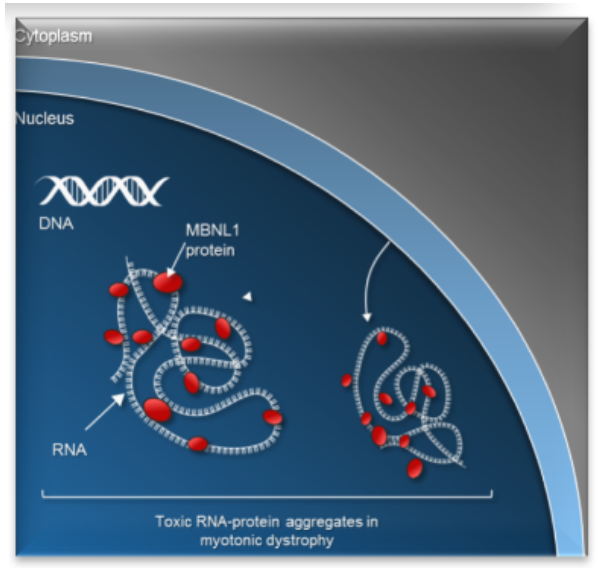
4. RepeatPrimed-PCR (RP-PCR)

Myotonic dystrophy type 1 – mechanism:

- toxic effect of expansion
- accumulation of RNA with expansions in the nucleus, sequestration of RNA-binding protein > formation of nuclear inclusions
- **altering mRNA splicing of other genes**



Turner, Journal of Neurology, Neurosurgery & Psychiatry 2010



Mignon, IONIS-DMPK Clinical Program in Myotonic Dystrophy

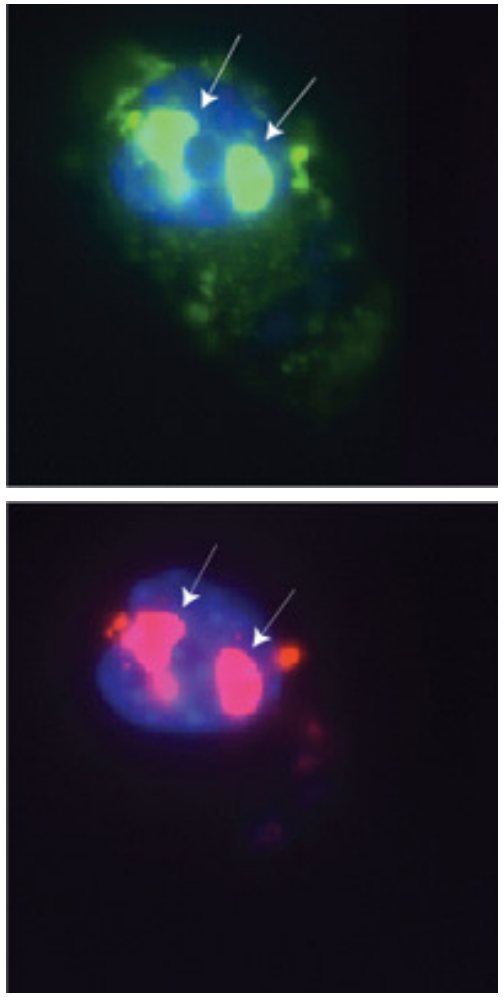


Image shows the location of the Mbnl1 splicing factor (green) and the second image shows the location of RNA repeats (red) inside the cell nucleus (blue). The white arrows point to two large foci in the cell nucleus where Mbnl1 is sequestered with RNA. Photos by Hongqing Du

4. RepeatPrimed-PCR (RP-PCR)

Example:

Patient is displaying subtle signs of myotonia as well as ptosis and weak eye lid closure. Neurologist requested testing for myotonic dystrophy 1.

Myotonic dystrophy type 1 – result:

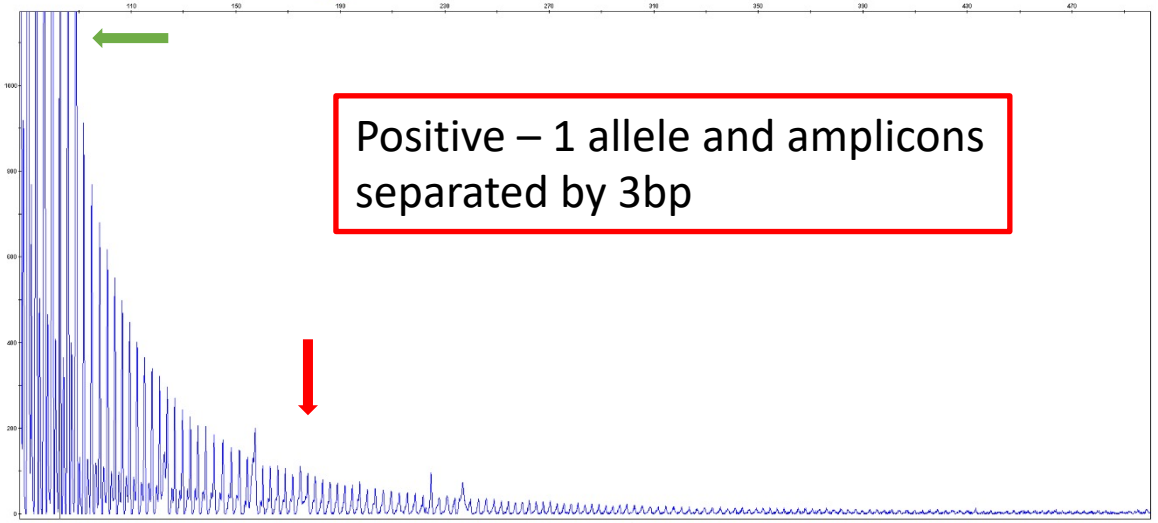
- presence of expansion

confirmed diagnosis of myotonic dystrophy type 1

- **not the length**

Solution: Southern blot and hybridization

RP-PCR and fragment analysis:



Molecular genetic diagnostics of monogenic diseases

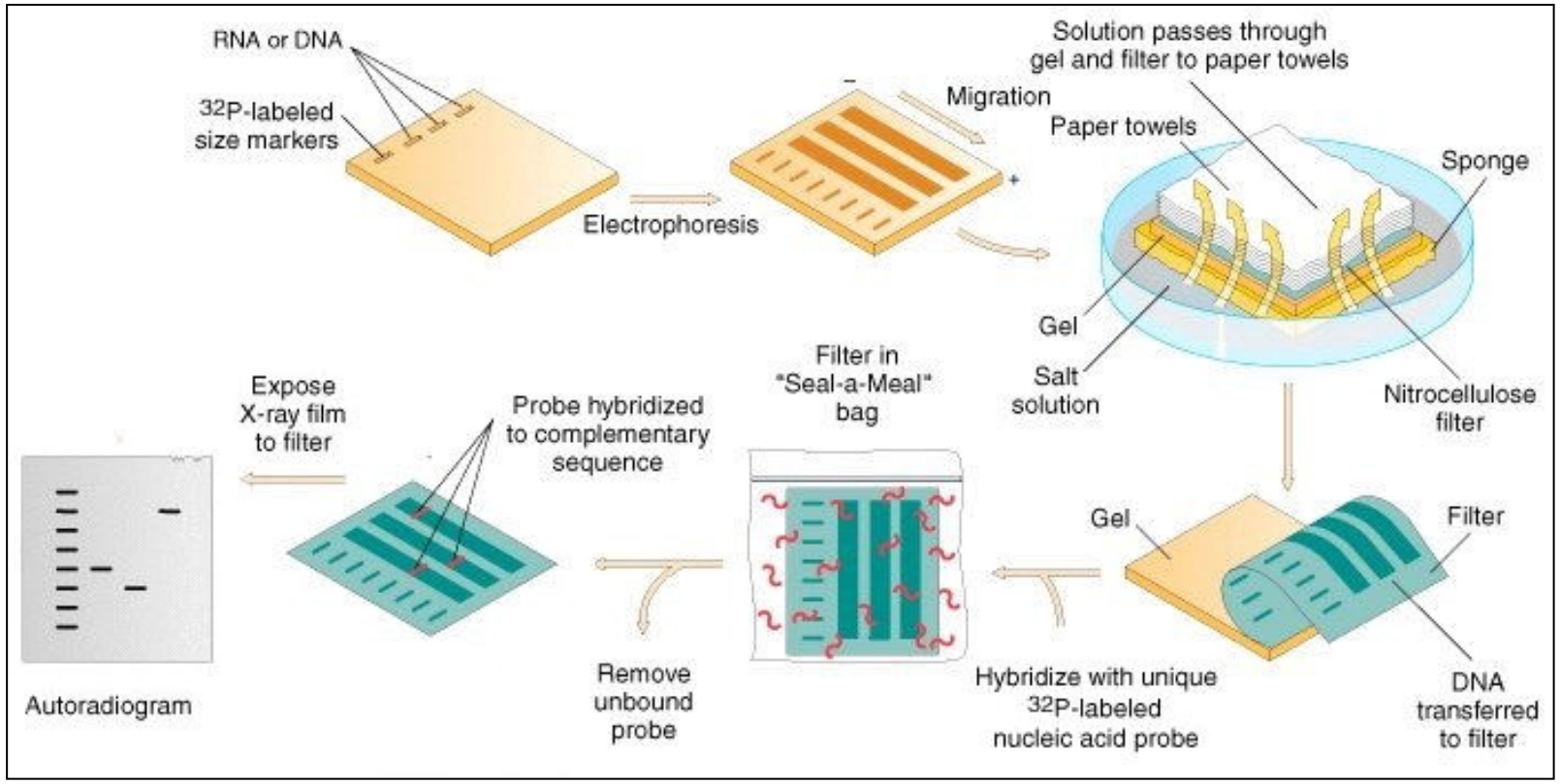
1. Classic Sanger sequencing
2. Next generation sequencing
3. MLPA – CNV detection
4. RP-PCR - detection of repeat expansions
5. Southern blot and hybridization - detection of repeat expansions / deletions

Material: DNA isolated from the whole blood

4. Southern blot and hybridization

- detection of repeat expansions / deletions
- determination of the size

- ***Method description:***
  - DNA is cleaved by a restriction endonuclease
  - electrophoresis
  - transfer to membrane
  - hybridization with radioactive labeled probe
  - autoradiography

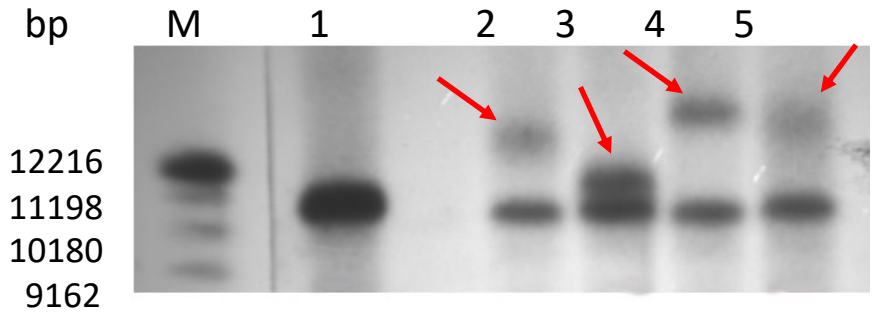


4. Southern blot and hybridization

Example:

Myotonic dystrophy type 1 – result:

- presence of expansion
- **not the length**



MD1: 1 – negative control
2-5 – expansion

- according to the size of the fragment, we determine the number of repeats

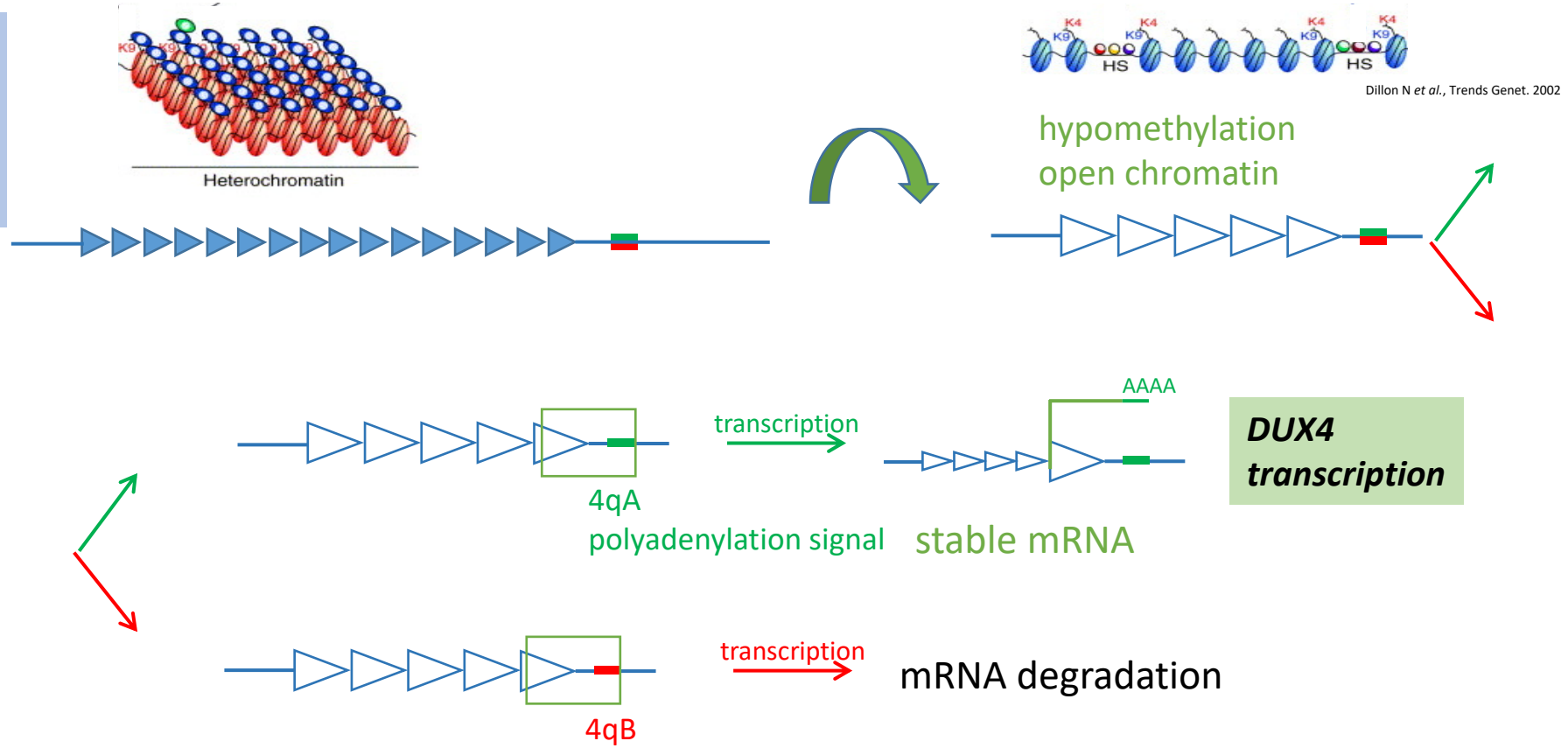
4. Southern blot and hybridization

Example:

Facioscapulohumeral dystrophy 1 (FSHD1)

- the third most prevalent muscular dystrophy, AD inheritance
- weakness and wasting of the face, shoulder and upper arm muscles, with later involvement of the trunk and lower extremities
- FSHD develops through complex genetic and epigenetic events that converge on a common mechanism of toxicity with mis-expression of the transcription factor DUX4

- 4q35
- repeats D4Z4 (contain DUX4 gene)
- 11-100 repeats → heterochromatin
- 1-10 repeats → chromatin conformational changes, hypomethylation

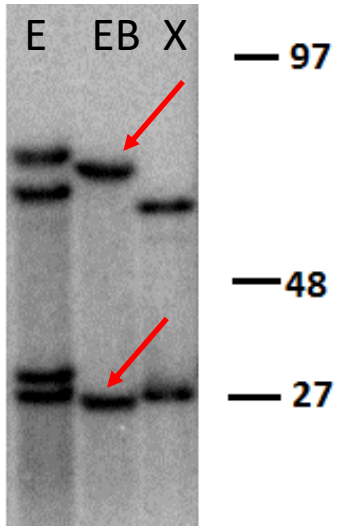


4. Southern blot and hybridization

Example:

Facioscapulohumeral dystrophy 1 (FSHD1)

- 4q35, repeats D4Z4 (contain DUX4)
- 11-100 repeats → unaffected
- 1-10 repeats → affected



we determine the number of D4Z4 repeats according to the size of the product

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

Example of specific treatment - certain disease, certain mutation

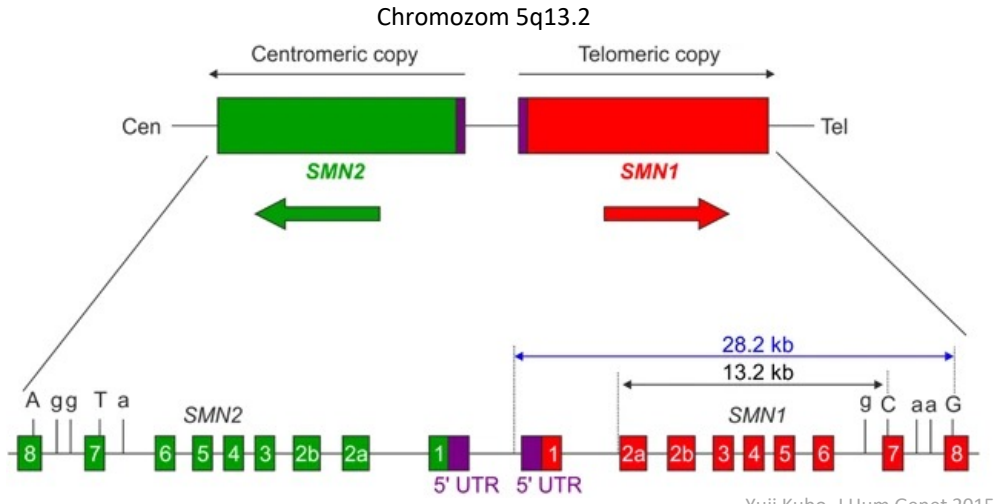
Spinal muscular atrophy (SMA)

Spinal muscular atrophy (SMA)

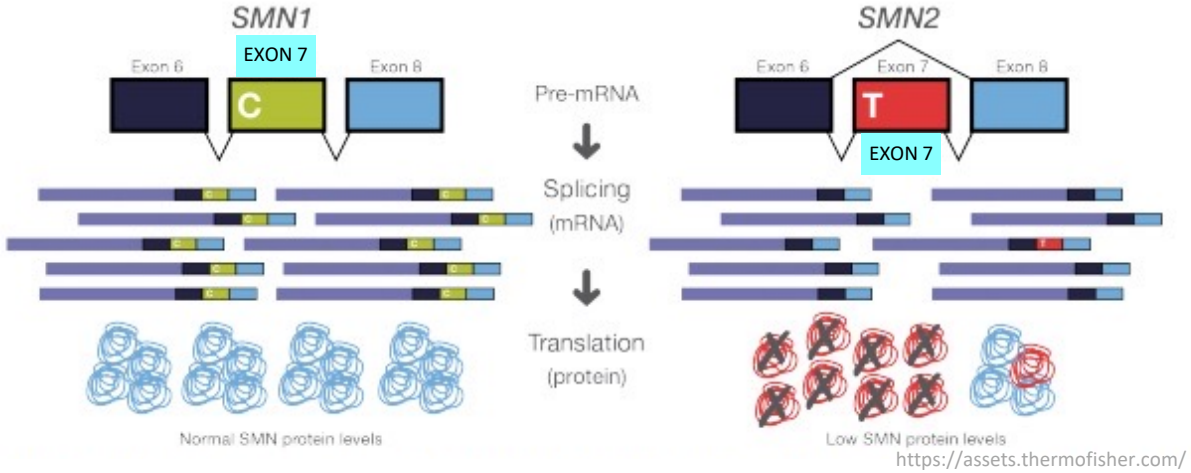
- gene *SMN1,* autosomal recessive disease
- incidence: 1 in 6,000 - 10,000 live births
- second most frequent fatal disease with autosomal recessive inheritance
- characterized by degeneration of alpha motor neurons
- **newborn screening – started last year**

- 95% caused by homozygous deletion of the *SMN1* gene
- *SMN1* has its almost identical copy – *SMN2* gene (*SMN1* and *SMN2* are homologous to except for few nucleotides)
- copy number variation of *SMN1* and *SMN2* in human genome

Not enough SMN protein, so the motor neurons shrink and die. As a result, the brain can't control voluntary movements, especially motion in the head, neck, arms and legs.



Yuji Kubo, J Hum Genet 2015

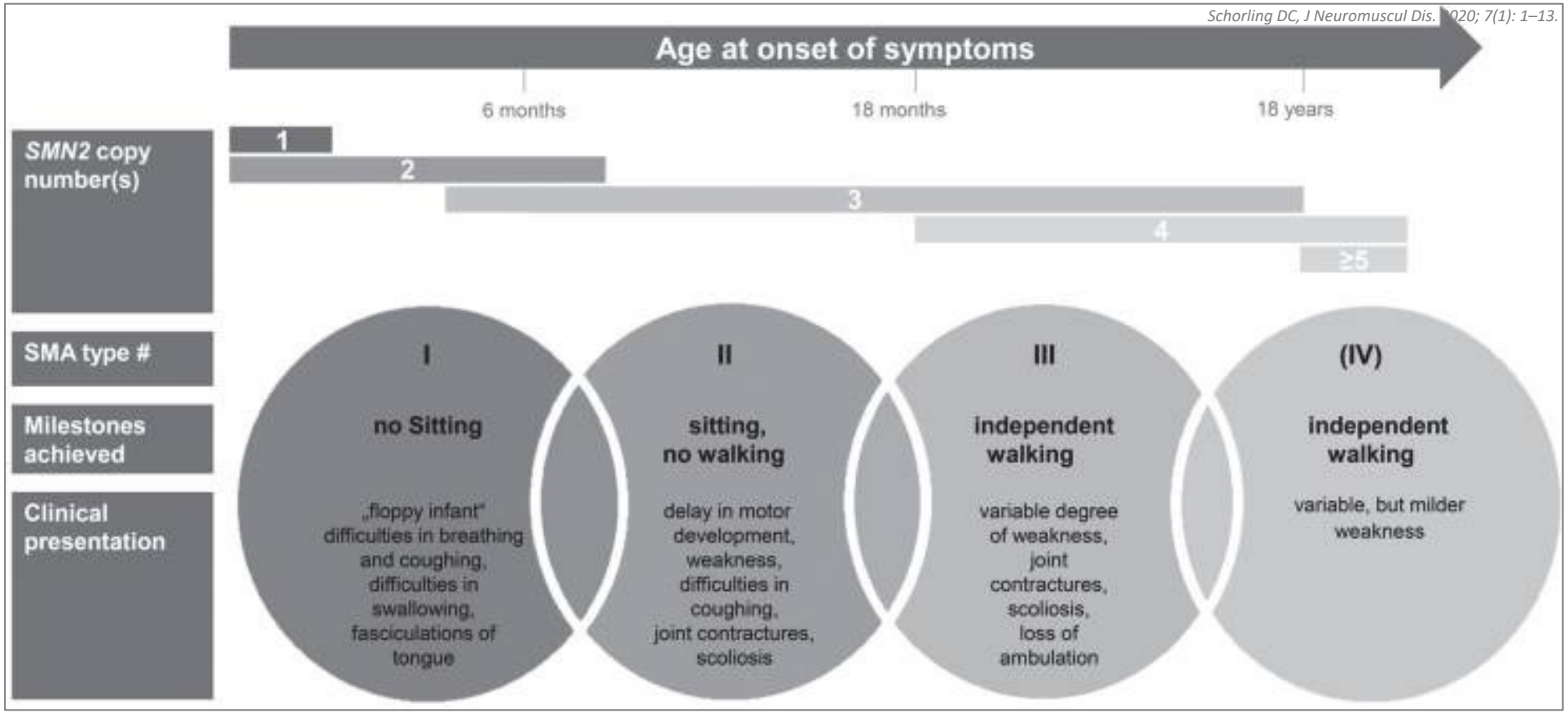


https://assets.thermofisher.com/

Clinical severity is modified by copy number the *SMN2* gene

Spinal muscular atrophy

4 clinical types of SMA



Type I – 60% of SMA, severe muscle weakness and hypotonia at birth or within the first 6 months; death from respiratory failure occurs usually within the first 2 years

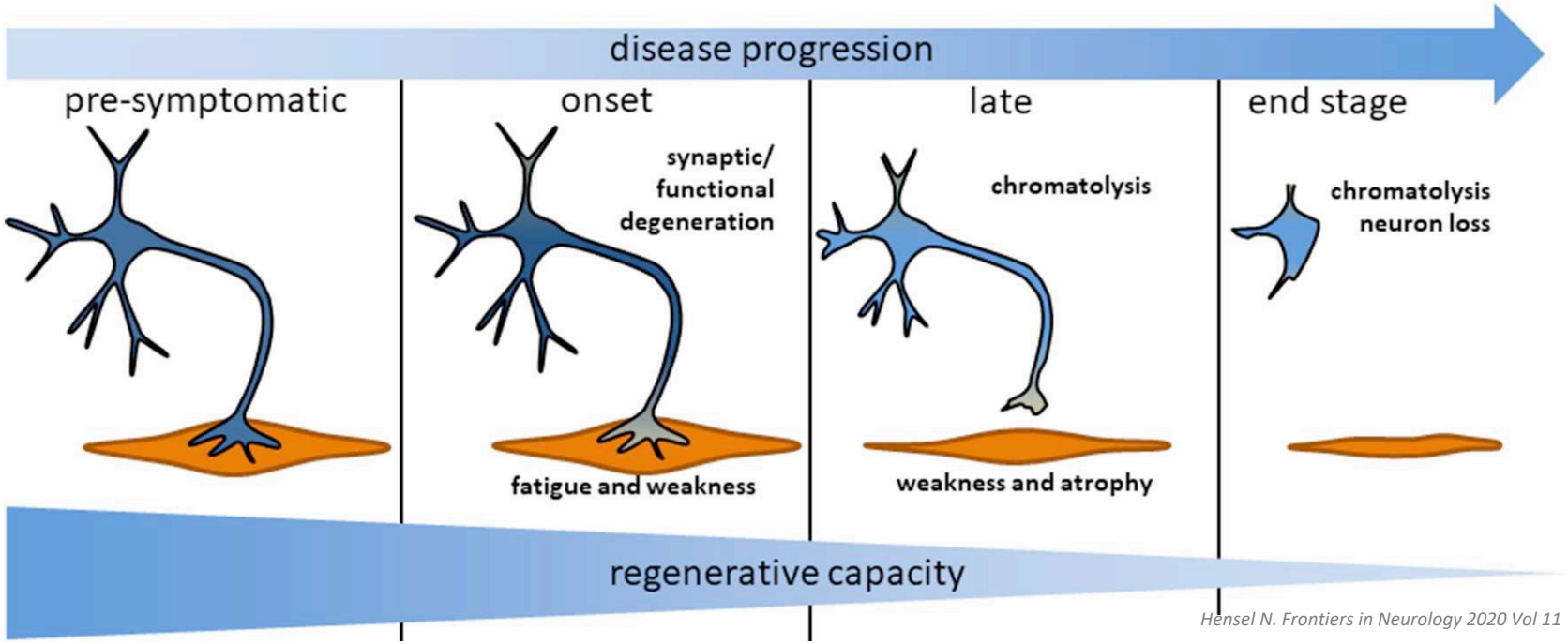
Type II – first symptoms begin 6-18 months, live into adulthood; patients are able to sit but unable to walk independently

Type III - first symptoms after 2 years of life; patients are able to walk but often wheelchair-bound; no significantly shorten life expectancy

Type IV – rare form, symptoms appear in adulthood; patients have mild motor impairment

Spinal muscular atrophy

Motoneuron degeneration in SMA



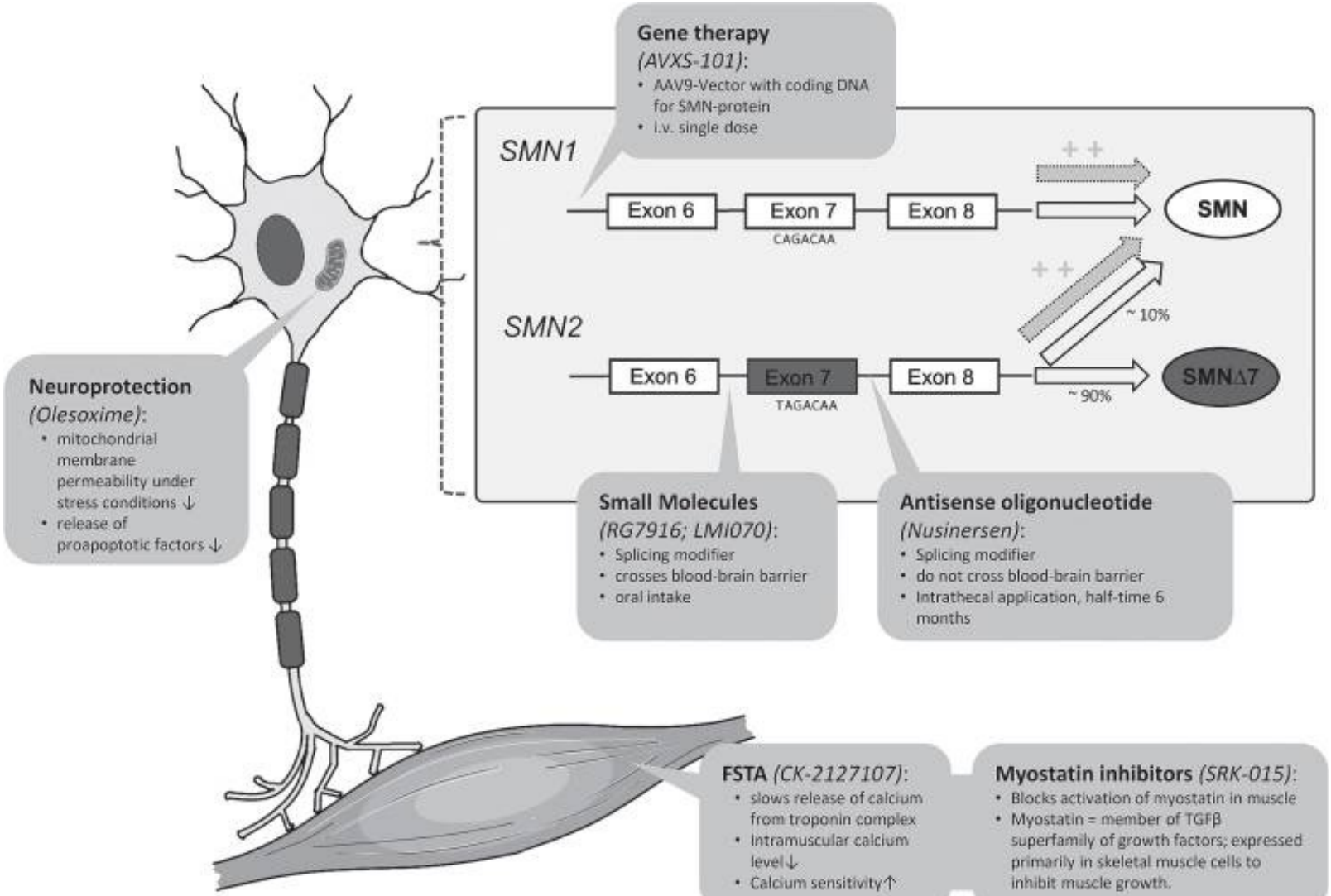
Hensel N. Frontiers in Neurology 2020 Vol 11

- functional degeneration of central synapses and neuromuscular junctions and subsequent axonal damage > motoneuron loss
- **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.**

Spinal muscular atrophy

SMA therapies

1. **modifying splicing of *SMN2* (production of more amount of full length mRNA)**
2. **replacing the *SMN1* gene**



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

Spinal muscular atrophy

SMA therapies

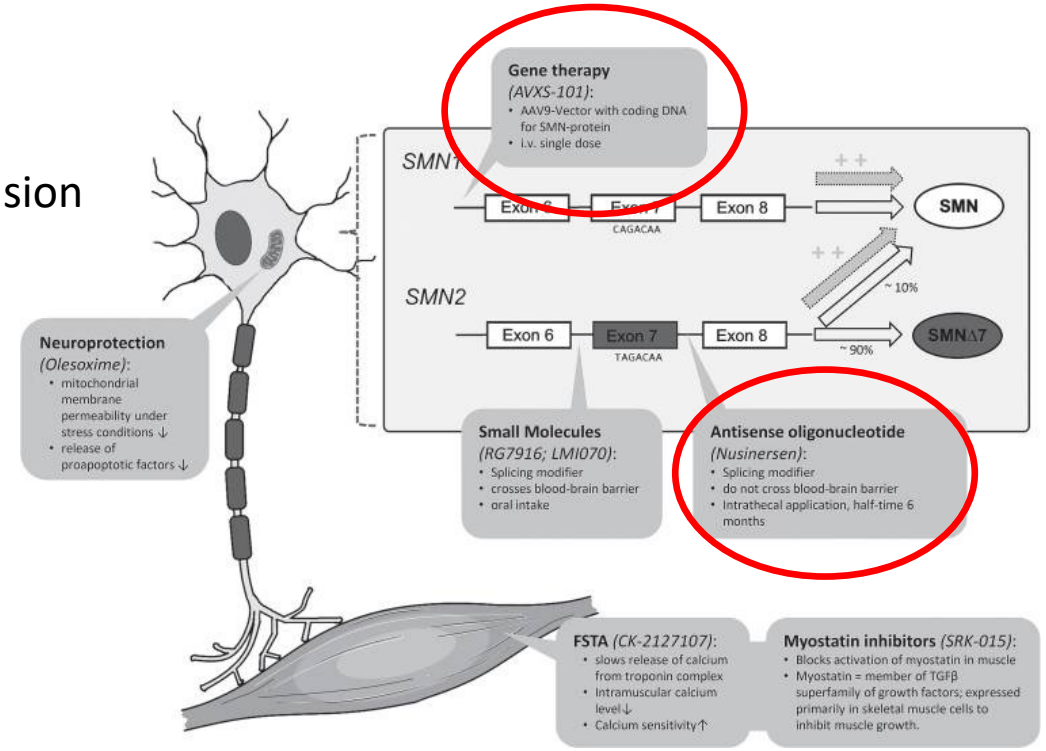
1. **modifying splicing of *SMN2* (production of more amount of full length mRNA)**

A. Nusinersen (Spinraza®)

- an antisense-oligonucleotide (ASO) that enhances the inclusion of exon 7 in mRNA transcripts of SMN2
- administered intrathecally

B. Risdaplam (Evrysdi®)

- administered orally



2. **replacing the *SMN1* gene**

C. Onasemnogene Abeparvovec-xioi (Zolgensma®)

- children younger than two
- one-time intravenous infusion
- adeno-associated virus 9 (AAV9) delivering cDNA which codes the full length SMN protein
- = replacement of a missing or faulty *SMN1* gene with a functioning gene

Spinal muscular atrophy

SMA neonatal screening pilot project in Czech republic

- early detection of neonates in the preclinical asymptomatic stage
- treatment before irreversible complete loss of motoneuron

QUESTIONS?