Molecular genetic studies on nemaline myopathy and related disorders

Vilma-Lotta Lehtokari

The Department of Medical Genetics,
University of Helsinki,
Helsinki, Finland
&
The Folkhälsan Institute of Genetics

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine of the University of Helsinki, for public examination in the Niilo Hallman Lecture Hall, Children's Hospital, Hospital District of Helsinki and Uusimaa, Stenbäckinkatu 11, Helsinki on 8th June 2009, at 12 o'clock

Helsinki 2009

Supervisors:
Docent Carina Wallgren-Pettersson, D.M.
The Folkhälsan Institute of Genetics and
Department of Medical Genetics
University of Helsinki
Finland

Docent Katarina Pelin, PhD
Division of Genetics
Department of Biological and Environmental Sciences
University of Helsinki
Finland

Reviewers:
Docent Marjo Kestilä, PhD
Department of Chronic Disease Prevention
Public Health Genomics
National Institute For Health and Welfare
Helsinki
Finland

Docent Mikaela Grönholm, PhD
Division of Biochemistry
Department of Biological and Environmental Sciences
University of Helsinki
Finland

Official opponent:
Professor Jaakko Ignatius, D.M.
Department of Clinical Genetics
University of Oulu
Finland

Contents

List of original publications

This thesis is based on the following publications. In addition, some unpublished results are presented.

I. **Lehtokari V-L**, Pelin K, Sandbacka M, Ranta S, Donner K, Muntoni F, Sewry C, Angelini C, Bushby K, Van den Bergh P, Iannaccone S, Laing N & Wallgren-Pettersson C. Identification of 45 novel mutations in the nebulin gene associated with autosomal recessive nemaline myopathy. *Hum Mutat.* 27(9):946-956, 2006

II. **Lehtokari V-L*,** Greenleaf R*, Dechene E, Kellinsalmi M, Pelin K, Laing N, Beggs A, Wallgren-Pettersson C. The exon 55 deletion in the nebulin gene – One single founder mutation with world-wide occurrence. *Neuromuscul Disord*. 19(3):179-81, 2009

III. Wallgren-Pettersson C, **Lehtokari V-L**, Kalimo H, Paetau A, Nuutinen E, Hackman P, Sewry C, Pelin K, Udd B. Distal myopathy caused by homozygous missense mutations in the nebulin gene. *Brain.* 130:1465-1476, 2007

IV. **Lehtokari V-L**, Ceuterick - de Groote C, de Jonghe P, Marttila M, Laing N, Pelin K, Wallgren-Pettersson C. Cap disease caused by heterozygous deletion of the beta-tropomyosin gene *TPM2. Neuromuscul Disord.* 17(6):433-442, 2007

V. **Lehtokari V-L**, Pelin K, Donner K, Voit T, Rudnik-Schöneborn S, Stoetter M, Talim B, Topaloglu H, Laing N, Wallgren-Pettersson C. Identification of a founder mutation in *TPM3* in nemaline myopathy patients of Turkish origin*. Eur J Hum Genet.*16(9):1055-61, 2008

The publications are referred to in the text by their Roman numerals.

*The authors contributed equally to the work.

The articles are reprinted with the permission of the copyright owners.

Abbreviations

Abbreviations of amino acids in appendix 1	
A	adenine
ACTA1	the gene encoding skeletal muscle specific α-actin
ACTN2 / 3	the gene encoding α-actinin 2 / 3
AD	autosomal dominant inheritance
ADP	adenosine diphosphate
AR	autosomal recessive inheritance
ATP	adenosine triphosphate
bp	base pair
C	cytosine
Ca^{2+}	calcium ion
CAPZA1 / 2	the gene encoding capping protein α1 / α2
CAPZB	the gene encoding capping protein β
cDNA	complementary DNA
CEPH	Centre d´Etude du Polymorphisme Humain
CFL2	the gene encoding cofilin 2
CFTD	congenital fibre type disproportion
CM	cap myopathy
cM	centi Morgan
CNV	copy number variation
DA	distal arthrogryposis
DES	the gene encoding desmin
dHPLC	denaturing high performance liquid chromatography
DMD	Duchenne muscular dystrophy
DNA	deoxyribonucleic acid
DNM	distal nebulin myopathy

e.g.	exempli gratia
EM	electron microscope
ESE	exonic splicing enhancers
ESS	exonic splicing silencers
etc.	et cetera
FNLC	the gene encoding filamin C
G	guanine
HE	hematoxylin & eosin (stain)
HGM	the human genome project
i.e.	id est
kb	kilobase
kDa	kiloDalton
KI	knock-in (mutation pointed to a specific gene)
KO	knock-out (gene function silenced)
LDM	Laing distal myopathy
LGMD	limb girdle muscular dystrophy
LM	light microscope
LMNA	the gene encoding lamin A/C
Mg^{2+}	magnesium ion
MLPA	multiplex ligation-dependent probe amplification
mm	millimeter
MYO	the gene encoding myotilin
μm	micrometre
mRNA	messenger ribonucleic acid
MyBP-C	Myosin-binding protein C
MYH	myosin heavy chain protein encoding gene(s)
MyHC	myosin heavy chain
MYL	myosin light chain
MYPN	myopalladin
NEB	the gene encoding nebulin

NEBL	the gene encoding nebulette	*TCAP*	the gene encoding titin capping protein
n	the number of		
nm	nanometre	TMD	tibial muscular dystrophy
NM	nemaline myopathy	*TMOD4*	the gene encoding tropomodulin 4
nt	nucleotide		
OBSCN	the gene encoding obscurin	*TNNC*	the gene(s) encoding troponin C (calcium binding) isoform(s)
p.	page		
PCR	polymerase chain reaction		
pH	pondus hydrogenii	*TNNI*	the gene(s) encoding troponin I (inhibitor) isoform(s)
RNA	ribonucleic acid		
RT-PCR	reverse transcriptase polymerase chain reaction		
		TNNT	the gene(s) encoding troponin T (tropomyosin binding) isoform(s)
RYR1	the gene encoding ryanodine receptor 1		
SEPN1	the gene encoding selenoprotein 1	*TPM1*	the gene encoding α-tropomyosin$_{fast}$
SERCA	sarcoplasmic reticulum Ca^{2+}-ATPase	*TPM2*	the gene encoding β-tropomyosin
SH3	Src homology domain	*TPM3*	the gene encoding α-tropomyosin$_{slow}$
SLN	sarcolipin		
SNP	single nucleotide polymorphism	*TTN*	the gene encoding titin
		YL1	the gene encoding vacuolar protein sorting 72
SR	sarcoplasmic reticulum		
SRF	the serum response factor	WT	wild type
SSCP	single-stranded conformation polymorphism	*ZASP*	the gene encoding lim domain-binding 3
T	thymine	$\varnothing$	diameter

Abstract

The aim of this thesis was to study the molecular genetics of nemaline myopathy and related disorders, and to investigate the molecular mechanisms by which the identified mutations cause muscle disease. This thesis comprises five publications on the molecular genetics, clinical features and histology of three different muscle disorders: nemaline myopathy, distal nebulin myopathy and cap myopathy. The molecular genetic studies performed led to the identification of mutations in three different genes encoding proteins of the thin filament of the muscle sarcomere: nebulin (*NEB*), and the tropomyosins α_{slow} and β (*TPM3* and *TPM2*). The patients studied exhibited variable clinical and histological features, but the muscle biopsies all displayed disorganised sarcomeric Z discs and/or aggregates of proteins.

Nemaline myopathy (NM) is a clinically and genetically heterogeneous group of disorders diagnosed on the basis of muscle weakness and the presence of protein aggregates called nemaline bodies or rods in the muscle fibres. Several genes are known to cause NM; these are *NEB*, slow skeletal α-actin (*ACTA1*), *TPM3, TPM2*, troponin T1 (*TNNT1*), and cofilin-2 (*CFL2*). In addition to these, there is at least a seventh NM gene yet to be identified. NM is usually the consequence of a gene mutation and the mode of inheritance varies between NM subclasses and different families. The disease can be inherited as an autosomal dominant or as a recessive trait. New dominant mutations in *ACTA1*, *TPM2* and *TPM3* are quite common also. The linkage and mutation analyses performed in this study included all known NM genes as well as several candidate genes. Mutations in *NEB* or *TPM3* were identified and published in 46 NM families. Including unpublished mutations a total of 115 different *NEB* mutations have been identified in 96 families.

Nebulin is a giant structural protein of the sarcomere which is encoded by *NEB* consisting of 183 exons, making mutation analyses demanding. Denaturing High Performance Liquid Chromatography (dHPLC) and sequencing proved to be efficient methods for the identification of heterozygous mutations along the whole length of the gene. In this study, the occurrence of one such mutation, deletion of the whole *NEB* exon 55 seen in the Ashkenazi Jewish population, was studied, and a haplotype segregating with this founder mutation, identified.

In a project utilizing genome-wide and candidate gene analyses, a homozygous deletion disrupting the termination signal was identified in *TPM3* in patients from two Turkish families. This is a likely founder mutation in the Turkish population.

Distal nebulin myopathy (DNM) is a novel disorder identified in four Finnish families during the course of the present study. It is a recessively inherited myopathy causing distal weakness (as opposed to the commonly proximal weakness in NM), and it also differs from

NM histologically; the biopsies of the patients did not display nemaline bodies in routine light microscopy, although some of them had small Z-disc-derived protein aggregates visible under the electron microscope. Two different homozygous mutations leading to the substitution of an amino acid in nebulin were found to underlie DNM. Both of the mutations were known to cause NM in compound heterozygous form, together with another, more disruptive *NEB* mutation. This study showed that *NEB* mutations may cause disorders other than NM.

Cap myopathy (CM) was described as a novel entity already in 1981. It is characterised by massive protein aggregates and disorganised sarcomeres forming cap-like structures under the muscle cell membrane on one side of the fibre. This disorder is variable and may overlap with NM both clinically and histologically. The patients may have nemaline bodies either within the cap structure or elsewhere in the fibre. The heterozygous *de novo* dominant in-frame deletion of one codon in *TPM2* described in the present study was the first genetic cause involved in CM.

NM, DNM and CM patients have variable clinical pictures in terms of both severity, distribution of affected muscles and age of onset. These may differ between and within the disorders as well as between individuals with mutations in the same genes, or even between patients sharing the same mutation. On the other hand, it is notable that patients with different diagnoses or causative genes may exhibit an overlap in their histological and/or clinical features.

The exact molecular mechanisms behind these disorders remain to be elucidated, but it is possible that some of the overlap could be explained by shared pathogenetic pathways. Based on the present study, these mechanisms might include altered interactions of the abnormal proteins with their binding partners within the sarcomere.

Review of the literature

1 Principles of molecular genetics and gene identification

Medicine and biology have always fascinated philosophers and scientists, and even today science relies on observations made during previous centuries: Carl Linnaeus created taxonomy, a system of ecological classification of species and published Systema naturae in 1759.[1, 2] Charles Darwin studied the hypothesis that organisms have evolved from the same origin, and published his On The Origin of Species in 1859, setting the basis for evolution.[3] How new species arise, what causes variation within a species and why members of a population often are alike was, however, not at the time understood. In 1865 Gregor Mendel carried out and described the first detailed formulas of inheritance,[4-6] and in 1909 Wilhelm Johannsen coined the term *gene* as the unit of heredity, and demonstrated that environmental adaptations are not inherited.[7] Genetics were molecularised in 1944 when Oswald Avery, Colin MacLeod, and Maclyn McCarthy[8] showed that genetic information, i.e. the material inherited from parent to offspring, is *DNA*. The molecular structure of DNA was resolved in 1953 by James Watson, Francis Crick and Rosalind Franklin who revealed DNA to be a 30 nm long ladder-like double helix in which the steps are formed by base pairs. Human DNA contains 3.2 billion base pairs. Their work made possible the understanding of inheritance and evolution at the molecular level. Genes are the recipes for thousands of different proteins encoded in DNA using four different bases (Adenine, Thymine, Cytosine and Guanine).[9] In 1956 Joe Hin Tijo and Albert Levan showed that human DNA is packed in the nucleus of each cell of a human body into 46 chromosomes (23 of them inherited from each parent).[10-12]

The major approaches used in molecular genetics are DNA amplification by polymerase chain reaction (PCR) and sequencing, i.e. reading through the genetic code of DNA.[13-16] In addition, cloning and recombinant technologies provide essential tools for biomedical research today.[17, 18] Research into the life sciences has developed at enormous speed during the past decades, and subsequently, knowledge in medicine and biology has increased exponentially.[19] This can be demonstrated unscientifically by viewing the articles in PubMed database; in the year 1950 the number of publications was 81580, in the year 1970, 216951 and in the year 2008, 803722.[20]

1.1 Genes and polymorphisms

DNA contains coding units, i.e. genes, which are first transcribed into mRNA molecules which in turn are usually translated into proteins, and elements regulating gene expression. Most genes have more than one expression pattern i.e. possess the ability to produce multiple proteins; the same gene may produce slightly different proteins in different tissues.[21, 22] Thus, genes define the function of each cell, allowing them to specialise into over the 300 different cell types found in the human body, for example neurons in the brain or osteoblasts in the bones. Together cells form tissues and organs.[23] Cells send and receive signals which instruct them to express the correct genes when needed.[24] Only a proportion of the genes are functional in each tissue, i.e. those encoding exactly those proteins needed in the cells of that tissue. The remaining genes may be functional at another point in time or expressed in other cell types.[25]

In addition to genes, the genome contains polymorphic (Greek: "having multiple forms") regions which contain the normal variation making us individuals. Polymorphisms are inherited according to the same principles as genes, and if close to (or within) the gene, they provide a useful tool kit for identifying a gene underlying a particular phenotype.[26-29] The polymorphic markers most often used for this purpose to date are microsatellites and single nucleotide polymorphisms (SNPs).[30-32] Microsatellites are short tandem repeat segments of DNA sequence (e.g. CT_n) providing several possible genotypes according to the length of the tandem repeat. Microsatellites have been estimated to comprise 3 % of the human genome.[33] The human genome contains an estimated 10 million SNPs of which 3 million have been identified.[22, 31] The importance of copy number variation (CNV), i.e. gains and losses of DNA segments, has only recently been acknowledged. CNVs have long been associated with disorders caused by chromosomal rearrangements, but recent twin studies have shown that otherwise identical healthy twins may have different copy numbers of DNA segments. Many CNVs include genes and thus, a copy number of a gene can vary between zero and ten. This observation suggests that CNVs may have a significant contribution to evolution, and to normal genetic and phenotypic variation due to the different expression levels of genes.[34] CNV arises during mitosis and studies on mouse embryonic stem cells have indicated that all somatic tissues in individuals can be CNV mosaics.[35] The non-coding, 95 %, of the genome was believed to be unimportant "junk DNA", but currently it seems likely that it contains important elements not yet well characterised.[36, 37]

1.1.1 Projects elucidating the genomic organisation of human and other organisms

In the year 1990, the sequencing of the whole human genome was launched as a massive international collaborative project, the Human Genome Project (HGM).[38, 39] By the year 2004 nearly 100 % of the gene-rich euchromatic human genome had been sequenced, and the physical and genetic *maps* created revealed the locations of 20 000 - 25 000 genes.[22] In addition, the HapMap-project[31] is identifying the locations and variations of SNPs in human sequences. Another project is identifying and locating CNVs of more than 1 kb in the human genome.[40] In addition to the human genome, genomes of other species have been and are being sequenced.[41, 42] These projects have enormously eased the work of the molecular geneticist trying to identify novel genes underlying a disease. Today, the identification of disease genes depends to a great extent on the physical and genetic maps of these genomes available in databases (e.g. NCBI) on the internet.

1.2 Mutations

Mutations, unlike polymorphisms, are changes in DNA which result in failed function of the gene or gene product, interfering with the sensitive system of the tissue or tissues where the mutated gene is expressed. External factors, such as radiation or chemicals, may cause novel changes in the DNA of somatic cells or gametes. Somatic mutations may cause disease, for example cancer, but may also be responsible for harmless variation as well. It is this kind of variation in gametes which gives rise to variation between individuals and makes evolution possible.[32, 36] If a new disease-causing mutation occurs in an egg or in sperm, giving rise to a new individual, this person may have a genetic disorder caused by that mutation.[43, 71]

In general, mutations often cause under- or over-expression of the gene product or alter its structure which leads to a failure of function and a disorder. Recessive mutations have been estimated to be 4:1 more common than dominant ones.[72, 73] This is, however, biased in favour of mild dominantly inherited disorders.[74] There are several ways in which a mutation can lead to altered function: Loss-of-function mutations are often recessive mutations causing a complete loss or reduced activity of the gene product underlying the disorder, i.e. the normal allele in heterozygous carriers produces enough of the gene product for the proper function. Haploinsuffiency refers to a heterozygous loss-of-function dominant mutation which results in half the quantity of gene product and causes phenotypic effects. Gain-of function mutations are dominant mutations which lead to increased levels of gene expression or the development of a new function of the gene product. Dominant-negative

mutations are common in proteins which form dimers or multimers, e.g. actin, tropomyosins and collagens. These altered gene products interfere with the function of the normal gene product.[25, 74]

Point mutations are mutations in DNA where one base of a nucleotide in changed to some other. Within the exon this may result in a missense, nonsense, silent or splicing mutation. A missense mutation substitutes one amino acid for another, often leading to an altered and disrupted conformation of the protein product. Nonsense mutations cause premature termination signals (TAA/TAG/TGA). Missense, nonsense and silent mutations in exons may also disrupt exonic splice signals such as exonic splicing enhancers (ESE) and silencers (ESS), and cause aberrant pre-mRNA splicing.[75-78] Splice-site mutations, however, are usually intronic mutations found in the acceptor or donor splice signals at exon-intron boundaries, or at branch-sites.[75, 79] Deletions and insertions, such as duplications within the exon usually cause a shift in the reading frame of the gene leading to a misread protein tail and a premature termination signal in the mRNA, but some of them are in-frame mutations evoking deletion or insertion of additional amino acids in to the protein without disrupting the reading frame.[25] Deletions or insertions in the intron may cause errors in the splicing patterns of the gene if regulatory elements of splicing are disrupted.[79] Missense mutations are approximately three and nonsense mutations six times more frequent than deletions/duplications/insertions and splice site mutations.[25, 43]

Mutations in genes essential for life may cause lethal or very severe disorders while genes which can either cope with the mutation or have several isoforms to compensate for the function of the faulty gene product, often cause milder disorders. When mutated, "housekeeping" genes expressed in several or even all tissues evoke multi-organ disorders, while genes specific to a particular tissue cause tissue-specific disorders.[43]

1.3 Modes of inheritance

There are four Mendelian modes of inheritance: autosomal (22 human non-sex chromosomes inherited from both parents) dominant (AD) and recessive (AR), X-linked (female sex chromosomal inheritance: females inherit one X-chromosome from each of their parents, males only one from their mothers) and Y-chromosomal (male sex chromosomal inheritance: present in and inherited from males only). Many textbooks separate dominant and recessive X-linked disorders as their own modes of inheritance, but since no true recessive X-linked disorders have been identified with certainty, they can be discussed as one mode of inheritance.[43]

A mutation in one allele is called dominant if it alone causes a disorder and recessive if mutations in both alleles are needed in order for the disorder to appear. In the case of dominant inheritance, one of the parents expresses the trait or disorder and has a 50 % possibility of passing it on to his/her offspring. The parents of a patient with a recessively inherited disorder are usually unaffected, while on average 25 % children are affected, 50 % are unaffected mutation carriers, and 25 % unaffected, not carrying the mutation.[25] Mutations can also be newly arisen, i.e. *de novo* mutations not seen in either of the parents. If such a mutation evokes a condition, it is a new dominant mutation.[43]

The whole concept of inheritance is, however, becoming more complicated and not all traits or disorders are inherited by Mendelian laws. For example in complex, multifactorial and polygenic disorders (such as autism and diabetes)[44, 45] the influence of more than one gene, environmental factors and normal genetic variation is recognised, but not yet well understood.[46] In addition, mutations in mitochondrial DNA,[47, 48] mosaicism (a mutation which is expressed in some, but not all, cells),[49, 50] X-chromosome inactivation and epigenetic gene silencing[51, 52] (epigenetic = gene expression is affected by mechanisms other than changes in the DNA sequence, e.g. gene silencing by promoter methylation) also show non-Mendelian inheritance. Chromosomal changes such as duplications or deletions of whole, or parts of, chromosomes are also the cause of several disorders. The most common chromosomal disorder is Down syndrome, i.e. trisomy of chromosome 21.[53]

The mutations described in this thesis are AR or AD or *de novo* mutations.

1.4 Gene identification

When a gene for a hereditary disorder is to be identified, the family history of the patients must be investigated and pedigrees drawn up. The more information that can be gleaned from several generations of family members, the more informative is the pedigree when elucidating the mode of inheritance. The mode of inheritance usually prescribes the subsequent methods to be used.[28] Sometimes knowing the population to which the patient genetically belongs, is helpful. For example, if the family shows AR inheritance and the family is consanguineous (the parents are relatives) or from an isolated population such as the Finns, this often indicates that the mutation is homozygous, i.e. the affected child has inherited the same mutation from both parents.[54, 55] If the family is not consanguineous, but has many affected members (familial occurrence), samples also from unaffected relatives are useful in further studies to exclude candidate genes and to identify linkage. This is the case even if the mutation inherited from the mother is different from that inherited from the father, i.e. the patient carries two different mutations (compound heterozygous).[28]

1.4.1 Genetic linkage and linkage analyses

Linkage can be defined as "the tendency of genes or other DNA sequences at specific loci (locations in the chromosome) to be inherited together as a consequence of their physical proximity on the same chromosome".[25] In other words, the closer the loci of the sequences are to each other in a chromosome, the more likely they are to be inherited together. This phenomenon can be utilised in disease gene identification based on the assumption that the disease gene is inherited together with a polymorphic marker more often than would be the case for independently inherited elements – in other words, these sequence elements are linked. Furthermore, recombination (crossing over between homologous chromosomes) between the two loci is more unlikely the closer the loci are to each other. Genomic distances can be measured by physical distances using base pairs or by genetic distances, i.e. centiMorgans (cM), which equals a 1 % probability of recombination in the formation of gametes via meiosis. If a distance between two loci is 10 cM, recombination between these sites occurs in 10 % of meioses.[25, 28]

Linkage analysis is used to identify the genomic regions shared between the members of a family/isolated population affected with a genetic disorder utilising polymorphic variation.[28, 30] Haplotypes are created for each family member by arranging their polymorphic markers on a physical map (haplotyping). These maps can be created by analysing fluorescently labelled markers using fragment analysis, but today, especially genome-wide analyses are usually performed using SNP arrays. If the material to be analysed is small, i.e. consists of one or a few families and few markers used, it is sometimes possible to detect positive (haplotypes are shared by the patients and none of the unaffected family members have the same set of haplotypes) or excluding (healthy and affected family members share the same haplotypes or the affected family members have different haplotypes) linkage by viewing the data.[56, 57] Most often, however, mathematical tools are needed to verify the result and to calculate the probability of linkage, especially if the material to be analysed is large, consisting of dozens of families and perhaps covering the whole genome (genome-wide linkage analysis).[58]

The mathematical measurement for the likelihood of linkage is the lod score which is the logarithm of the likelihood (logarithm of odds) for the linkage, assuming the inheritance follows Mendelian laws, and taking into account the recombination fraction (θ). Recombination fraction means the proportion of the meioses in which a given pair of loci are separated by recombination. If the lod score is +3 or more, the region can be considered to be linked (to the disorder), and if -2 or below, linkage can be said to be excluded. The lod score can be calculated using the following formula:[25]

$$LOD = \log_{10} \frac{\text{probability of birth sequence with a given linkage value}}{\text{probability of birth sequence with no linkage}} = \log_{10} \frac{(1-\theta)^{NR} \times \theta^{R}}{0.5^{(NR+R)}}$$

θ = recombination fraction

N = number of non-recombinant offspring

R = number of recombinant offspring

The success and reliability of linkage analysis is dependent on the markers chosen, the number and the variation of the markers, the individuals analysed and the mathematical approach chosen.[26, 27, 58, 59] The number and types of markers included in the analysis are of major importance. Microsatellite markers carry more variation compared with SNPs, but the distance between them in genome-wide analysis is usually 10 cM which is often too sparse for identification of linked regions, and therefore denser, for example 1 cM distance SNP scans are more efficient.[27, 59, 60]

There are different ways to perform meiotic mapping and linkage analyses when attempting to identify a gene. The most common methods are: 1. Parametric (model-based) two-point or multipoint lod score analysis, 2. Non-parametric (model-free) two-point or multipoint lod score analysis, and 3. Association studies. Generally speaking, the two-point linkage analysis evaluates the linkage between the disease locus and the marker, and it is usually the method of choice when analysing candidate gene loci with only a relatively small number of markers and family members.[61] Multipoint linkage analysis examines the linkage between the disease locus and more than one marker simultaneously, which overcomes the errors caused by uninformative markers.[62, 63] This is useful when analysing for example genome-wide SNP data. Association studies are effective when analysing data of complex diseases.[64] Parametric analyses are more reliable but require a knowledge of inheritance and other parameters. Non-parametric analyses rely more on calculated likelihood on the basis of data with less background information. When analysing incomplete data, non-parametric single point analysis or the multipoint approach, is the method of choice.[63]

The term linkage equilibrium refers to the situation where a mutation is specific for one family, i.e. several families may be affected by the same disorder caused by mutations in the same gene but have different mutations and haplotypes segregating with the mutation. Linkage disequilibrium points to situations where all the families with a certain disorder share the same mutation and haplotype segregating with the disorder in question. This is a typical phenomenon in isolated populations, and homozygous areas shared by all the affected patients are sought in order to identify the causative gene. The method is called homozygosity mapping.[28]

1.4.1.1 Computational tools in linkage analysis

Today, researchers have multiple different types of mathematical applications to choose from in order to create haplotypes and calculate lod scores from their "raw" genotype data. The calculation of the lod score is challenging because the calculation of the likelihood for recombination is very complicated.[28] Bioinformatics provide several efficient tools to create pedigrees and haplotypes and there also are programmes to carry out the massive arithmetical operations involving several variables (such as consanguinity, modes of inheritance, penetrance, gene frequencies, distances of the markers used) when analysing the linkage data. Some (parametric) mathematical and computational approaches are purely logical and rule-based, but some (non-parametric) are based on likelihood and/or conditional probabilities.[56, 62, 62] For example MLINK (LINKAGE toolkit) is useful when calculating two-point lod scores.[65, 66] Genehunter calculates multipoint lod scores involving dozens of markers in complex pedigrees. Inheritance information in Genehunter allows the reconstruction of maximum-likelihood haplotypes for all individuals in the pedigree, but due to the more complex mathematical operations of the multipoint lod score, the pedigree size must be moderate.[62, 67] Merlin is used for non-parametric multipoint linkage analysis. It has the ability to detect genotyping errors and omit the uninformative data.[68]

1.4.1.2 Linkage analysis utilizing genetic maps

The candidate region(s) for a disorder provided by genome-wide linkage analysis are usually first roughly delineated and then narrowed further by analysing more markers in the promising areas detected.[25] When a candidate region has been identified, the genetic maps begin to play an important role.[38] The genes in the candidate regions, their expression patterns and possible functions are scrutinised, and those playing a role in the affected tissue and in the particular mechanism believed to be disturbed in the disorder under study, are chosen for mutation analysis. Nowadays this information can be found in public databases.

1.4.2 Candidate gene approach

Understanding the molecular biology of the affected organ or tissue is important when attempting to identify a causative gene for a phenotype. The histology of the affected tissue, the mode of inheritance and the clinical symptoms can also provide valuable hints, indicating the failure of some specific molecular mechanisms to evoke the disorder.[25]

Naturally occurring and gene-modified, knock-in (KI) (mutation pointed to a specific gene) and knock-out (KO) (gene function silenced) animal models can also provide valuable information and clues to the possible causative genes when expressing phenotypes and/or histology similar to the disorder being studied.[69, 70] Today, expression arrays play an important role in elucidating the expression patterns of different genes in different tissues. Sometimes it is possible to identify a gene based on these data, alone, and no linkage analysis needs to be performed. In addition, not all families are suitable for linkage analysis. This is the case if a DNA sample is available only from the affected child of a non-consanguineous family, there is no knowledge of the mode of inheritance, or the mutation is likely to be *de novo*.[28] In these cases, the candidate gene is analysed directly for the mutations.

1.5 Mutation identification

Identification of the disease-causing mutations is often important for the diagnosis, and the family concerned. From the biological point of view, characterising the gene and the mutation causing the condition provides valuable information on the function of the gene. The most common and straight-forward method for mutation identification in both genomic and mitochondrial DNA is sequencing.[14, 15] If the gene does not express several isoforms in the tissue studied and if the geneticist has access to the tissue affected, the mRNA of the gene can be extracted from the tissue, converted to complementary DNA (cDNA) using RT(reverse transcriptase)-PCR and sequenced, reducing the length of the DNA fragments that need to be analysed.

When analysing large genes with tens or hundreds of exons, or a large number of patient samples, a pre-screening method for genomic changes in the DNA is preferred. In addition, several methods have been developed in order to detect different types of mutations. These kinds of mutation screening methods are for example SSCP (Single-Stranded Conformation Polymorphism), dHPLC (denaturing High Performance Liquid Chromatography) and MLPA (Multiplex Ligation-dependent Probe Amplification).

SSCP is based on different secondary conformations of denatured DNA single-strand fragments containing a SNP or a mutation when the fragments move in an electric field towards the positively charged end in a non-denaturing polyacrylamide gel.[80] The method was previously widely used for mutation and SNP detection, but since it is such a time-consuming method, SSCP has largely been replaced by more modern methods such as dHPLC.[81]

dHPLC detects small heterozygous variations in heteroduplex DNA fragments. The principle of the method is presented in *Figure 1*: Heteroduplex fragments are produced by denaturing PCR-fragments using heat followed by slow cooling. This leads to the formation of heteroduplex and homoduplex DNA fragments when a heterozygous mutation (or SNP) is present in the sample DNA. Due to the different chemical properties and charges of the homo- and heteroduplexes, they adhere differentially to the hydrophobic electrostatically neutral matrix of the stationary phase (polystyrene-divinylbenzene copolymer beads in the column of the chromatography), and elute with the hydrophobic buffer (acetonitrile-TEAA-buffer - a running liquid phase) from the column at different time points. This is detected by an inbuilt spectrophotometer, and seen as peaks in the chromatograms.[82, 83]
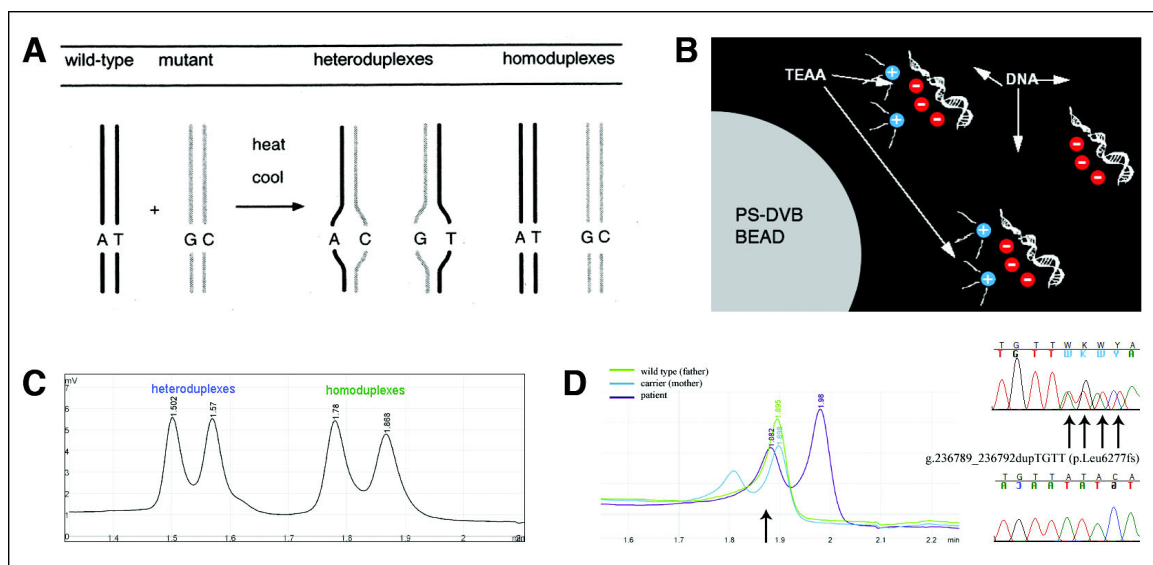

Figure 1. The principle of dHPLC. **A.** Formation of the heteroduplexes by heating and slowly cooling the PCR-product. **B.** Separation of the hetero- and homoduplexes in the column of the dHPLC. (Transgenomics Inc) **C.** Chromatogram peaks of optimised standard samples. **D.** An example of the mutation seen in a patient and his mother. The curve produced by the fathers sample resembles the wild type and he does not carry the mutation. The frameshift mutation identified shown in the nucleotide sequence.

dHPLC is able to scan hundreds of samples a day and when the analysis temperatures for each fragment are optimised carefully, it is a reliable method for screening mutations.[84] Several variations of this method have been developed for e.g. detection of mitochondrial mutations or mutational mosaicism.[85]

Large copy-number changes consisting of larger segments of genomic DNA covering one or multiple exons cannot be detected by the above methods. When such mutations are suspected, the method of choice could be MLPA or Southern blotting. Southern blotting is

based on the hybridisation of labelled cDNA probes to restriction enzyme-digested and subsequently size-separated target DNA transferred to a filter membrane.[86, 87] In MLPA, a copy is made of each target sequence by hybridisation of two probes to the target DNA and ligating them. The target sequences are amplified in a multiplex-PCR reaction using universal fluorescent-labelled primers. The PCR fragments are analysed by fragment analysis methods, and the relative copy numbers of the target are calculated.[88] Like dHPLC, MLPA also has several applications, e.g. for detection of variations in the methylation patterns of a gene.[89]

High-throughput sequencing and microarrays provide more advanced mutation identification tools utilising the latest and rapidly developing technologies for mutation detection and expression studies. With high-throughput sequencing, the whole genome can be analysed relatively quickly. This can, for example, be performed in a system where multiple DNA segments are bound to microscopic beads and amplified in one reaction (pyrosequencing or sequencing-by-synthesis), and the sequence analysis is automated by the software comparing the reference sequence to the sequence analysed.[90] Microarrays are chip-based systems where, for example, wild-type DNA probes (such as the exons of a gene) are bound to the membrane of the chip and the DNA studied is hybridised to it. The level of the affinity of the hybridisation is analysed using bioinformatic tools. The expression of a gene studied can be analysed by hybridising mRNA to the probes bound to the membrane.[91, 92]

1.5.1 Verification of the mutation

When a possible mutation has been identified, it is first verified by analysing control DNA samples from healthy individuals. Evaluation of the conservation of the change, especially of a point mutation, is performed by comparing the amino acid sequence to animal orthologs, and in the case of missense mutations, to protein homologues using computational database-based BLAST tools.[93] It is assumed that the more conserved the amino acid is between different species is, the more relevant it is for the function of the gene.[25]

RT-PCR reveals the outcome of the mutation at the mRNA level. Faulty mRNA can lead to an altered protein product, but the RNA may also be degraded before being translated into protein. The expression levels of the gene can be studied using RNA-based methods, for example by northern blotting and TaqMan (real-time PCR-method developed to quantify differences in mRNA expression).[94, 95] If possible, the effects of the mutations identified in genes should be studied in the protein. Alternatively, the change caused by the mutation can be predicted by computer programmes designed for detecting possible misfolding or loss of important domains in the protein. Western blotting is the method for primary analysis of protein size differences by running the protein samples in a gel where they drift differentially

according to their size. Two-dimensional gels (with pH gradient in one and size separation in another direction) can be used to investigate changes in the polarity and conformation of the protein. Several approaches such as two-hybrid methods have been developed to study the binding of proteins to their targets (enzyme-substrate or multimerisation studies). In addition, it is possible to analyse the amino acid sequences of the protein by mass spectrometry-based tools.[43]

2 Muscle tissue

According to the body and trace fossils found in the White Sea, it has been estimated that the evolution of the muscle tissue began 555.3 ± 0.3 million years ago to meet the requirements of locomotion.[96] Studies by computational tools on reconstructed phylogenetic trees and analysis of the six genes expressed in muscles suggest that the cardiac and skeletal muscle tissues share a common (genetic) ancestor which existed already before the divergence of the arthropods (e.g. spiders, insects) and the vertebrates, while smooth muscle seems to have evolved independently, and is believed to be the most primitive type of muscle tissue. Today, invertebrates have smooth and skeletal muscle tissues, while vertebrates have, in addition, a specialised cardiac muscle tissue found in heart[97] (personal communication with Prof. Frank Corsetti). In humans, there are approximately 650 muscles producing force for the movement of the human body and its internal organs with their own histological and functional characteristics.[43]

Skeletal/striated muscles (*Fig 2*), comprising 40 % of the human body mass, are usually attached to bones by tendons. Skeletal muscle is built up of cylindrical units inside one another: the muscle is formed of bundles of fascicles surrounded by thin layers of connective tissue. Fascicles are formed of bunches of muscle fibres ($\varnothing$ = 0.01 – 0.1 mm), which in turn are packed with myofibrils ($\varnothing$ = 1 - 2 μm). Myofibrils again are formed of the smallest units of the muscle bundled together, i.e. highly organised contractile units, the sarcomeres. The sarcomeres are responsible for the striated appearance of the skeletal muscle tissue. Actin and myosin filaments are the main components of the sarcomere. Muscle fibres (*Fig 3A*) have multiple subsarcolemmal (cell membrane) nuclei sometimes difficult to distinguish from satellite cells which are small mononuclear progenitor cells found between the sarcolemma and basal membrane. These cells are able to differentiate into new fibres. The contraction of the skeletal muscle is dependent on conscious nervous stimuli.[23, 98, 99]
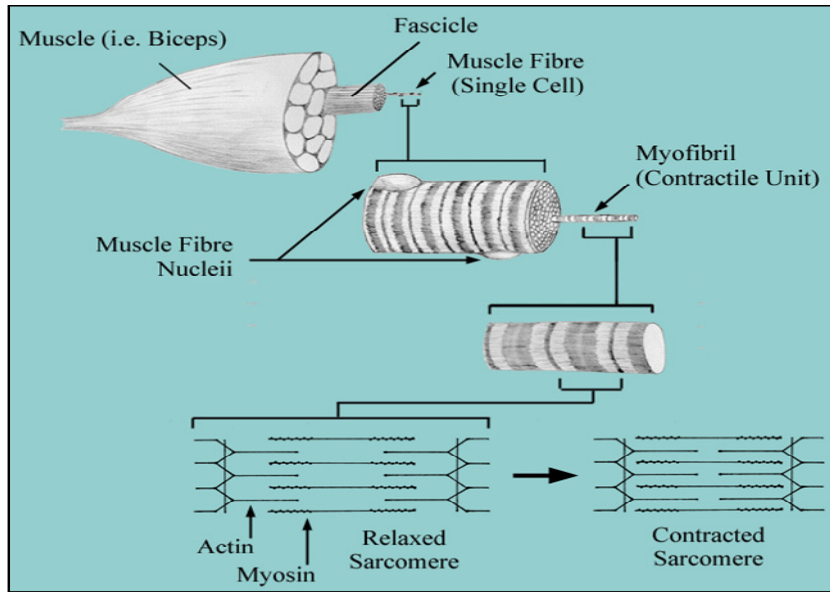
Figure 2. The organisation of the skeletal muscle. (Figure kindly provided by Professor Sandra K. Ackerley, University of Guelph, Canada)

Smooth muscle, found for example in the walls of the blood vessels and internal organs contracts involuntarily and very slowly, often in response to signals from the autonomic nervous system. It comprises 3 % of the body mass. The smooth muscle cell has a single central nucleus and the cells form neat parallel but less organised structure than skeletal muscle cells (*Fig 3B*). **Cardiac muscle** is unique to the heart (*Fig 3C*). It shares features with both skeletal and smooth muscle; it is striated and multinucleate, but the nuclei are centrally located. One feature unique to cardiac muscle is the presence of branched cells which are joined to one another via intercalated discs, making sequential contraction of the heart possible.[23]
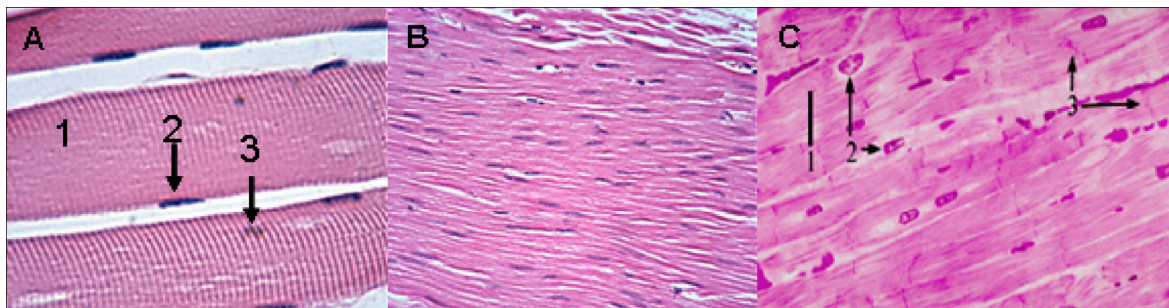

Figure 3. **A.** Skeletal muscle **1.** a single fibre, **2.** a satellite cell, **3.** a nucleus. **B.** Smooth muscle tissue **C.** Cardiac muscle **1.** a single cell, **2.** nuclei, **3.** intercalated discs. (Adapted from http://www.kumc.edu/instruction/medicine with permission)

2.1 Skeletal muscle fibre types

Skeletal muscle fibres have specialised as oxidative (red) slow twitch, type I (or 1) fibres and glycolytic fast-twitch, type II (or 2) fibres (*Fig 4A*). **Type I** fibres are dominant in large muscles rich in mitochondria and myoglobin, and responsible for most of the aerobic long term/static activity. Type II fibres are divided into subtypes: **Type IIA** fibres can create energy by both aerobic and anaerobic metabolism, **type IIB** (white) fibres use anaerobic metabolism to create energy for producing powerful bursts. They have high contraction frequencies but they fatigue quickly. This fibre type is dominant in rodents. **Type IIX** is the fastest muscle fibre type, being able to contract most quickly and to use anaerobic metabolism to generate short-term power. The tiring of type IIX fibres during bursts of activity causes the pain nicknamed "lactic acid pain". The fibre type is determined by the type of neuron which innervates the muscle, and fibres of different types have different gene expression patterns. All muscle fibres of a motor unit are of the same type (*Fig 4B*).[98, 100]
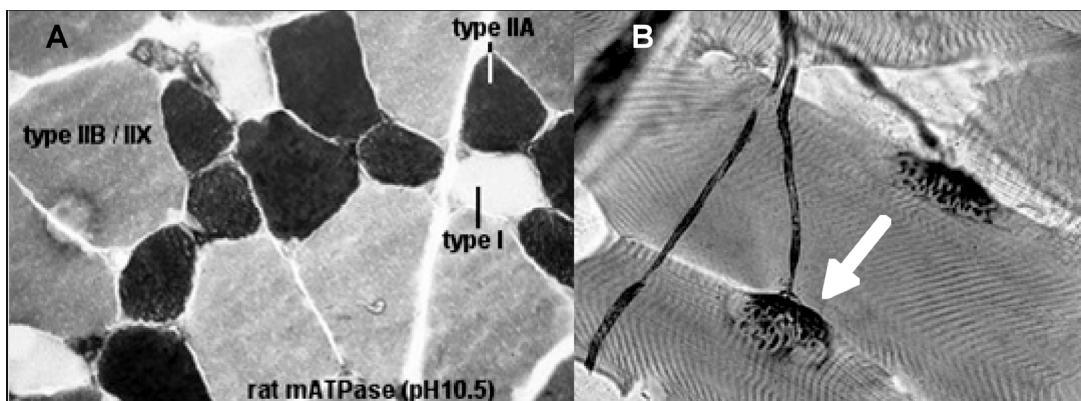
Figure 4. **A.** Type I, IIA and IIB/X fibre seen in a cross-section of rat skeletal muscle tissue. **B.** Innervation (pointed by an arrow) seen in longitudinal section of skeletal muscle fibres. (Figures kindly provided by Prof. Roger Wagner, University of Delaware, USA)

2.2 The skeletal muscle sarcomere

Microarray studies on skeletal muscle have estimated that some 3500 genes are expressed in this type of tissue. At least 1000 genes of these are believed to be specific for muscle tissue and essential for muscle cells, and most of them encode proteins for the sarcomeres. The largest proteins encoded by the largest genes (such as titin, nebulin and obscurin) in vertebrates are found in the muscles.[101-105]

Figure 5. **A.** The structure of the muscle sarcomere: Actin molecules in green, myosin molecules in red, tropomyosin and nebulin molecules as black lines spanning the thin (actin) filaments. Thin filaments in muscle sarcomeres are anchored at the Z-disc by the cross-linking protein α -actinin (orange) and are capped by CapZ seen in pink in the Z discs. The thin-filament pointed ends terminate withi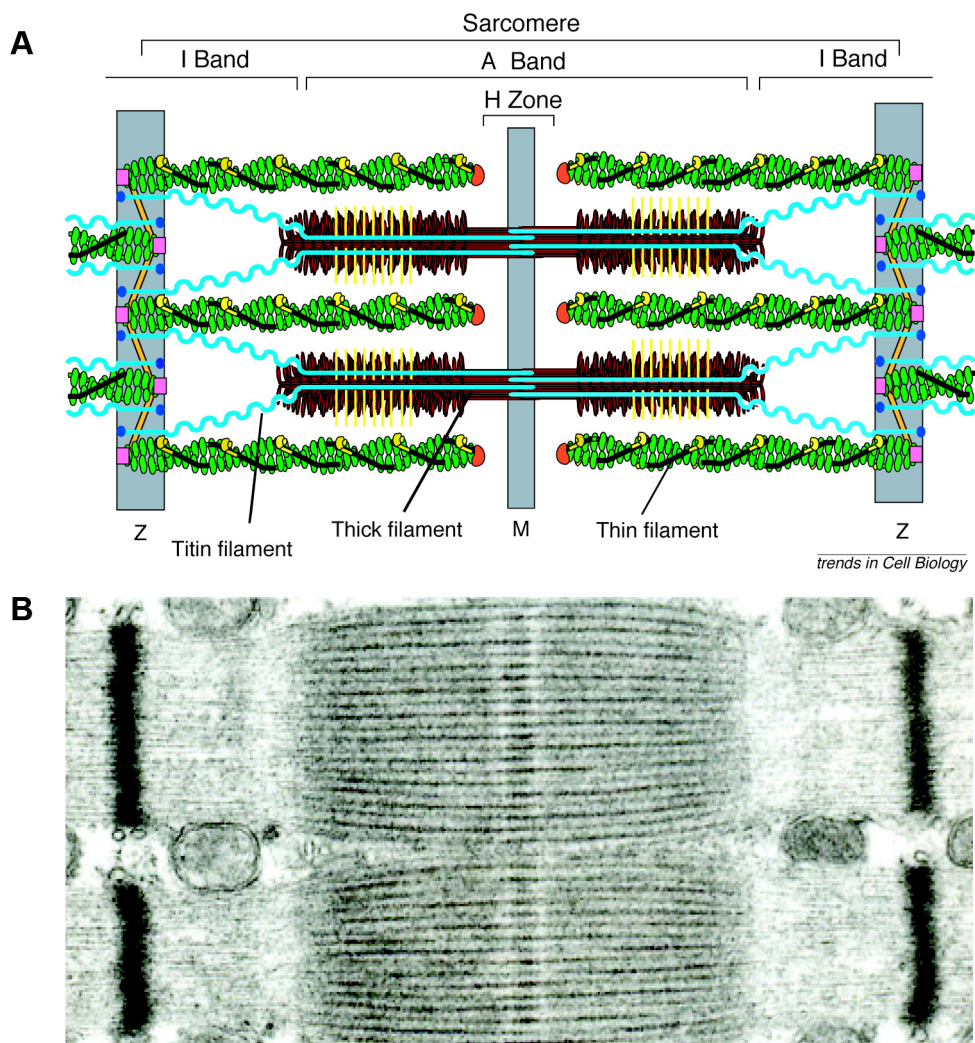n the A band and are capped by tropomodulin (bright red). Myosin-binding protein C (MyBP-C) as vertical yellow lines. Titin (the third filament system) is in turquoise. (Reprinted from Gregorio et al.,2000, Trends in Cell Biology, with permission from Elsevier) **B.** Electron microscopic view of a sarcomere (Reprinted from Ottenheijm et al., 2008, Respiratory Research with permission of BioMed Central)

The basic structure of the sarcomere has been known for decades,[106, 107] but many structural and functional details remain to be resolved even today. The sarcomere consists of a meshwork of hundreds of structural proteins, and proteins functioning in signalling cascades between the sarcolemma and the sarcomere and in contraction (*Fig 5A and 6*).[108] The sarcomere can be divided into sections according to the bands seen under the electron microscope (EM) (*Fig 5B*). The thin filament is attached to the Z disc and extends to the I band seen in the EM view as a pale section next to the Z disc ending at the A band of the sarcomere. The A band is seen as a denser area of the sarcomere where the thin, thick and

third filaments overlap and interact. Here the molecular events of the muscle contraction take place i.e. actin filaments slide along the myosin filaments toward the H band and mid-line, M line. During the contraction, the I band and the H band are shortened while the length of the A band remains the same.[107, 109]
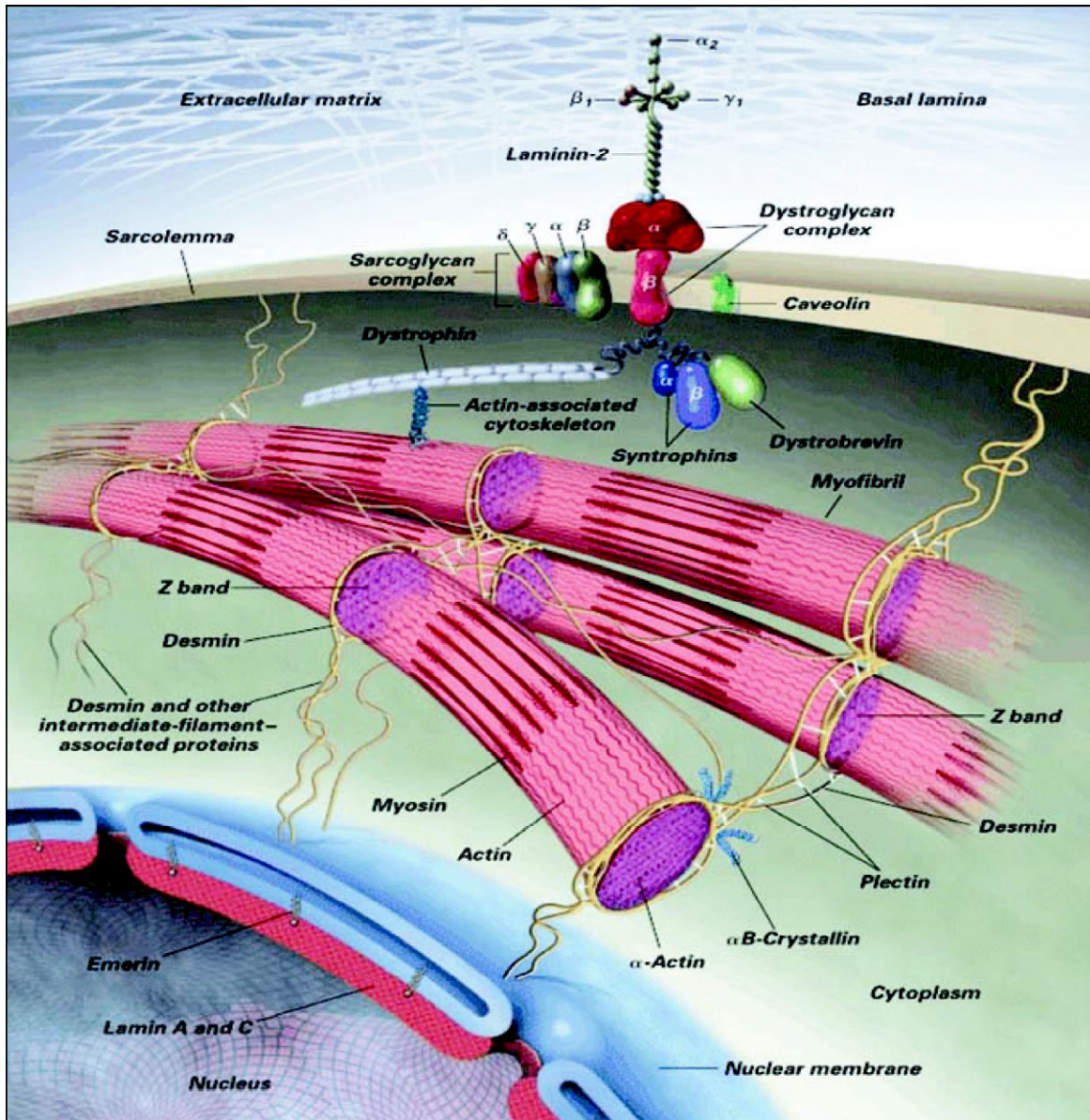

Figure 6. An overview to the skeletal muscle cell illustrating the filamentous systems, interactions and connections between sarcolemma, sarcomere and nucleus (reproduced with permission from Dalakas et al. The New England Journal of Medicine 2000. Copyright © Massachusetts Medical Society)

2.2.1 The Z disc

Z discs, seen in EM (*Fig 5B*) as the darkest lines, define the border of the two adjacent sarcomeres, link sarcomeres to the sarcolemma, allow the force to be transmitted along a myofibril during contraction, and function in stretch sensing and signalling.[98]

Z discs are composed of zigzag layers formed by the connections of the oppositely oriented thin (actin) filaments of the adjacent sarcomeres (*Fig 5*). The thickness, defined by the number of layers of the Z disc indicates different fibre types: in fast fibres it is ~30–50 nm and in slow and cardiac fibres ~100–140 nm thick. It has been proposed that the number and/or composition of the N-terminal domains of titin (the backbone of the third filament system) and the C-terminal domains of nebulin (the thin filament ruler) within a single sarcomere could differ and influence the thickness of the Z discs.[110-112]

α-**actinin 2** belongs to the α-actinin protein family. In muscle, it exists as a calcium-insensitive protein forming the backbone of the Z disc, having the ability to bind many structural and signalling molecules.[113-117] It is essential not only for mature muscle as a cross-linker of multiple proteins, but together with titin it is needed for the proper assembly of the developing sarcomere, i.e. sarcomerogenesis.[118] In mature muscle, actin and titin filaments are anchored together and cross-linked in the Z disc via α-actinin 2.[116] Other important Z-disc components include titin-binding and -capping **telethonin**,[112, 119, 120] the **capZ-complex**, which binds to nebulin and caps the actin filament at the Z disc,[112] as well as **myotilin** and **myopalladin**.[114] In addition nebulin and titin are bound to each other in the Z disc.[121]

2.2.2 The thin filament

The backbone of the thin filament is an actin polymer (*Fig 5*). The lengths of the thin filament display some variability according to the fibre type and muscle tissue, though usually it is longer than 1 µm. Proteins bound to actin, determine the length of the thin filament and make muscle contraction possible.[122, 123] The proteins most relevant for this thesis are discussed further.

2.2.2.1 Actin

Six different genes encode six actin proteins: α-skeletal, α-cardiac, α-smooth muscle and γ-enteric actin are tissue specific. In addition, β- and γ-actins are expressed in almost all vertebrate cells where they act in maintaining the cell structure by forming an actin cytoskeleton.[124] Actins are very conserved proteins, and do not seem to tolerate changes in the nucleotide and/or amino acid sequence. The gene encoding skeletal thin filament actin, *ACTA1,* is in chromosome 1. It has six exons and the length of the mRNA is 1374 bp.

Two polymerised skeletal muscle α-actin filaments coiled around each other form the backbone of the thin filament. Each globular actin molecule can bind to four other actins; two to the same and two to the second actin polymer of the same filament.[125] In addition, actin has three binding sites for nebulin, and it binds troponin/tropomyosin complexes as well as several other proteins, which anchor the filament to the Z disc, and cap it in the H band (tropomodulin).[126, 127] An α-actin molecule contains Mg^{2+}- or Ca^{2+}-ions and ATP/ADP-binding sites, making energy-dependent muscle contraction possible due to the interaction with the myosin heavy chains of the thick filament.[125] The structure of the actin polymer and filament is illustrated in *Figures 5* (*p. 26*), *9* (*p. 33*) and *10* (*p.35*).

2.2.2.2 Nebulin

Nebulin is required for the proper assembly of the thin filaments and Z discs in mature muscle tissue as well as for defining and maintaining the correct lengths and contractile function in different fibre types.[112, 128-131] Studies performed on nebulin fragments already over a decade ago, revealed that its domains periodically bind actin, calmodulin, tropomyosin and troponin complexes suggesting that the segmental structure of nebulin plays an important role in defining or maintaining the length of, as well as stabilizing the thin filament of the sarcomere.

It has been estimated that there are two nebulin molecules spanning one thin filament.[132] The ~20 kDa C-terminus of nebulin is anchored into the Z disc.[133] It contains a conserved SH3 domain, which binds CapZ, titin and myopalladin. CapZ caps the Z disc end of the thin filament, titin forms the third filament, and myopalladin cross-links nebulin via α-actinin to the Z disc (among its other functions).[112, 121, 134-138] The peripheral C-terminus binds desmin.[139] Desmin is an intermediate-filament protein forming a bridge between the sarcolemma and the Z disc.[140] The function of multiple phosphorylation sites seen in the C-terminal nebulin is still not fully clarified.[141] With respect to nebulin, 97 % (mainly in the I and A band regions) consists of 30 – 35 amino acids long α-helical simple repeats.[122, 129, 132] Each of

the 5.5 nm long simple repeats contains an actin-binding site (SDXXYK-motif). Depending on the isoform, nebulin may contain 179 – 239 simple repeats, each of which is capable of binding 179 – 239 actin monomers of the thin filament. Most of the simple repeats are arranged into super repeats, each seven simple repeats long with the potential to form 22 – 30 super repeats (*Fig 7*). There is a WLKGIGW-motif present in every super repeat at 38.5 nm intervals most probably for binding tropomyosin and troponin.[122, 142] The binding of nebulin to the tropomyosin and troponin complex was deduced to form a calcium-linked regulatory complex.[122, 129, 132, 133, 143] The 8 kDa N-terminus at the H band region contains unique domains for binding tropomodulin, which caps the pointed end of the thin filament.[144-146]
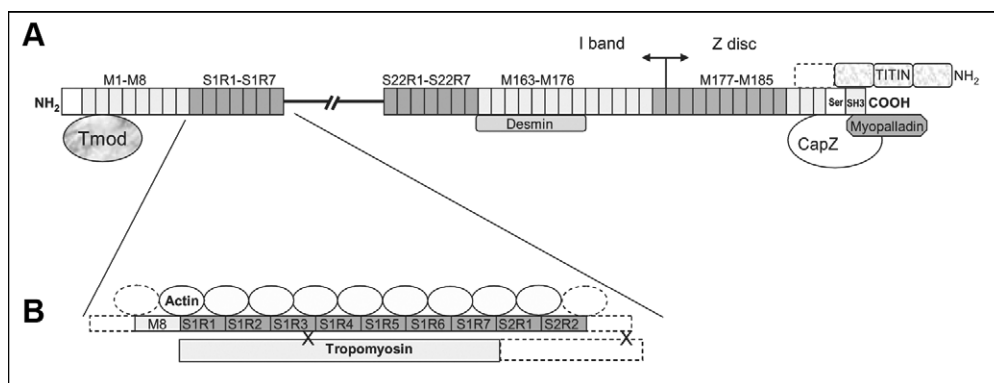

Figure 7. **A.** The protein structure of nebulin and the binding-partners (tropomodulin (Tmod), desmin, CapZ, myopalladin and titin shown) of nebulin. M1-M8 and M163-M176 are simple repeats not organized into super repeats; S1R1-S1R7, S22R1-S22R7 are super repeats of seven simple repeats, Ser = serine rich and SH3 = Src homology domain. **B.** A detailed view of one nebulin super repeat (S1) consisting of seven simple repeats (R1-R7). The actin-binding motifs (SDXXYK) are present in the simple repeat boundaries and tropomyosin-binding motifs are present in the third simple repeat (R3) of every super repeat. (Adapted from The Sarcomere and Skeletal Muscle Disease, ed. Nigel Laing, with permission from Landes Bioscience and Springer Science)

The calmodulin regulated interaction of nebulin with actomyosin suggests a role for nebulin in the regulation of muscle contraction via a calcium-linked system.[143] It was shown that the N-terminus of nebulin inhibits the sliding of acting over myosin in *in vitro* mobility assays, while the C terminus located in the Z disc did not inhibit the sliding. The nebulin KO mice have provided new knowledge on the role of nebulin *in vivo.*[112, 147] The results of two separate groups working with nebulin KO mice show that the animals which do not express nebulin do assemble sarcomeres prenatally, but the thin filaments are disorganized and 15 – 25 % shorter than normal, and the Z discs are abnormally thick. The mice died at approximately two weeks of age resembling both clinically and histologically human patients with severe NM caused by mutations in *NEB.*[112, 147, 148] It was hypothesised that in the muscle lacking nebulin, altered Ca^{2+} homeostasis would lead to dysfunction of the muscle. It was

noted that the levels of the sarcoplasmic reticulum (SR) Ca^{2+}-ATPase (SERCA) –inhibitor, sarcolipin (SLN), were upregulated in nebulin KO mice.[112] As explained in Chapter 2.3, muscle contraction is triggered by the release of Ca^{2+} from the SR and relaxation occurs when Ca^{2+} is taken back up into to the SR by SERCA. Further investigations supported the hypothesis that nebulin has a role in handling Ca^{2+} in muscles and in regulating muscle contraction. The studies showed that if nebulin is not present in the muscle, SLN was up-regulated, while the expression of other proteins involved in Ca^{2+} pathways were not significantly altered. In nebulin deficient muscle, the speed of Ca^{2+} uptake decreased and the relaxation time was significantly longer.[131]

The gene encoding nebulin, *NEB*, (*Fig 14, p. 55*) is located to the chromosomal region 2q22 and it is one of the biggest genes known,[149] containing 183 exons in an area of 249 kb of genomic sequence. Translation begins at exon three and ends at exon 183. An 8.2 kb genomic region in the middle of the gene encompasses a triplication of a segment containing eight nearly identical exons (exons 82 – 89, 90 – 97, 98 – 105) and introns. These, as well as exons 63 – 66, 143 – 144 and 167 – 177, point to alternatively spliced exons theoretically giving rise to thousands of different nebulin isoforms. The splicing patterns of the triplicated area are not yet known in detail, but it has been predicted that the region may produce seven different length variants.[150] Exons 63 – 66 form one cluster of exons which are either all included or excluded from the transcript (exon 62 is spliced to exon 67 or all exons 63 - 66 included). Exons 143 and 144 are mutually exclusive, i.e. the transcripts always contain either of them, never both, while exons 167 – 177 are independently spliced, i.e. they are included or excluded independently of each other.[151] The vast nebulin isoform diversity probably meets the different requirements of prenatal vs. adult muscle, different muscles and muscle fibre types.[142, 152, 153] Due to its extensive splicing, the size of the nebulin protein varies between 600 and 900 kDa.[142, 151] Nebulin is mainly expressed in the striated muscle thin filament,[122, 154] but minor expression has been detected in the heart[152] and possibly in the brain[155].

2.2.2.3 The tropomyosins

Tropomyosins are α-helical coiled-coil homo- or heterodimers which form a long filament by polymerising head-to-tail (*Fig 8*). The dimerised polymers run along the length of the actin molecule aside nebulin (*Fig 9, p. 33 and 10, p. 35*).[157-160] Tropomyosins bind to actin, stabilising the thin filament, and together these molecules regulate muscle contraction.[125, 127, 161] A more detailed investigation of the tropomyosins reveals heptapeptide repeats (*abcdefg*) underlying the coiled coil structure. The *a* and *d* residues are generally non-polar and form the interhelical space or core of the double-stranded structure.[127]
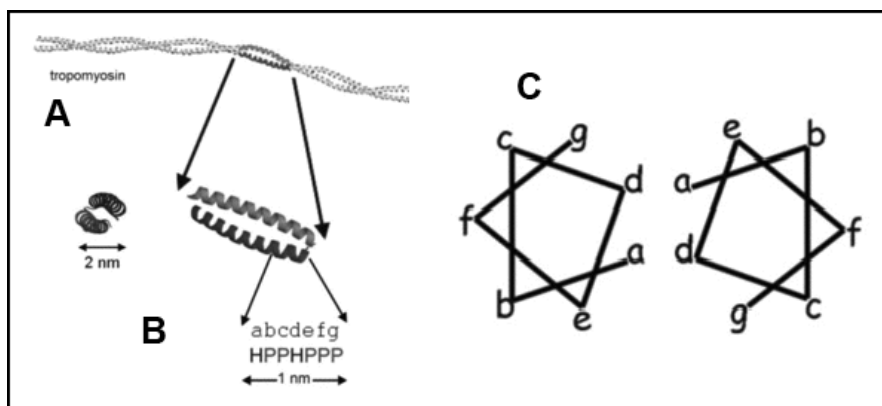
Figure 8. A. The coiled-coil structure of the dimerised tropomyosins. **B.** The heptad repeats (abcdefg) of coiled coils; HPPHPPP where H are hydrophobic amino acids in the core of the helix, and P are polar exposed outwards from the molecule. One heptad is 2 nm in diameter and 1 nm long. **C.** Cross-section view of the coiled coils. (Reproduced with permission from Ryadnov MG, Biochem. Soc. Trans. 35, 487-491, 2007. Copyright © the Biochemical Society)

Four different genes, all containing 10 exons encode the tropomyosins: *TPM1*, *TPM2*, *TPM3*, and *TPM4*. These genes can encode over 40 different tropomyosin isoforms due to alternative promoters and splicing.[162-165] The genes *TPM1, TPM2* and *TPM3* encode the skeletal, exactly 284 amino acids long muscle-specific isoforms α-tropomyosin$_{fast}$, β-tropomyosin, and α-tropomyosin$_{slow}$. α-tropomyosin$_{fast}$ (encoded by *TPM1*) is expressed in type II muscle fibres and is the most abundant isoform in the heart. β-tropomyosin (encoded by *TPM2*) is present in both muscle fibre types; more abundantly in type I and less in type II muscle fibres (and in small amounts in heart). α-tropomyosin$_{slow}$ (encoded by *TPM3*) is found in type I fibres only. When both α- and β-tropomyosins are expressed, $\alpha\beta$-heterodimers are formed preferentially over $\alpha\alpha$-homodimers, and $\beta\beta$-homodimers are rare.[127]

2.2.2.4 The troponin complex

Troponin C (calcium binding), I (inhibitor) and T (tropomyosin binding) form a complex which regulates muscle contraction (*Fig 8 and 9*).[125] There are several genes encoding troponins. Troponins I and T have specific isoforms for type I, type II and cardiac fibers (*TNNI1*, *2*, and *3* & *TNNT1*, *2* and *3*), but troponin C has one gene encoding type I and cardiac fibre isoforms (*TNNC1*), and another gene encoding type II fibre isoform (*TNNC2*). These different isoforms differ from each other by only a few amino acids.[162]

2.2.2.5 The cofilins

Cofilin 1, cofilin 2 (the skeletal muscle-specific isoform encoded by *CFL2*) and destrin belong to a protein family which regulate actin filament dynamics. In the thin filament, cofilin 2 acts together with the actin depolymerisation factor catalysing the depolymerisation of the actin filament.[166]

2.2.3 The thick filament

Myosin and myosin-binding proteins form the thick filament of the sarcomere (*Fig 9*). Myosin is the main component, but hundreds of other components contribute to the actin-myosin interaction during contraction, acting as accessory proteins and stabilizing the structure of the thick filament.[167]
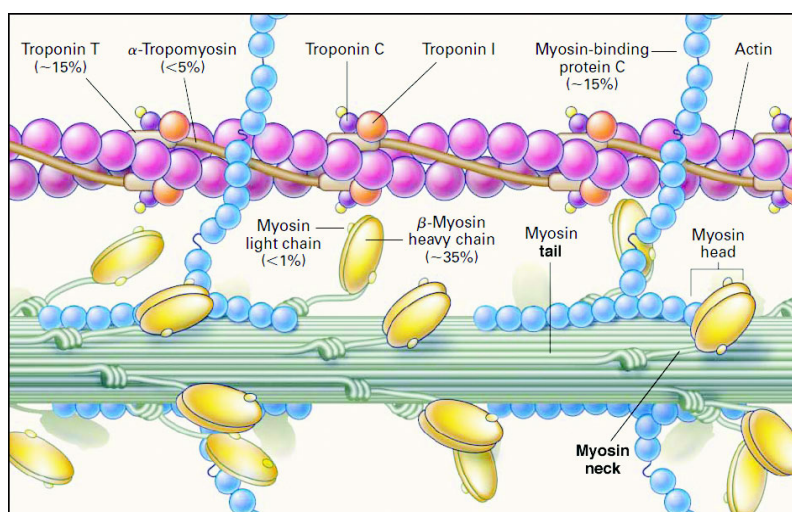
Figure 9. Thin and thick filament structures and interaction. Head, neck and tail domains of the myosin molecule pointed out. (Reproduced and modified with permission from Spirito et. al. The New England Journal of Medicine 1997. Copyright © Massachusetts Medical Society)

Myosins form a protein family of molecules able to bind actins and hydrolyse MgATP for energy, for example for cell crawling. Myosins consist of two identical heavy (MyHC) chains and two different pairs of light chains (MYL), and there are three domains in the myosin molecule: Head, neck and tail. MyHCs are divided into classes I and II, with the seven isoforms expressed in vertebrate skeletal muscle belonging to class II. Myosins are encoded by a total of ~40 genes belonging to ~12 classes according to their head and tail domain structures.[168] The various MyHC isoforms in different muscle fibre types bind actin with the head domain during muscle contraction and force production, the neck binds the light chains

33

which act as a regulatory platform, and the tail maintains the structure of the molecule and provides the anchoring point. In the sarcomere, myosin molecules exist as a chain of 147 myosin molecules. The myosin chains of the opposite thick filaments are bound together by M-band proteins such as myomesin, and 294 myosin molecules form an exactly 1.59 μm long and 1 nm thick filament. Another protein worth mentioning is myosin-binding protein C, which enables the muscle to regulate its contraction (*Fig 9*).[168-170]

2.2.4 The third filament - titin

While nebulin defines the length of the thin filament, titin (*Fig 5A, p.26*) is the ruler of the whole sarcomere. Titin, 3600 kDa in size, forms the backbone of the third filament of the sarcomere. It provides the elastic stabilising spring which spans the distance from the Z disc to the M line, i.e. a titin filament spans half of the sarcomere, and binds and interacts with several proteins along its length. It consists of several domains with different functions needed in different parts of the sarcomere.[171-174] About 100 kDa of the N-terminus of titin is bound to the Z disc, being cross-linked to it via several proteins, such as calpain 3 and α-actinin.[119, 120, 175] A 1500 kDa section of the I-band titin includes the calpain 3-binding N2A-region and functions as an elastic spring which is able to shorten during contraction in relation to the shortening of the whole I-band.[174, 176] In the A-band, 1750 kDa of titin is cross-linked to the myosin filament which in turn is linked to actin by myosin-binding protein C.[177] During contraction, the length of the A-band remains stable.[108]

The more complex 250 kDa M-line part of titin serves as a signalling platform for several pathways which function in the coordination of the sarcomere. This region of titin interacts with myomesin and contains sequences of several signalling domains such as the kinase domain. The kinase domain, in turn, is a binding region for several signalling complexes, which e.g. provide a mechanical sensor for muscle contraction activity. The muscle sarcomere can react to external stimuli according to the binding state of the signalling molecules present in these titin M-line bound complexes. For example, if a signalling molecule is bound to titin, it is inactive – if unbound it shuttles to the nucleus and regulates the transcription of other muscle genes via SRF (the serum response factor).[178] In the M-line, the two titin molecules of the opposite halves of the sarcomeres are bound together.[179] Among other functional roles of titin, it has a role in stabilizing and activating calpain 3. Calpains are calcium-dependent intracellular proteases of which calpain 3 is the most abundant isoform in skeletal muscle. In the sarcomere, titin-bound calpain 3 is believed to act as a sarcomeric sensor which is involved in its maintenance and repair.[180]

The protein is encoded by the gene *TTN* which contains 336 exons and the whole gene covers 294 kb of genomic DNA in chromosomal region 2q31. The gene is extensively spliced to meet the requirements of different muscle fibre types and in different phases of the maturation of the muscle cell. It is also one of the first genes expressed during sarcomerogenesis and seems to be essential for this process.[177, 181-184]

2.3 Muscle contraction

Due to the complex cascades evoked by the action potential of the motor neuron, the sarcolemma of the muscle fibre is depolarized, which in turn triggers a reaction elevating the calcium level within the muscle cell. Calcium binds to troponin C, which causes conformational changes of the complex, and the movement of tropomyosin to expose myosin-binding sites on the actin filament (*Fig 10*). Myosin can now tightly bind to actin (*Fig 9, p. 33* and *10*). the subsequent reactions release ADP and inorganic phosphate from myosin to generate the force needed for the myosin head domains to drag the thick filament along the thin filament i.e. pulling the Z-discs closer to each other. Next, ATP is bound to myosin which detaches from the actin, and then ATP is hydrolysed to ADP and the myosin is ready to interact with actin again. This continues until no calcium or ATP is left, and the troponin-tropomyosin complex changes its conformation back to the resting state. In resting muscle, troponin T is bound to tropomyosin, blocking the myosin-binding sites of the thin filament.[125, 185]

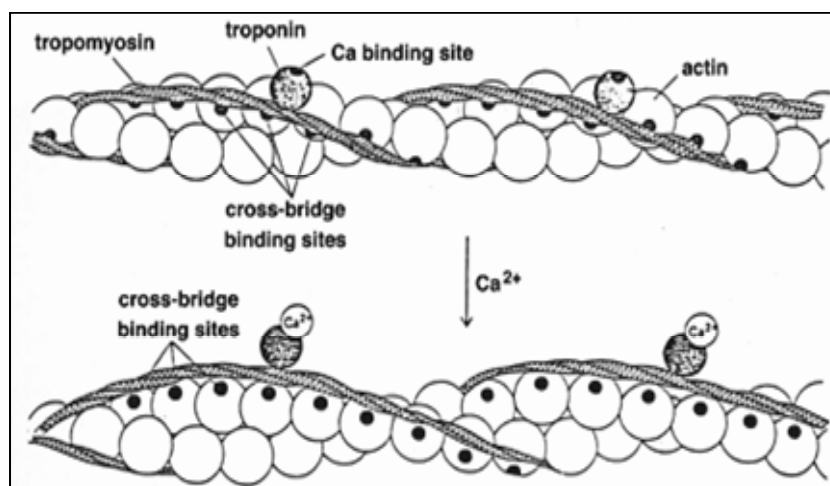

Figure 10.
The principle of the interaction of the thin filament and Ca^{2+} during muscle contraction. Nebulin runs along the tropomyosin dimer. (Figure kindly provided by Professor David Warshaw, University of Vermont, USA)

3 Nemaline myopathy and related disorders of the muscle sarcomere

Is has been estimated that there are at least 1000 muscle-specific genes in the human genome. If an error occurs in any of these genes, the likely outcome is a muscle disorder. Most often mutations causing muscle disorders interfere with muscle cell structure, contraction and/or energy metabolism.[101-105]

Neuromuscular disorders constitute a wide spectrum of diseases characterised by muscle weakness and hypotonia. The primary defect may lie in the anterior horn cell of the spinal cord, in the nerve innervating the muscle, in the neuromuscular junction or in the muscle cell itself. Generally speaking, the primary defect causing muscle weakness in myopathies lies in the muscle itself, while the defect causing muscle weakness in neuropathies is located in the nerve. The weakness and the muscles affected vary between different disorders and even between individuals with the same disease. Usually these disorders are painless, or pain is not the predominant symptom. The disorder can affect some muscle groups specifically (e.g. proximal or distal) or even some specific muscles only. Neuromuscular disorders have been classified on the basis of clinical and histological criteria into several different groups and sub-groups (e.g. *Dystrophies; Duchenne muscular dystrophy (DMD), Myotonic dystrophy etc*), which often in turn are divided according to additional criteria, for example the degree of severity and underlying genetic defect. This thesis concentrates on the molecular genetics of some of the *congenital* (inborn) *myopathies* (Greek: myo = muscle, pathy = suffering/disease), in particular muscle disorders caused by mutations in genes encoding proteins for the muscle sarcomere.[107, 186] The disorders in focus are nemaline myopathy,[187, 188] distal nebulin myopathy (II) and cap myopathy[189].

Congenital myopathies are neuromuscular disorders which include clinically, histologically and genetically variable disorders defined on the basis of microscopic structural abnormalities in the muscle fibres.[190, 196] They may be caused by mutations in one or several possible genes. The symptoms are not directly correlated with the histological characteristics. In the congenital myopathies, the muscle tissue does not show any signs of inflammation or progressive muscle fibre death (necrosis).[43]

In this chapter, the most relevant disorders and entities for this thesis are briefly presented. Although in this project disorders of the thin filament were studied, some disorders of the thick and the third filament, which are interesting from the point of view of the thesis, are reviewed as well.

3.1 Thin filament disorders

Dysfunction of the structural or functional proteins of the thin filament is seen in a variety of muscle disorders, most of them shown to be attributable to mutations in the genes encoding these proteins. In addition, mutations in the same gene may cause several different disorders, while mutations in different genes can cause a very similar clinical outcome. For example, mutations in *ACTA1, NEB, TPM2,* and *TPM3* are all known to cause nemaline myopathy (NM), but can also cause a variety of different and overlapping disorders.[191] Of these, nemaline and cap myopathy are included here.

3.1.1 Nemaline myopathy (NM)

Nemaline (rod) myopathy is one of the most common forms of congenital myopathy. It was first described in 1963 with its characteristic protein aggregates, called nemaline bodies, seen with LM in Gomori trichrome-stained muscle fibres.[187] NM was, however, perhaps first seen by Dr. Reye 1958 in a patient in whom a mutation in *ACTA1* was later identified.[192] The weakness of NM patients is usually non-progressive or slowly progressive and proximal muscles are typically more severely affected than the distal counterparts, but the distribution of the affected muscles as well as the severity of NM varies enormously between patients.[193-196] The occurrence of NM is difficult to estimate reliably, but it is very rare. The incidence in Finland has been estimated to be 0.02 per 1000 live births.[197] Based on the clinical variability of NM, it is divided into six subclasses: severe, intermediate, and typical congenital forms of NM, and mild childhood or juvenile onset, adult onset NM and other forms.[193] The disease can be inherited as an AD or as an AR trait, and new dominant mutations are also quite common.[198-200] Nemaline body-like aggregates can be seen in other disorders, e.g. in the muscles of AIDS patients.[201]

3.1.1.1 Classification of NM

The European Neuromuscular Centre International Consortium on NM has categorised NM into six clinical classes:[202] The **severe form of NM** (*Fig 11A* and *B*) is often noticed even before birth due to the immobility of the foetus, or immediately after birth due to respiratory problems, severe muscle weakness, hypotonia and lack of spontaneous movements, difficulties in suckling and swallowing, and in some patients contractures and occasionally fractures. Recurrent pneumonias at a young age can be lethal, and many patients with severe

NM die soon after birth because of general severe muscle weakness and respiratory insufficiency. **The intermediate form of NM** is intermediate in severity between the severe and typical form. Muscle weakness is evident at birth, but spontaneous movements are present, and there are no major contractures or fractures. The child may never be able to sit unassisted or walk and possibly needs a wheelchair before the age of eleven. **Typical NM** (*Fig 11C and D*) is often evident at birth. The newborn is floppy and weak, but spontaneous movement and respiratory efforts are present. Feeding problems are common. There are no contractures or fractures. The most severely affected muscles are usually the facial, neck and proximal muscles of the limbs and axial muscles of the body. Later the distal muscles may become affected as well. The patients often have an expressionless face and a nasal voice due to the affected facial and bulbar muscles. These children achieve their motor milestones later than healthy children, and some need to move in a wheelchair in their teens. Patients with typical NM usually live rather normal lives despite their muscle weakness, but they may experience severe breathing problems. In **mild childhood or juvenile onset NM,** the onset is later, the muscle weakness milder and the facial muscles may not be involved. In **adult onset NM,** the patients can display a variety of symptoms. This form of NM might not have a genetic background at all. The patients belong to the group of **other forms of NM** if they, in addition to diagnosed NM, have for example cardiomyopathy, ophthalmoplegia, or other unusual clinical symptoms or histological findings, for example, intranuclear nemaline bodies observed in the muscle biopsy.[203] Sometimes the form of NM is difficult to define due to the overlap between different forms NM and symptoms of individuals.[193, 202]

3.1.1.2 Molecular genetics of NM

Mutations underlying NM have been identified in six genes which encode thin filament proteins of the sarcomere. These are the nebulin (*NEB*),[199] α-actin (*ACTA1*),[200] α-tropomyosin$_{slow}$ (*TPM3*),[198] troponin T1 (*TNNT1*),[204] β-tropomyosin (*TPM2*)[205] and cofilin 2 (*CFL2*)[206] genes. According to candidate gene and genome-wide linkage analyses, it can be assumed that at least a seventh NM gene is still to be identified.[206]

      NM-causing *NEB* mutations had been published, prior to the current study, in approximately 20 NM probands, causing all forms except for the adult-onset form of NM. The most common form of NM caused by *NEB* is the typical form.[194, 199, 207] All the mutations identified are AR and ~90 % compound heterozygous, and most often both of the mutations are unique to the family.[199, 207, 208] In addition, a founder mutation, a deletion encompassing the entire exon 55, has been identified in the Ashkenazi Jewish population.[208]

Mutations in *ACTA1* are thought to be responsible for 20 % of congenital myopathies.[209, 210] Most of the over 140 mutations identified in *ACTA1* cause NM.[210] Mutations in *ACTA1* can, however, cause a variety of other disorders as well: actin myopathy (patients have aggregates consisting of actin in their muscle fibres),[200, 211] intranuclear rod myopathy,[210, 212] congenital fibre type disproportion (without rods) (CFTD) and myopathy with core-like lesions (*Fig 12*)[195, 213] *De novo* dominant **mutations** are the most common cause in severe, even lethal, forms of NM, but *ACTA1* mutations can also be recessive or inherited in a dominant fashion, and underlie mild forms of NM.[194, 200, 210, 214, 215] Somatic mosaicism has been seen in two families with mildly affected parents and severely affected children.[216]

NM-causing mutations in **the β-tropomyosin gene, *TPM2*,** were first identified in 2002 in two families. One of the mutations was a likely *de novo* AD mutation causing NM and another was an AD mutation identified in an affected mother and her affected son leading to CFTD.[205, 217] A dominant mutation said to cause distal arthrogryposis (multiple congenital contractures of the joints) type 1 (DA1) was identified soon after this, in 2003.[218] Mutations in *TPM2* are known to cause other entities as well.[219] These will be discussed further in the section "Results and Discussion".

Mutations in the gene for α-tropomyosin slow, *TPM3*, have been identified in both recessive and dominant forms of NM and in CFTD.[198, 219-221, 260, 263] These will be discussed further in "Results and Discussion".

Only one mutation, a founder mutation of the Old Order Amish population, has been identified in the **troponin T1 gene TNNT1**. The mutation detected is a nonsense mutation in exon 11 of the gene in homozygous form causing progressive, severe NM which typically leads to the death of the patient during the second year of life. The incidence of the disorder among the Old Order Amish is approximately 1:500.[204] One AR mutation has been identified in the gene encoding **cofilin-2, CFL2**, in one sib pair from a large family of Middle Eastern origin. The mutation is a missense mutation A35T and in its homozygous form it caused a clinical picture resembling the typical form of NM, although no foot drop or weakness of the facial muscles was observed in these patients. The biopsies of the patients displayed nemaline bodies and a few minicores.[206]
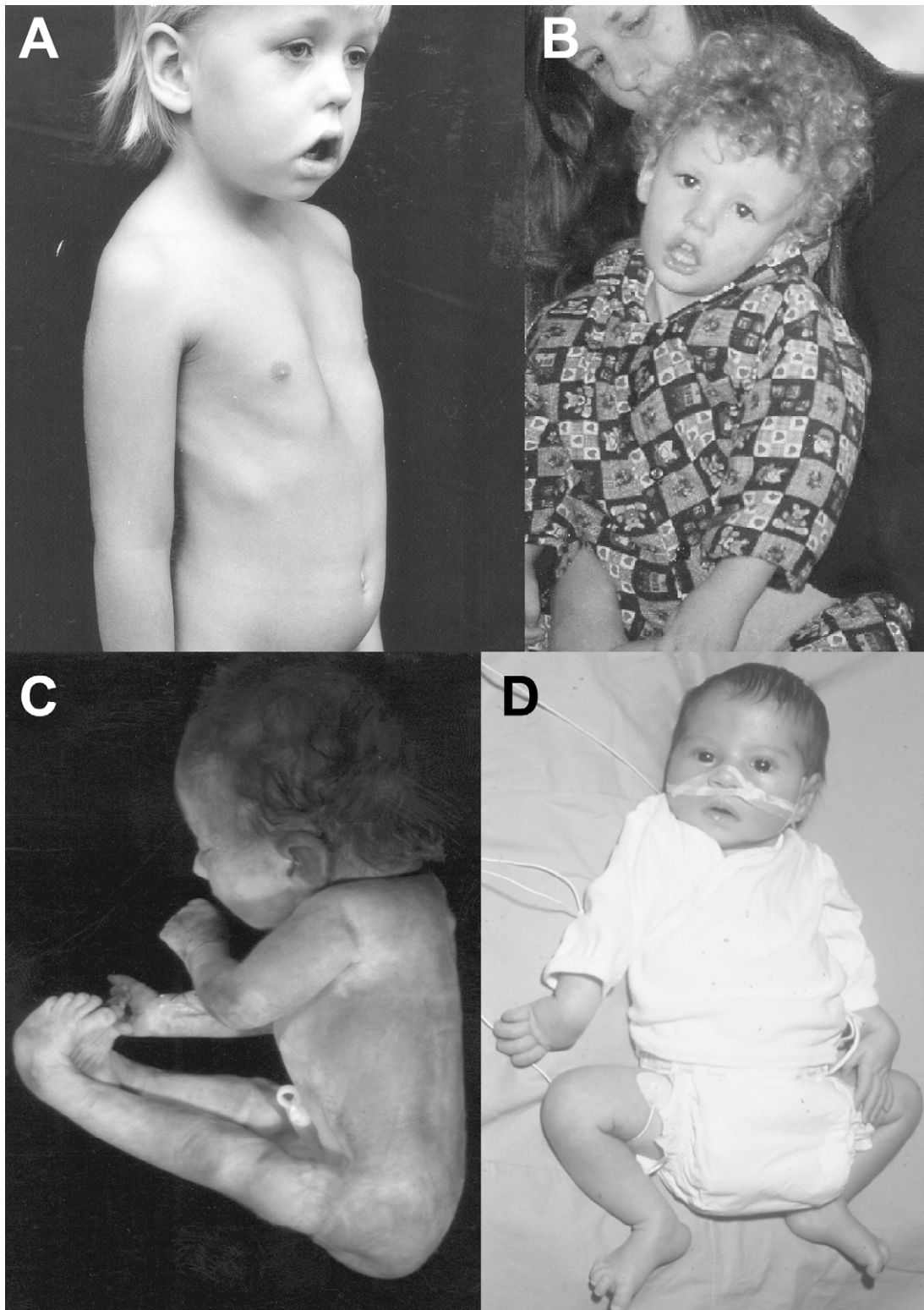
Figure 11. A. Typical NM caused by mutation of *NEB*. **B.** Typical NM caused by mutation of *ACTA1*. **C.** Severe NM caused by mutation of *NEB.* **D.** Severe NM caused by mutation of *ACTA1.* (Printed with permission of the families)

3.1.1.3 Histology of NM

Histological features mentioned in this chapter are nemaline bodies (rods), fibre type disporotion (CFTD) and cores: **Nemaline (rod) bodies**, (*Fig 12A*) are protein aggregates best seen in the cross-sections of Gomori trichrome-stained muscle fibres. They usually arise from the Z discs, but may in some cases be intranuclear. In **CFTD** *(Fig 12B)* describes the situation in which the type I fibres are smaller than type II fibre sizes [220, 223] **Central cores** or **cores** *(Fig 12C)* are compact regions or zones seen in the centres of cross-sections of myopathic fibres and, being devoid of mitochondria, they are easily identified in oxidative enzyme reactions. In longitudinal section the cores run along the length of the fibre (*Fig 12C2*). *Minicores* are smaller than (central) cores in both in cross-section and in longitudinal sections.[99]
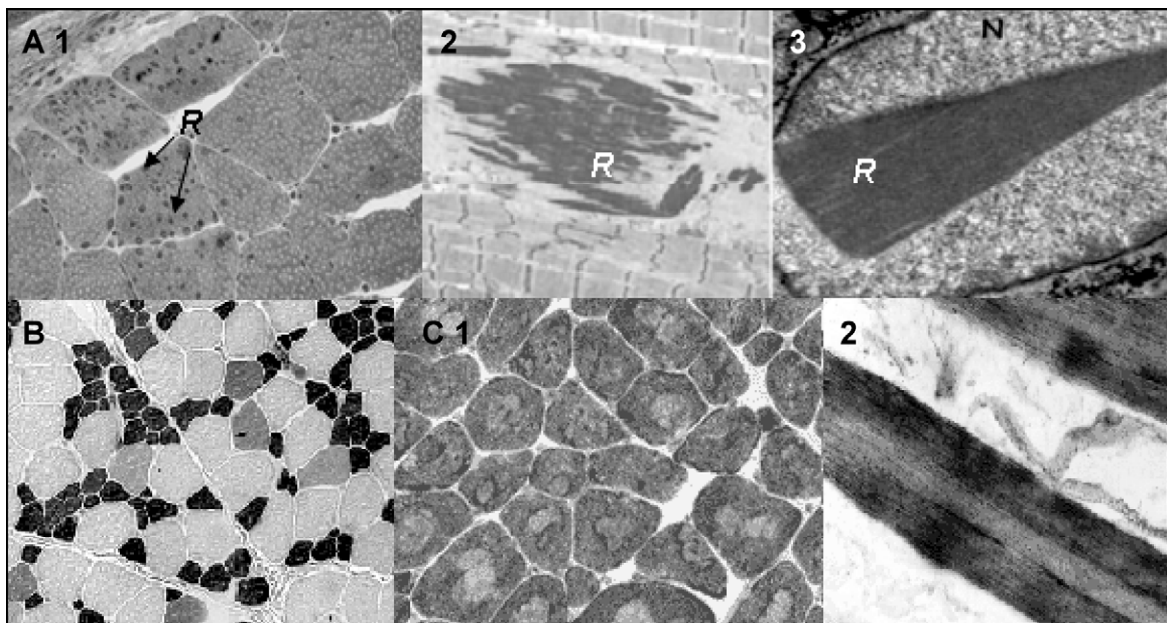
Figure 12. **A. Rods** R = rod **1.** Nemaline rods / bodies (dots pointed by arrows) in Gomori trichrome-stained in cross-section of muscle fibres (LM), **2.** in longitudinal section of muscle fibre (EM), **3.** intranuclear rod: N = nucleus, (EM). **B. CFTD**: small type I fibres appear dark in ATPase (pH 4.3) stained cross-section. Type II fibres appear pale. (LM), **C.** central core structures: 1. in longitudinal section (LM), 2. in cross section (LM) 3. in longitudinal section (EM). (Figures kindly provided by Professor Kathryn North and Dr. Nigel Clarke, University of Sydney, Australia, Professor Norma Romero, Groupe Hospitalier Pitié-Salpêtrière, Paris, France and Professor Caroline Sewry, Imperial College, London, UK) LM = light microscope, EM = electron microscope.

The nemaline bodies are the defining characteristic of NM. Nemaline bodies (*Fig 12A*) are protein aggregates containing α-actinin, α-actin and other Z-disc and thin-filament proteins. The nemaline rods have no enzyme activity. They usually accumulate under the sarcolemma, but they may be centrally located as well.[187, 196, 222] If one is using the **light microscope** (LM), the best method to identify nemaline bodies (rods) is the modified Gomori trichrome stain, where the bodies appear as very dense dark red rods against a turquoise background (*Fig 12A1*). Gomori's trichrome staining procedure combines a plasma stain and a connective fibre stain. Hematoxylin and eosin (HE) stained nemaline bodies may appear as non-specific pink bodies. It can be said that in the HE section, hematoxylin stains nucleic acids and other basophilic structures purple-bluish, whereas eosin stains protein-rich eosinophilic structures such as cytoplasm red-pink.[99] Viewed using the **EM** (*Fig 12A2*), the smaller nemaline bodies appear as thickened disorganised Z discs but larger bodies are not clearly connected to the Z discs. Nemaline rods are osmiophilic. Osmium tretroxide is used to enhance contrast in the EM; the rods appear darker when they bind osmium.

Both or one of the type II muscle fibre types are usually deficient in NM, and the type I fibres tend to be small, in many cases fulfilling the mathematical definition of CFTD (i.e. the difference between type I and type II fibre sizes should be at least 12 %)[220, 223] (*Fig 12B*). CFTD may be considered a disorder *per se*, if no rods or other specific structural abnormalities are present in the biopsy.[221] Nemaline bodies are mostly seen in type I fibres, especially in childhood cases, but they can also be seen in type II fibres.[202] The proportion of the affected fibres *in situ* may vary greatly between different muscles and even within the same biopsy, and may change with age. Fibres containing nemaline bodies are often quite atrophic (wasted) or hypotrophic (decreased in size), but both atrophic and hypertrophic (increased in size) fibres may be affected. Muscle biopsies of patients with NM can, in addition, show overlapping histology with central core (*Fig 12C*) / core-rod disease, and other related entities, complicating the histological definition of NM in some cases. Clinically the patients affected with the entities mentioned may be very similar to NM patients. Usually these patients with additional histological features are classified into the category of "other forms" of NM.[196] Nuclear rods (*Fig 12A3*) and cardiac nemaline bodies can rarely be seen and they tend to correspond with more severe forms of NM. The quantity of nemaline bodies does not correlate with the severity of NM.[196, 222, 224-226]

3.1.2 Cap myopathy

"Cap myopathy", also known as "cap disease" is a congenital myopathy first described in 1981 by Fidzianska and colleagues in a seven-year-old boy who was examined for non-progressive muscle weakness, hypotonia, skeletal dysmorphism and respiratory insufficiency. The boy died at the age of eight years because of respiratory complications.[189] Subsequently after reporting the first case of Fidzianska's, she has reported three additional cases. All the patients had childhood onset but clinically variable cap myopathy.[227] In addition, in 1992 Gibbels reported a case with possible cap myopathy or an unusual form of NM[228] and in 1996 Martland presented a cap myopathy sib pair.[229] In total, the cases published to date include 15 cap myopathy patients in 13 families.

The more recent descriptions of cap myopathy patients show that this entity can vary clinically and histologically and may overlap with NM. The cases are often sporadic but there are familial cases showing AR or AD inheritance. Marked differences in severity and distribution of the muscle weakness have been observed between different individuals, but the facial and the neck muscles are affected in the majority of patients.[230-233] In some families, certain affected family members have NM or an undefined muscle disorder, while others are diagnosed as having cap myopathy.[228, 230, 232] The molecular genetic aspects of cap myopathy are discussed under "Results and Discussion".

3.1.2.1 Histology of cap myopathy

The disorder is identified on the basis of the cap-like structural abnormalities lacking ATPase activity on one side of the cell/fibre under the sarcolemma (the muscle cell membrane) (*Fig 13, p. 45*). These caps have been revealed to be massive protein aggregates consisting of disarranged thin filaments with enlarged Z discs, while myosin filaments seem to be absent, and in the trichrome stain, "cap" structures exhibit a granular appearance.[189, 227, 234] The caps are often positively labelled with antibodies against desmin, α-actinin, α-actin and tropomyosin, and have been shown to be negative for adult isoforms of myosin.[227, 232, 233]

In cap myopathy patients, type I fibres are often hypotrophic and type I uniformity/predominance is common.[189, 227-233] The histological findings may vary even in biopsies of the same person taken at different time points and/or from different parts of the body.[228, 230, 232] It has been postulated that the severity of the disorder may correlate with the number of affected fibres.[227]

Table 1. Cap myopathy families and patients published

Family/ Patient	Onset & Clinical features	Biopsied at	Biopsy findings
1	At birth; progressive, death at 8 yrs from respiratory failure	7 yrs	Caps in 70-75 % of fibres
2	At birth; progressive, death at 14 yrs	12 yrs	Caps in 70-75 % of fibres
3	Childhood; slowly progressive	18 yrs	Caps in 20-30 % of fibres
		21 yrs	Caps in 20-30 % of fibres
4	Childhood ; slowly progressive	15 yrs	Caps in 20-25 % of fibres
5 cons.,	Neonatal; Slight improvement of muscle strength reported at age 7	20 months 7 yrs	Caps in 13 % of fibres Rods in 1 % of fibres, caps in 0,1 % fibres
6.1.(S1) cons.	Neonatal; slowly progressive. Ambulant at 15 yrs.	5 yrs 16 yrs	Normal Caps in 40 % of fibres
6.2.(S2) cons.	Infantile; slowly progressive. Ambulant at 33 yrs.	5→15 yrs 33 yrs	Normal Caps in 20 % of fibres
6.3.(H.C)	Neonatal; **severe NM,** death at 4 yrs from pneumonia	30 days	Nemaline bodies in 60 % of fibres
7	Infantile, non-progressive	4 yrs	Caps in 20 % of fibres.
8.1.(S1)	Congenital; mild non-progressive.	13 yrs	Caps
8.2.(S2)	A brother with similar symptoms.		
9	Neonatal: slowly progressive	12 yrs	Caps
		33 yrs	Caps
10.1 (D)	Neonatal; slowly progressive	2,5 yrs	No aggregates
		26 yrs	Caps
10.2.(M)	Neonatal;**slowly progressive NM**	32 yrs 57 yrs	No aggregates nemaline bodies
11	Infantile; slowly progressive. At 42 yrs, uses a wheelchair outdoors.	15 yrs 30 yrs	No aggregates Caps in 15 % of fibres
12	Infantile; slow improvement	3 yrs	No aggregates seen using LM, but in EM small caps.
13	Neonatal; non-progressive	5 yrs	Caps

S = sibling, H.C.= half cousin, D = daughter, M = mother, cons. = consanguineous1-4:[189, 227]5:[228] 6:[230] 7:[231] 8:[229] 9:(IV) 10:[232]11,12,13:[233]
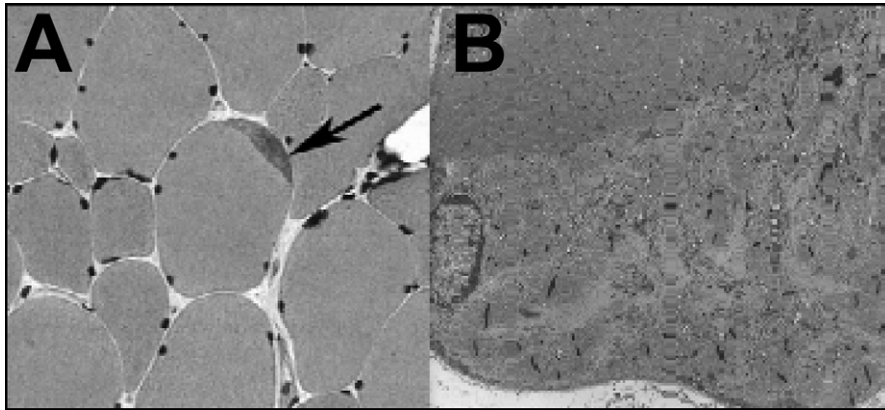
Figure 13. Cap structures consist of disorganised thin filaments and usually locate on one side of the cell/fibre under the sarcolemma (the muscle cell membrane). **A.** LM, **B.** EM

3.2 Thick filament disorders

Thick filament myopathies include a group of disorders called hereditary myosin myopathies. Mutations in any of the genes encoding myosin heavy or light chain proteins could potentially underlie a thick filament myopathy. To date, mutations have been identified in the genes for two embryonic isoform (*MYH3* and *MYH8*) and two adult isoform myosin heavy chain genes (*MYH2* and *MYH7*). These mutations cause a wide range of disorders from arthrogryposis to distal myopathies, progressive muscle weakness as well as cardiomyopathies or ophthalmoplegia, i.e. paralysis of the extraocular muscles. At present, no mutations have hitherto been identified in myosin light chain genes.[235, 236]

3.2.1 Laing distal myopathy and other *MYH7*-related disorders

MYH7 (in chromosomal region 14q12) encodes the β-cardiac and type I muscle fibre-specific myosins.[235] Most, over 200, mutations in *MYH7* cause cardiac myopathies without any skeletal muscle weakness.[236] Some, however, underlie Laing distal myopathy (LDM).[237] LDM is usually a childhood onset distal myopathy (although a few adult-onset cases have been reported) typically first diagnosed due to weakness of the toes ("hanging toe") and ankle dorsiflexors. This usually slowly progressive disorder next affects finger extensors and neck flexors, and finally the weakness spreads to the upper muscles of the limbs and the facial muscles. In most cases patients retain their ability to walk till the fifth decade of life.[237, 238] Two patients have been reported to have both LDM and cardiomyopathy.[239, 240] No specific histological features define the diagnosis, but some biopsies have revealed excessive

45

variation in fibre size, central nuclei and streaming of the Z discs.[241] Most of the mutations causing LDM have been identified in the tail region of *MYH7*, in exons 32, 34, 55 and 36. These exons encode the light meromyosin region of the protein.[238] Recent mutations, however, in patients with LDM with cardiomyopathy have been located in exons 14 and 16 encoding the globular head of MYH7.[239, 240] The mechanisms by which the mutations cause the disorder were first believed to be due to interrupted coiled-coil structure of the protein, but the discovery of the mutations in exons 14 and 16 indicate that disrupted binding to titin, myomesin or M-protein may play important roles in the pathogenesis of LDM.[236]

In addition to LDM, mutations in *MYH7* may cause myosin storage myopathy in which all mutations identified to date are located close to the assembly competence domain (functioning in the polymerisation of the molecule) of the protein in the distal rod region of the myosin heavy chain. It is believed that this situation prevents myosin dimers from assembling into the thick filaments.[242, 243]

3.3 Third filament disorders

The backbone of the third filament is titin encoded by *TTN*. In addition to mutations in *TTN*, mutations in calpain 3, myotilin and telethonin are known causes of a variety of third filament diseases, such as the limb girdle muscular dystrophies (LGMDs) (a group of disorders predominantly affecting the pelvic and shoulder girdles),[244-246] cardiomyopathies,[247, 248] tibial muscular dystrophy (TMD),[249, 250] and hereditary myopathy with early respiratory failure.[178, 251] The first mutation in a third filament protein-encoding gene, calpain 3, was identified in a LGMD2A-patient in 1995. To date some 300 mutations have been found in this gene.[244, 252]

3.3.1 Tibial muscular dystrophy and LGMD2J

Tibial muscular dystrophy (TMD) is an adult-onset muscular disorder. Typically the onset is after the mid-forties and the disease specifically affects the anterior tibial muscles, weakening the dorsiflexion of the ankles, and later the long toe extensors. The muscle biopsy displays variability in fibre size and central nuclei. During the later course of TMD, fibrosis and fatty replacement become evident. The sarcomere structure does not seem to be disturbed.[249] TMD is an AD disease caused by heterozygous mutations of the M-line titin region. To date, six TMD-causing *TTN* mutations have been identified.[250, 253, 254] Recent studies suggest that truncating mutations in C-terminal *TTN* would lead to a more severe form of TMD.[254]

While one of the Finnish founder mutations (an in-frame deletion-insertion of 11 nucleotides in the last exon, exon 363, of *TTN* [250]) in heterozygous form is known to cause TMD, in the homozygous state, it causes a more severe muscle disorder, LGMD2J. In this disorder, the homozygous *TTN* mutation leads to loss of ambulation by the age of 25 years and dystrophy with fatty replacement affecting all the large muscles.[249]

Aims

The aims of this study were to:

1. Optimise the mutation screening of *NEB* using denaturing High Performance Liquid Chromatography (dHPLC).

2. Estimate the occurrence in an international sample collection of the deletion of 2 502 bp including *NEB* exon 55 identified in the Ashkenazi Jewish population by Anderson and colleagues (2004).

3. Identify novel mutations in *NEB* using dHPLC and to discern any emerging genotype-phenotype correlations.

4. Detect either a novel NM gene or a founder mutation in a known NM gene in consanguineous Turkish patients using linkage analysis and sequencing of candidate genes.

5. Identify causative gene(s) associated with NM-related myopathies.

Materials and methods

1 Families and control individuals

A total of some 300 families were included in this study. The families were from a variety of countries and of various ethnic origins. The research project has been approved by the Ethics Committee for Pediatrics, Adolescent Medicine and Psychiatry of the Helsinki University Central Hospital. The patients or their guardians have given their consent for having their samples and data included in the study. Control DNA samples were provided by Centre d´Etude du Polymorphisme Humain (CEPH) and the Finnish Red Cross Blood Transfusion Service. All control individuals were anonymous.

1.1 Families included in the dHPLC analysis (I)

Patients and families were selected on the basis of linkage results and/or the clinical and histological picture of the patient. All families selected for *NEB* mutation analysis had likely AR NM and mutations in *ACTA1* had previously been excluded by sequencing (personal communication with Prof. Nigel Laing). In familial and/or consanguineous families, *TPM2*, *TPM3* and *TNNT1* had been excluded by sequencing or linkage analyses as part of this study or in previous projects. The first analysis batch (I) included samples from 45 NM patients and/or their parents. All known Finnish NM families were included. Samples from 42 patients or their families were included in the second dHPLC series, when the dHPLC analyses were further optimised and each PCR-fragment was run at two or three analysis temperatures (unpublished).

1.2 Screening for the deletion of exon 55 (II)

In total, 208 probands in our cohort, i.e. all the NM probands with no previous genetic cause identified or with the second of the *NEB* mutations unidentified. As part of this project, results of our collaborators in Perth, Australia, and Boston, USA, were collected and summarised. In all, the results of 355 probands were included in this study.

1.3 Identification of distal nebulin myopathy (III)

Two of the four Finnish families with the distal nebulin myopathy described in this study were first selected for another ongoing study, and due to their geographic origin, their DNA samples were used as ontrols in the haplotype analyses performed using SSCP. Altogether, in this study, 12 Finnish and 12 families from other countries were included.

1.4 Families included in candidate gene analyses (IV)

As part of this study, some 300 families were classified for analyses according to their clinical and histological data (the form of NM or other entity, the mode of inheritance, ethnic origin, and any useful clinical or histological clues). On the basis of this classification, linkage studies and/or direct sequencing of possible candidate genes were performed. In total, six known NM and 20 candidate genes were analysed either by haplotyping or by sequencing (*Table 2*).

1.5 Families included in the genome-wide linkage analysis (V)

In order to identify the seventh NM gene and/or a founder mutation in a known NM gene causing NM in the Turkish population, 12 Turkish families were included in the study; ten consanguineous and two non-consanguineous multiplex families.

2 Methods

2.1 Summary of the common methods used

DNA isolation and purification   **I –V**
Saliva: Oragene^{TM} saliva kits (*DNA Genotek Inc., Ottawa, Canada*)
Blood, fibroblasts of cultured skin biopsies or myoblasts cultured from muscle biopsies: Genomic
DNA Purification Kit (*Gentra Systems, Minneapolis, USA*)
RNA isolation and purification   **I, III-V**
Muscle biopsy or a myoblast culture: RNeasy Mini Kit (*Qiagen, West Sussex, UK*)

Whole-genome amplification   **I, III-V**

Phi29-DNA-polymerase / Illustra GenomiPhi DNA Amplification Kit *(GE Healthcare, Chalfont St. Giles, UK)*

PCR and RT-PCR   **I, III-V**

Basic procedures used and described in the articles

Sequencing   **I, III-V**

Sequencing performed using an ABI 3730 DNA Analyzer (*Applied Biosystems, Fosters City, USA*)

Sequence analysis programmes used:

- BioEdit (*Ibis Biosciences, Carlsbad, USA*)

- Sequencher 4.1. (*Gene Codes corporation, Ann Arbor, USA*) programmes.

If a myoblast cell culture or a muscle biopsy was available, RT-PCR was performed and the patient's *NEB* cDNA sequenced.

Analysis of the deletion of the whole exon 55   **II**

As described by Anderson and colleagues.[208]

Cloning and minigene method   **I**

Described in details in the article

2.2 Genotyping, creating haplotypes and linkage analysis (I – V)

Genotyping: SNPs were used when analysing known NM-causing genes, and microsatellite markers when analysing candidate genes. SNP analysis was performed by SSCP or sequencing. The SNPs were chosen on the basis of data gained from the previous studies and from the SNP consortium database. The genotypes are available at http://www.ceph/fr/cephdb. Using fluorescently labelled primers, microsatellite markers were analysed using GeneMapper version 5.0 (*Applied Biosystems, Foster City, USA*). Markers closest to the gene were selected using the NCBI MapViewer (http://www.ncbi.nlm.nih.gov/mapview/). The genome-wide analysis was performed at the Finnish Genome Centre, Biomedicum Helsinki, using a 10 cM map of microsatellite markers. **Haplotyping:** In the candidate gene analyses, haplotypes were created manually for all members available of the multiplex and/or consanguineous families. If no other family members were available for the analysis, the sample of the affected member alone of the consanguineous families was included. The Finnish Genome Centre provided the haplotypes of the genome-wide scan. **Linkage analysis:** The pair-wise linkage analysis was performed using LINKAGE tool (V). The multipoint analysis was performed using Genehunter (V).

2.3 dHPLC analysis (I)

For *NEB* analysis, primers were designed to 159 of the 183 *NEB* exons (the 8 exon triplication in the middle of *NEB* was excluded from the dHPLC analysis) covering one or two exons per PCR reaction. According to the melting temperatures of each PCR fragment, unique analysis conditions were designed (running temperature, acetonitrile/TEAA buffer gradient etc) for each fragment using the software of dHPLC provided by Wave Transgenomic Ltd (*Transgenomic, Omaha, USA*). Chromatograms were analysed manually comparing the result of each individual chromatogram peak to the other peaks in the analysis plate exon by exon. Samples showing an abnormal chromatogram peak were sequenced for a possible mutation.

2.4 Web-based tools used (I, III-V)

Primer design:
Primer3 (versions 0.3.0 and 0.4.0) (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
Analysis tools to predict the effects of the possible mutations:
- Sequence comparison tools NCBI-BLAST at http://blast.ncbi.nlm.nih.gov/Blast.cgi
- Analysis of ESE factors using ESEfinder 2.0 at http://rulai.cshl.edu/tools/ESE2/
- Coiled coil analysis tool COILS at http://www.ch.embnet.org/software/coils/COILS_doc.html
- Phosphorylation site search NetPhos 2.0 at http://www.cbs.dtu.dk/services/NetPhos/
- Interspersed repeats of DNA sequences RepeatMasker at http://www.repeatmasker.org/
- The tolerance of a particular amino acid for a change was predicted using SIFT at http://blocks.fhcrc.org/sift/SIFT.html

Table 2. The known NM genes and candidate genes analysed by haplotyping and/or sequencing (and/or in the case of *NEB* analysed using dHPLC) in this study (known NM genes in bold). The numbers indicate the number of families and/or probands analysed for the given gene.

GENE (CHR. LOCATION)	HAPLOTYPED	SEQUENCED
***ACTA1* (1q42.13)**	**41**	
***CFL2* (14q12)**	**20**	
***NEB* (2q22)**	**95**	**97** (seq or dHPLC)
***TNNT1* (19q13.4)**	**25**	**4**
***TPM2* (9p13.2-1)**	**18**	**40**
***TPM3* (1q21.2)**	**39**	**45**
ACTN2 (1q42-43)	8	
ACTN3 (1q13.1)	3	
CAPZA1 (1p13.2)	6	
CAPZA2 (7q31.2-3)	6	
CAPZB (1p36.1)	5	
DES (2q35)	19	4
FLNC (7q32-35)	11	
LMNA (1q21.2-3)	6	
MYO (5q31)	29	
MYPN (10q21.3)	8	
NEBL (10p12)	2	
OBSCN (1q42.13)	6	
RYR1 (19q13.1)	41	
SEPN1 (1p36.13)	19	24
TCAP (17q12)	4	
TPM1 (15q22.1)	2	
TTN (2q31)	7	
YL1+TMOD4 (1q12)	20	19
ZASP (10q22.3-23.2)	4	10

Results and discussion

1 Mutation analyses of *NEB* (I - III)

Using dHPLC, RT-PCR and direct sequencing, 115 *NEB* mutations had been identified in 96 families by March 2009 (*Fig 14, Table 3*): the typical form of NM was diagnosed in 44 % of cases, severe in 25 %, mild in 9 %, intermediate in 5 %, other/unusual forms of NM in 3 %, and 9 % of the analysed probands had not been classified because of insufficient data (I and unpublished results). Distal nebulin myopathy (DNM) was diagnosed in four (III) and core-rod myopathy in one family (*Romero&Lehtokari et al., accepted in Neurology*) (5 %). Mutations in *NEB* have not hitherto been identified in patients with adult onset NM. All the mutations identified to date are summarised and discussed here in order to have all available data included and provide a broader view on the phenotypes caused by these mutations. NM-causing mutations, i.e. 92 % of the *NEB* mutations identified, are compound heterozygous and most often both of the mutations are unique to the family. Mutations are found along the whole length of the gene and no true mutational hotspots or recurrent mutations are evident. In this genomic area of 249 kb, the occurrence of different kinds of mutations seems to be rather equal (point mutations vs deletions/insertions). Point mutations account for 59 % of all the *NEB* mutations (missense 14 %, nonsense 17 % and splice-site 28 % of mutations). Small deletions, duplications or insertions (38 %), leading to either frameshifts (all except one in-frame deletion) and subsequent premature termination codons, or to disturbed splice signals (2 %) account for 39 % of the mutations identified. A 2,502 bp deletion erasing the whole exon 55 identified in the Ashkenazi Jewish population[208] accounts as a single mutation for approximately 1 % of the mutations. This mutation was analysed in a total of 355 probands (of the sample cohorts of Helsinki, Boston and Perth) and found in 14 of these. This is 2 % of all 702 NM probands from which these three main NM research centres had DNA samples. All the patients carrying the deletion shared the same haplotype and thus the mutation was shown to be a founder mutation (II). If the mutations are organised according to their predicted or known effect on the RNA or protein, truncating (nonsense and frameshift) mutations account for 54 %, splice-site mutations for 30 % and missense mutations for 14 % of all identified *NEB* mutations. The majority of the splice-site mutations skip one exon from the transcripts (in-frame exon skipping). The missense mutations are all located in known binding sites for other proteins and/or change highly conserved amino acids.
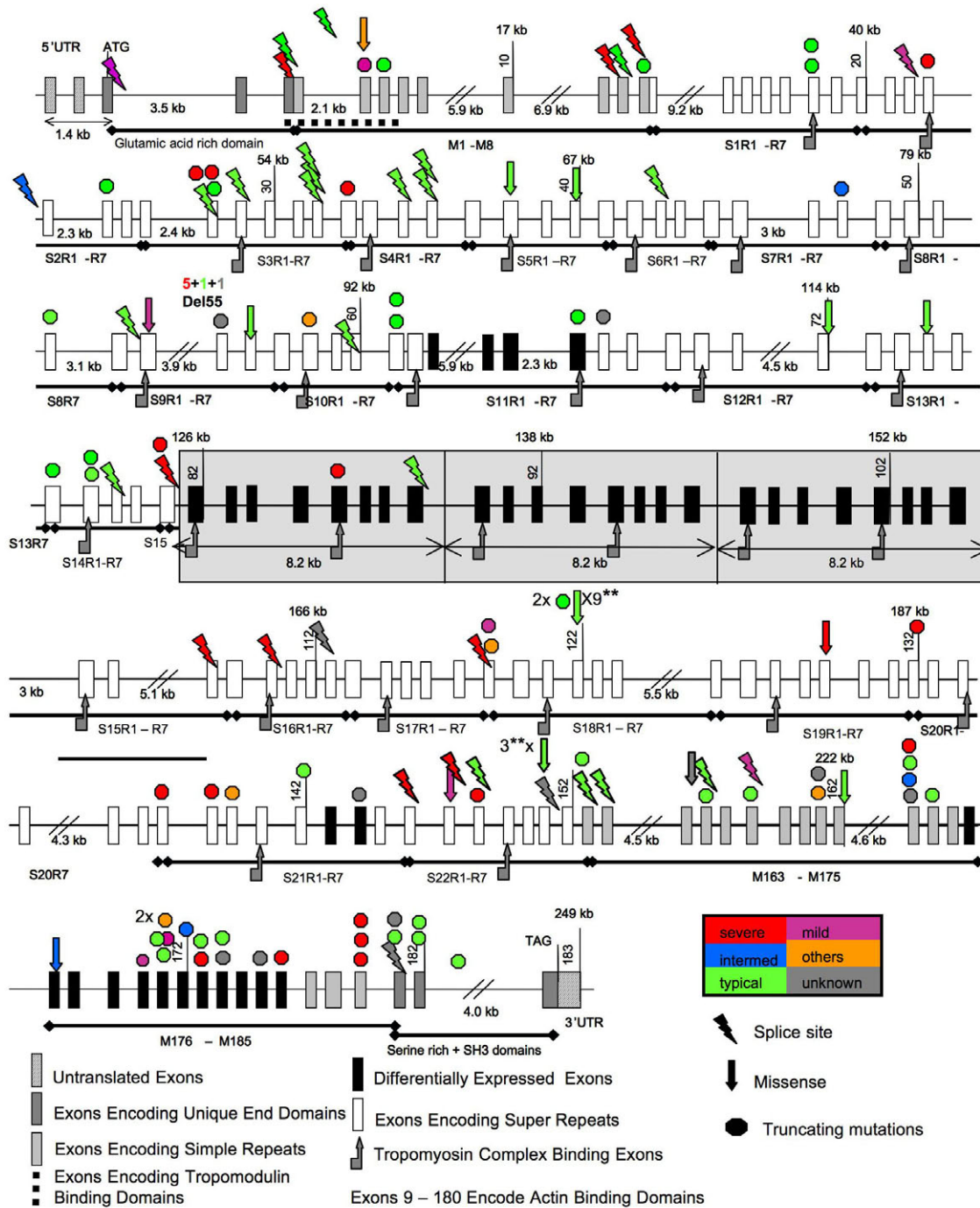
Figure 14. The nebulin gene and mutations identified within this gene. ** Finnish mutations identified in more than one family. The triplication of 8 exons on grey background.

Nonsense and frameshift mutations have been claimed to result in truncated nebulin molecules or nonsense-mediated mRNA decay. Immunohistochemical studies with antibodies against nebulin domains in the C-terminal region (Z-disc-binding region) and domains near the I band indicate, however, that a nebulin allele with mutations in the 3' end is usually expressed and the C terminus of the protein is usually present.[199, 255-257] One of the reasons explaining this, may be the differences in the efficiency of different termination codons and the nucleotides following it.[258] In one sib pair with a homozygous nonsense mutation in exon 181, staining with an SH3 antibody detecting the last C-terminal domain of nebulin, was negative, whereas the other antibodies produced positive results.[255] Such analyses have not been performed in cases of truncating mutations identified in the 5' or mid-region of nebulin. It remains to be established whether premature termination codons far from the normal stop codon result in nonsense-mediated mRNA decay.

Table 3. Families with different forms of NM or other myopathies, and the types of nebulin mutations identified in them. Note that some of the mutations have been identified in more than one family and may underlie different forms of NM.

Disorder	SEV NM	INT NM	TYP NM	MILD NM	OTH NM	UNK NM	DNM	CRM
Families (n)	30	9	51	12	5	14	2	2
Mutations (n)	24	5 *	42 *	9	3	9	4	1
Truncating (%)	*57*	*45*	*47*	*42*	*80*			
All (n)	**17**	**4**	**24**	**5**	**4**	**8**		**2**
Hez	15	4	22	4	4	8		2****
Hoz	2	-	2	1***	-	**-**		
Splice site (%)	*33*	*11*	*33*	*33*				
All (n)	**10**	**1**	**17**	**4**	-	**4**		
Hez	10	1	15	4	**-**	4		
Hoz	**-**	**-**	3**	-	**-**	**-**		
Missense (%)	*7*	*33*	*18*	*17*	*20*			
All (n)	**2**	**3**	**9**	**2**	**1**	**1**	**2**	
Hez	2	1	9	2	1	1	-	
Hoz	-	**-**	-	**-**	**-**	**-**	2	
Del55(!)/i.f. (%)	*3*	*11*	*2*	*8*	**-**			
All (n)	**5**	**3**	**1**	1 **i.f.**		**1**		
Hez	4	3	1			1		
Hoz	1	**-**	**-**			**-**		

SEV = severe, INT = intermediate, TYP = typical, OTH = other, UNK = unknown form of NM, DNM = distal nebulin myopathy, CRM = core-rod myopathy, n = number. ! the deletion of exon 55 is the same in all patients. i.f. = in-frame deletion, Hez = heterozygous, Hoz = homozygous *In one family one sibling has the intermediate and another the typical form of NM ; **The same mutation found both in homozygous and heterozygous form; *** the same mutation (in exon 171) causes the typical form of NM as compound heterozygous form in another family. **** The same mutation in ex171 identified in a proband classified as belonging to the group "other" forms of NM.

The performed RT-PCR and minigene analyses (I) support the assumption that splice-site mutations would lead to skipping of one exon from the transcript. Although most transcripts lack one exon, a small proportion of the transcripts are normal, i.e. the splice-site mutations are leaky (I). Most (69 %) of the splice-site mutations, as well as the deletion of exon 55 (II), shorten a super-repeat with 35-66 amino acids, disrupting its organisation. This, in turn, is predicted to lead to impaired nebulin-tropomyosin interactions, and to reduce actin-binding sites in the molecule and thus to interfere with the contractile function of the sarcomere (I). The splice-site mutations in intron five, causing skipping of exon five, likely disrupt the N terminus of nebulin. This region binds to tropomodulin, a crucial capping protein defining the proper length of the thin filament at its pointed end.[123, 144, 146, 259] The N-terminal mutations interfering with the tropomodulin-nebulin interaction may result in incorrect assembly and unstable thin filaments. A splice-site mutation leading to skipping of exon 3, in which the translation initiation codon lies, was identified in a patient with mild NM (unpublished). This indicates the possible usage of the next initiation codon (ATG) the transcription machinery encounters, i.e. codon 82 in exon 5. The same patient has another splice-site mutation on the other allele and this is predicted to cause skipping of exon 22 (unpublished). In addition to the mutated form of nebulin, the leaky nature of the splice-site mutations would result in normal nebulin isoforms also in the muscles of this patient, which may explain the mild disease phenotype.

It is difficult to determine with certainty whether a missense change in nebulin is a pathogenic mutation or not. The missense mutations identified and presented here are all located in binding sites of actin or tropomyosin, or are predicted to affect the α-helical secondary structure of nebulin.[142] *NEB* is rich in polymorphic changes; over 200 SNPs and various di- or trinucleotide repetitions have been identified. The gene seems to be quite tolerant to missense changes and if the change does not alter a conserved amino acid, it is usually not expected to be harmful. On the other hand, some of the rare missense changes we have identified might be mutations, but due to lack of knowledge of all significant domains, e.g. binding sites of many proteins binding to nebulin, we cannot be sure whether the change is a SNP or a mutation. Hopefully, the ongoing functional analyses will shed light on the effects of these changes on the nebulin protein as well as on the pathogenesis caused by the known mutations. In addition, since nebulin is rich in polymorphisms, the significance of the different SNPs and haplotypes as modifiers of the phenotype is not known. Especially in very mildly affected patients with NM, it may transpire that some particular SNP or haplotype in fact affects the phenotype. Naturally, the surrounding genome and environmental factors modify the outcome of the disorder, as seen in the sib pairs with the same mutations but different phenotypes. Possible changes deeper in the introns, promoter or regulatory elements, in addition to larger genomic changes (such as the deletion of exon 55) should be considered in

the cases where the first causative mutations have been detected but the second not. In addition, the new knowledge suggesting that nebulin has a significant role in Ca^{2+} homeostasis[131] and subsequently in regulating muscle contraction diversifies the possible mechanisms through which mutations in *NEB* might lead to a muscle disorder.

Project I revealed the true need for a method for effective *NEB* analysis. Using dHPLC and direct sequencing of DNA or, when possible, cDNA, we have detected more than 100 mutations, some of them previously missed by SSCP. Careful optimisation of dHPLC is essential. We have identified three changes detectable at one, but not at another running temperature. Due to the nucleotide composition and sequence of some fragments, it is impossible to create conditions in which 100 % of the given fragment would be optimally analysed. dHPLC should, however, identify 98 % of changes despite the analysis conditions not being theoretically fully optimal.[84] Since dHPLC detects any heterozygous change, it cannot distinguish between mutations and normal genomic variation. That a normal variation may mask additional changes in the DNA has been observed in at least one case. dHPLC is the most optimal method for scanning large numbers of patients for known mutations. The need for additional mutation analysis methods is due to: 1) *NEB* is a large gene, and using dHPLC or DNA sequencing only, the exons and exon-intron boundaries are covered. Large deletions, replications, insertions or inversions are missed, as well as mutations deeper in the introns possibly affecting the splicing of the gene. These kinds of mutations can be detected by RT-PCR, but usually no RNA source, i.e. muscle biopsy or myoblasts, is available. 2) Mutations further upstream in the promoter area or affecting the methylation. Such changes would go unidentified. We have been setting up the multiplex ligation-dependent probe amplification (MLPA) method in order to identify large copy number changes and changes in the methylation pattern of *NEB.* In addition, a project to design microarrays for mutation identification has been launched. Since it is such a gigantic, highly repetitive gene, *NEB* is very challenging to analyse.

1.1 *NEB* mutations in Finnish NM families (I)

To date, NM has been diagnosed in 21 families resident in Finland. In three of these families, however, one of the parents is not of Finnish origin. All parents are healthy which indicates that NM to is inherited in an AR fashion (or the mutations possibly to be *de novo* AD). One of the families has an *ACTA1* mutation (*de novo* AD), another has a likely *TPM3* mutation, and in three families no mutations have yet been identified. The remaining 20 patients from 16 Finnish families have mutations identified in *NEB*; in six families the first mutation and in ten families both of the mutations have been identified. Most of the patients (18) among the

Finnish NM families with identified *NEB* mutations exhibit the typical form of NM. In family 11, one of the siblings has typical NM and another has the intermediate form of NM. In addition, the patient of family 3 has intermediate NM, while none have severe NM. The mutations identified include 15 different mutations (*Table 4*). None of the mutations have been found in homozygous form in any of the Finnish NM patients. Three of the mutations are shared by more than one NM family: a missense mutation (Ser to Ile) in exon 122 in nine families, a missense mutation (Thr to Pro) in exon 151 in three families, and a deletion of ten nucleotides in exon 122 in two families. The remaining mutations identified seem to be unique to each familiy. No two NM families share the same combination of mutations.

Previous histological follow-up studies performed on the biopsies of 13 Finnish NM patients (all included in the *NEB* mutation analyses) detected clear histological changes between the first biopsy and the second biopsy taken 5-18 years later.[222] Deficiency of type II fibres was seen in all of them except for the patients of three families (6, 11 and 13). In fact, the follow-up biopsies of the patients of these families showed hypertrophic type II fibres and small type I fibres.[222] Families 11 and 13 have the same deletion of ten nucleotides in exon 122. The second mutation in family 11 is the previously mentioned missense mutation in exon 151, and in family 13 a splice-site mutation in intron 32. Family 6 has the missense mutation in exon 122 as the first mutation identified, while the second mutation is still unidentified. Thus, all of these families have a mutation in exon 122 but the nature of the mutations differ. It is difficult to draw any conclusions about why these patients have large type II fibres in addition to their small type I fibres while this is not the case in the other families who share the same exon 122 mutations. In addition, three patients from three families (1, 3 and 8) were using wheelchairs at the time the histological follow-up study was performed and showed, logically, small type I fibres and no hypertrophic fibres.[222] They all have the missense mutation in exon 122 and, interestingly, two patients of families 1 and 3 have truncating mutations (different ones) in exon 163. The second mutation present in family 8 is yet to be identified. Patients in two separate non-Finnish families share another truncating mutation in exon 163. One of them has a severe and another has an unknown form of NM.[222]

Comparisons of histological features between patients with *NEB* and *ACTA1* mutations did not reveal any clear differences in their fibre type proportions. Intranuclear rods, however, have to date been reported only in patients with *ACTA1* mutations. As previously mentioned, *ACTA1* mutations more often cause severe NM, while *NEB* mutations seem usually to cause typical NM. The ankle dorsiflexors have been shown to be especially weak in patients with *NEB* mutations while in patients with *ACTA1* mutations, these muscles seem to be relatively spared. *NEB* mutations seem to affect knee extensors more than knee flexors with the situation being the opposite with the *ACTA1* mutations. Otherwise, no clear general differences can be seen.[194]

Table 4. *NEB* mutations identified in Finnish families.

Family	Form of NM; other notes	Mis	Del	Dup	Non	Splice
1	T ; W	Ex122		Ex163		
2	T	Ex122				In54
3	I ; W	Ex122			Ex163	
4	T	Ex122				Int5
5	T	Ex122	Ex68			
6	T ; C	Ex122				
7	T	Ex122				
8	T ; W	Ex122				
9	T	Ex122				
10	T	Ex151		Ex61		
11	I+T ; C	Ex151	Ex122			
12	T	Ex151				
13	T ; C		Ex122			Int32
14	T		Ex61		Ex182	
15	T	Ex76		Ex173		
16	T				Ex142	

Mis = missense mutation, Del = deletion, Dup = duplication, Non = nonsense mutation, Splice = splice-site mutation, T = typical NM, I = intermediate NM, C = CFTD, W = the patient uses a wheelchair. Mutation shared by two or more families highlighted by shaded background. The exact locations of the mutations in the genomic and protein sequences can be found in the corresponding article (I).

1.2 *NEB* mutations in Finnish distal nebulin myopathy families (III)

Analyses of Finnish families revealed a novel, recessively inherited distal myopathy named distal nebulin myopathy (DNM) in six patients from four families, caused by two different homozygous missense mutations in *NEB*. The first mutation is the missense mutation in exon 122 and the second is the missense mutation in 151 (both mentioned above). Combined with more disruptive mutations in compound heterozygotes, the same missense mutations cause NM. The muscle weakness in DNM involves predominantly the neck flexors, the lower leg and the finger extensor muscles (*Fig 15*), with the initial symptom being childhood or adult-onset foot drop. The distribution of affected muscles differs from the weakness seen in NM caused by *NEB* mutations, and also from other inherited distal myopathies, such as tibial muscular dystrophy. As patients with Laing myopathy, however, show a similar distribution of muscle weakness as that of patients with DNM, Laing myopathy should be considered as the diagnosis in dominantly inherited and sporadic cases.[237]
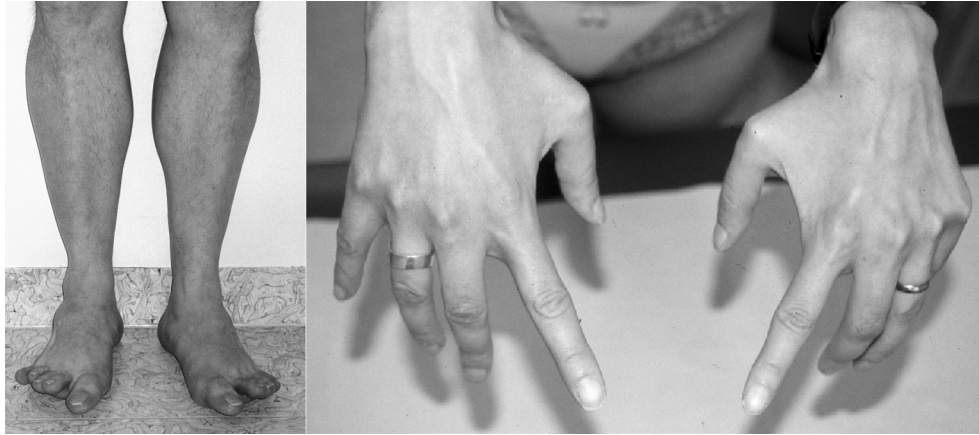
Figure 15. **A.** Atrophy of the tibialis anterior muscles and weakness ankle dorsifexion in patient with homozygous missense mutation (Thr to Pro) in exon 151. **B.** Atrophy of finger extensors of the patient with a homozygous missense mutation (Ser to Pro) in exon 151.

Nemaline bodies (rods) are not detectable in the biopsies of the patients by LM. Some sections viewed under the EM (*Fig 16*) revealed Z-disc-derived aggregates, but they differ from nemaline bodies.
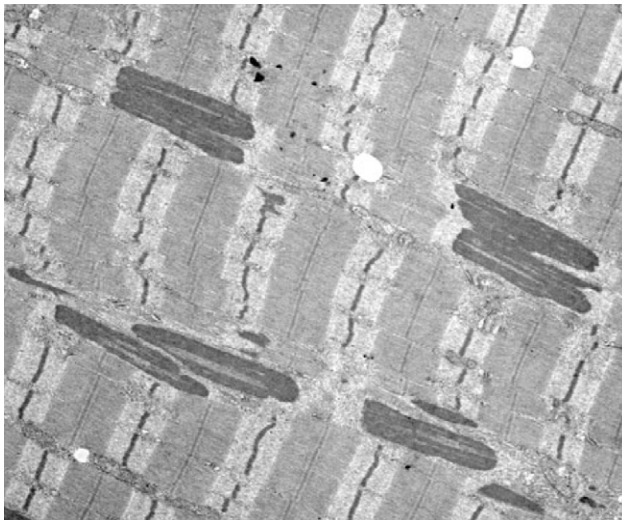

Figure 16. Z-disc streaming visible in the biopsy of a DNM patient

1.3 Genotype-phenotype correlations in patients with *NEB* mutations (I-III)

Based on the results of the present study, it is still difficult to draw reliable conclusions about genotype-phenotype correlations in NM and related disorders. This might be difficult or even impossible until knowledge of all the important domains has been achieved and other proteins binding to nebulin have been identified. Viewing the mutations in *Figure 14 (p. 55)* and in

Tables 3 (p.56). and *4 (p.60)*, it can be seen that all types of mutations are involved in all forms of NM. In addition, mutations underlying different phenotypes have been identified all along the gene and in exons encoding super and simple repeats, as well as in alternatively spliced exons, in exons in the triplicated area, and in unique exons.

Viewing and summarising the mutations causing the **severe form of NM,** it can be said that mutations in exon and intron 81 (the last exon before the triplication of the eight exons in the middle of the gene), and mutations in exon 180 (the last exon encoding the last simple repeat) seem to cause the severe form of NM. Of the 29 mutations causing severe NM, 19 have been found towards the 3' end from exon 81 onwards. Six out of eight patients with the Ashkenazi deletion in homozygous form had severe NM, and two the typical form of NM (II). The **typical form of NM** is the most common form caused by mutations in *NEB,* with 51 mutations identified to date in 42 families. A truncating duplication CAAA in exon 171 caused the **mild form of NM** in one family when in the homozygous form, but combined with a nonsense mutation in exon 162, the same mutation has been found in a patient with the typical form of NM (I). Exon 171 is an alternatively spliced exon, which appears to be expressed in rare isoforms.[151] Another truncating mutation in another family with mild NM is in exon 170 (I), which, too, is in a rarely expressed alternatively spliced exon.[151] On the other hand, a homozygous truncating mutation in the alternatively spliced exon 173 causes the severe form of NM (unpublished). Exon 173 seems to be more commonly expressed than exons 170 and 171.[151]

One of the families classified as having an **unusual or "other" form of NM** is an interesting family with a sib pair affected with ophthalmoplegia, which is a very unusual symptom in NM. The first mutation identified in this sib pair was a small deletion in exon 58 (I), which is the only mutation identified in this exon to date. The second mutation remains to be found. It would be of interest to find out the cause of the eye muscle paralysis and studies towards this are ongoing. It might turn out that some regions in nebulin are essential for the specialised function of the eye muscles. It is also possible that these patients have an additional mutation in another gene. One patient belonging to the category of other forms of NM has a duplication TCAA in exon 171, and a nonsense mutation in exon 119 (I). A patient with core-rod myopathy has the same duplication in exon 171, and another truncating mutation in exon 140 (submitted). The combination of the two mutations thus affects the disease outcome.

Until 2007, mutations identified in *NEB* were known to cause NM, only. The identification of DNM caused by homozygous missense mutations in *NEB* was possible due to the genetic isolation of Finland. This work not only described a novel myopathy, but showed that mutations in nebulin may cause disorders other than NM. Our recent data in fact provide further evidence for this, since we have recently identified mutations causing core-rod

myopathy in one patient. The milder outcome caused by homozygous missense *NEB* mutations in DNM seems logical, if missense mutations are considered to be less disturbing for the function of the protein. However, the overlap between the different NM forms caused by different/the same mutations and the mechanisms of nemaline body or core-rod formation remain to be resolved.

2 Candidate gene and genome-wide linkage analyses (IV & V)

2.1 Candidate gene analysis: Cap myopathy (IV)

One of the aims of this project was to identify underlying genes in families with rare myopathies related to NM using a candidate gene approach. Sequencing genes known to cause similar myopathies, a mutation was identified in the β-tropomyosin (*TPM2)* gene in a patient with cap myopathy. The histological overlap with NM suggested *TPM2* to be the possible causative gene for this disorder, while immunohistochemical analyses performed on the biopsy of the patient displayed strong reactivity of the caps to nebulin. The mutation identified was a heterozygous in-frame deletion of one codon (delGAG) in exon 4 of the β-tropomyosin gene (*TPM2*), removing a single glutamate residue from the β-tropomyosin protein (p.Glu139del). This is expected to disrupt the seven amino-acids long repeat essential for making a coiled coil, and thus to impair the tropomyosin-actin interaction. Sequencing *TPM2* in the remaining six cap myopathy DNA samples in our sample cohort did not lead to the identification of other mutations.

To date, 15 cap myopathy patients in 13 families have been reported (*Table 1, p. 44*). The discovery of families with members affected by both NM and cap myopathy (and some with unspecified muscle disorders)[228, 230, 232] led to a discussion about the overlap between these entities; is cap myopathy a disorder of its own or is cap formation one histological feature of NM? Could the same genetic cause underlie these two disorders? Cap myopathy has previously established a position as an entity of its own. The patient of this study showed clinical and histological overlap both with previously reported cap myopathy patients and with NM patients; the distribution of muscle weakness was similar to that of patients with the typical form of NM, including facial weakness and both proximal and distal weakness of the limbs, as well as respiratory compromise out of proportion to his general muscle weakness. In addition, the MRI scan showed a pattern of thigh muscle involvement similar to that of NM patients with *NEB* mutations. However, neck and knee flexors, as well as rectus femoris, were spared, which is usually not the case with NM. The general body habitus of the patient of the

present study was also different from that seen in the typical form of NM. An unusual feature in this patient was ptosis, which is not a common symptom either in NM nor in previously reported cap myopathy patients.

After the identification of the initial cap myopathy-causing mutation, four additional mutations causing this disorder have been found in *TPM2*.[232, 233]

2.2 Genome-wide linkage analysis in Turkish families (V)

Twelve Turkish families with AR NM were included in a genome-wide linkage analysis using microsatellite markers, and subsequent mutational analyses were performed. As expected, the genome-wide scan revealed no haplotype shared by all the families included in the study. In fact, two different features were seen; those families in which the parents were cousins and originated from isolated populations (such as a village), displayed large homozygous regions covering areas too large to be analysed further. Another group of families did show promising homozygous regions, but subsequent fine mapping with additional markers failed to confirm the homozygosities, revealing false positive linkage in many chromosomal regions. This showed that the initial map of 300 markers (on average the markers were 10 cM apart from each other) was too sparse.

One of the homozygous regions identified did, however, lead to the identification of a possible founder mutation in two of the families in the gene encoding α-tropomyosin$_{slow}$ (*TPM3*), in the chromosomal region 1q12-21.2. Sequencing of this gene revealed a homozygous deletion in the muscle-specific, last exon of *TPM3*, c.913delA. This mutation resulted in severe NM in one and intermediate NM in another family. It was identified neither in the remaining Turkish families nor in the control individuals. The mutation was verified using RT-PCR showing expression of the mutant transcript. Unfortunately, no muscle biopsies, on which protein analysis could be conducted, were available from the patients with the mutation.

The homozygous deletion of an adenine (c.913delA) removes the last nucleotide before the termination codon. A similar, but heterozygous, mutation previously reported to cause NM in a North American patient was analysed using Western blot analysis. This showed that reading through the termination codon leads to a longer protein product.[260] As the mutation previously described,[260] the mutation identified in this project changes the reading frame of the gene leading to read-through of the termination codon and likely to a 73 amino acids longer protein. This longer protein is believed to be unable to form coiled-coil dimers, and subsequently the protein would not be incorporated into the sarcomere, since the

terminal regions of tropomyosins are essential for the proper head-to-tail polymerisation of the protein, but the the mutation might affect the acting-binding properties, as well.[261, 262]

If the extended α-tropomyosin$_{slow}$ is present in the patients' sarcomeres, it should be non-functional, and the situation would be similar to the previously reported patient with a homozygous nonsense mutation Q31X in *TPM3*, and severe NM.[263] In either case, the patients would have no functional α-tropomyosin$_{slow}$ in their type I muscle fibres. This correlates with the histology of the patients' muscle biopsies, showing severely hypotrophic type I muscle fibres. The patients should, however, express functional β-tropomyosin in these fibres. Type II fibres, in which *TPM3* is not expressed at all, are relatively unaffected. The altered protein in these patients is expressed and may affect the function of the two cytoskeletal isoforms of the gene expressed in small quantities also in other tissues,[163] possibly explaining some of the patients' unusual clinical features.

2.3 Mutations in *TPM2* and *TPM3*

The dimerised coiled-coil alpha-helix conformation, as well as polymerisation head-to-tail, are essential for the function of the tropomyosins expressed in the muscle sarcomere. The amino acids in the tropomyosin polymers, arranged into heptad repeats, must be in the exactly correct positions (*a b c d e f g*) for proper head-to-tail binding and for correct dimerisation of the pairing tropomyosins. The positions *g* and *e* of the opposite tropomyosins are bound together by salt bridges and act in stabilising dimer, and they point outwards from the molecule (*Fig 8, p.32*).[157, 158, 261, 262, 264] Mutations in these positions may interrupt the formation of dimers. Other crucial regions, which have the potential to cause a disorder if interrupted, are the actin-binding residues of the tropomyosin in the outer surface acidic residues[265] (such as the region of the E139del) and the tropomyosin-tropomodulin[266] binding sites (*Fig 19, p.70*).

TPM2 is expressed in both slow (type I), and, to a lesser extent, fast (type II) muscle fibres while *TPM3* solely encodes the slow-specific isoform α-tropomyosin$_{slow}$.[127] This might be one of the explanations why *TPM2* mutations are known to cause a variety of different disorders, while mutations in *TPM3* are known to underlie two entities only. Mutations in *TPM3* are known to cause NM and CFTD. The known *TPM2* and *TPM3* mutations and the disorders caused are listed in *Tables 5 (p. 67)* and 6 (p. 68), and shown in *Figures 17 – 19 (p. 67-69*. Mutations in *TPM1* are known to cause cardiomyopathies, but even though the gene is expressed in type II fibres, no skeletal muscle myopathies are known to be caused by *TPM1* mutations.[267-269]

A variety of NMs, (distal) arthrogryposis, CFTD and cap disease are known to be caused by mutations in *TPM2*[205, 217, 218, 270, 271](and personal communication with Prof. Anders Oldfors and Dr. Nigel Clarke) (*Table 5, Fig 17, p. 67*). Except for the situation in one family, the mutations are all dominant (mostly AD *de novo* mutations), and located in the exons expressed in all *TPM2* splice variants. The only recessive mutation in *TPM2* identified to date, a homozygous nonsense mutation causing NM associated with the Escobar syndrome, is in the muscle-specific exon 6B.[271] Escobar syndrome is a rare disorder of orthopaedic and cranial anomalies, e.g. short stature, craniofacial anomalies, joint contractures, skin folds and low-set ears.[272] In family 2 (*Table 5*), a mutation causes NM in one, and cap disease in another family member.[232] Two *TPM2* mutations seem to be recurrent: the mutation E139del first identified in the patient included in this PhD project (IV), has lately been identified as *de novo* AD in two more patients (supported by haplotype results): one with cap myopathy, and another with NM (unpublished data). Another recurrent mutation, R133W,[270] recently identified in another patient, underlies two different forms of arthrogryposes (unpublished data).

Mutations in *TPM3* include both AR and AD mutations, and mutations have been identified both in muscle isoform-specific exons and in exons expressed in isoforms expressed in muscle and other tissues (*Table 6, Fig 18, p. 68*). *TPM3* was the first gene to be identified to cause NM.[198] The patients with *TPM3* mutations include a clinically and histologically variable group of patients, but one common feature shared by the patients are small and abnormal type I fibres. The AD mutations in *TPM3* result in milder forms of NM than AR mutations. The phenotype and the histology does, however, vary between the AR and AD NM cases and even between patients with the same mutation, as is the case with the patients described in this study.[198, 221, 260, 263, 273, 274] The codon encodin to amino acid 168 of α-tropomyosin$_{slow}$ seems to be a mutational hotspot. It has shown to be mutated independently in different individuals six times due to different missense mutations affecting amino acid 168 (unpublished).[221, 273, 274]

Table 5. *TPM2* mutations expressed in β-tropomyosin. All mutations except n:o 13 are dominant, and in exons expressed in all *TPM2* variants.

	Mutation	L	Fam	Disorder
1	p.K7del	*g*	1	NM*
2	p.E41K	*f*	2.1	NM (mother)
			2.2	CAP (daughter)[232]
3	p.K49del	*g*	3	CAP[233]
4	p.G52dup	c	4	CAP[233]
5	p.R91G	*f*	5	(possible) DA1[218]
6	p.E117K	e	6	CFTD[205, 217]
7	p.E122K	*c*		CFTD**
8	p.R133W	g	7	DA2B[270]
			8	Arthrogryposis***
9	p.E139del recurrent	*f*	9	CAP(IV)
			10&11	CAP [i.p.] & NM***
10	p.Q147P	*g*	12	NM [205]
11	p.L148P	*a*	13	NM***
12	p.N202K	*f*	14	CAP[233]
13	p.Q210X homoz, AR	*g*	15	NM/Escobar syndrome[271]

L = location in the coiled-coil.* personal communication with Dr. Nigel Clarke and ** Prof. Anders Oldfors, *** unpublished own data, i.p. = Clarke et al., 2009 in press (Neuromuscul. Disord). Mutational details can be found in the corresponding publications.

```
      abcdefgabcdefgabcdefgabcd...
  1  MDAIKKKMQMLKLDKENAIDRAEQAEADKKQAEDRCKQLEEEQQALQKKLKGTEDEVEKY
 61  SESVKEAQEKLEQAEKKATDAEADVASLNRRIQLVEEELDRAQERLATALQKLEEAEKAA
121  DESERGMKVIENRAMKDEEKMELQEMQLKEAKHIAEDSDRKYEEVARKLVILEGELERSE
181  ERAEVAESKCGDLEEELKIVTNNLKSLEAQADKYSTKEDKYEEEIKLLEEKLKEAETRAE
241  FAERSVAKLEKTIDDLEDEVYAQKMKYKAISEELDNALNDITSL*
```

Figure 17. Alterations caused by mutations identified in β-tropomyosin protein: Every other heptad repeat of the coiled coil in bold and underlined and every other as normal text. Alterations highlighted in grey, recurrent ones highlighted in black, following the order of *Table 5*.

Table 6. *TPM3 mutations expressed in α-tropomyosin$_{slow}$*

	Mutation & mode of inheritance	L	Isoforms affected	Family	Disorder
1	p.M9R AD	a	muscle	1	NM[198]
2	p.Q32X homoz AR	c	muscle	2	NM[263]
3	p.S88F AD ?	c	muscle and nonm.	3	NM*
4	p.L100M AD	a	muscle and nonm.	4	CFTD[221]
5	p.R168H AD	f	muscle and nonm.	5 6 7	NM[273] NM[274] CFTD[221]
6	p.R168C *de novo* AD	f	muscle and nonm.	8 9	CFTD[221] NM*
7	p.R168G *de novo* AD?	f	muscle and nonm.	10	CFTD[221]
8	p.K169E *de novo*AD	g	muscle and nonm.	11	CFTD[221]
9	p.R245G *de novo* AD	f	muscle and nonm.	12	CFTD[221]
10	1.int9 splice mutation 2.X286Next57		muscle	13	NM[260]
11	AR		muscle		
12	Turkish X285NextX73 homoz AR		muscle	14&15	NM (V)

L = location in coil, nonm. = non-muscle, * = unpublished data. Mutational details can be found in the corresponding publications.

```
      abcdefgabcdefgabcdefgaccd...
1    MMEAIKKKMQMLKLDKENALDRAEQAEAEQKQAEERSKQLEDELAAMQKKLKGTEDELDK
61   YSEALKDAQEKLELAEKKAADAEAEVASLNRRIQLVEEELDRAQERLATALQKLEEAEKA
121  ADESERGMKVIENRALKDEEKMELQEIQLKEAKHIAEEADRKYEEVARKLVIIEGDLERT
181  EERAELAESKCSELEEELKNVTNNLKSLEAQAEKYSQKEDKYEEEIKILTDKLKEAETRA
241  EFAERSVAKLEKTIDDLEDELYAQKLKYKAISEELDHALNDMTSI*
```

Figure 18. Alterations caused by mutations identified in α-tropomyosin$_{slow}$ protein: Every other heptad repeat in the coiled coil is underlined and in bold and every other in thin. Alterations highlighted in grey, recurrent ones highlighted in black, following the order of *Table 6*.
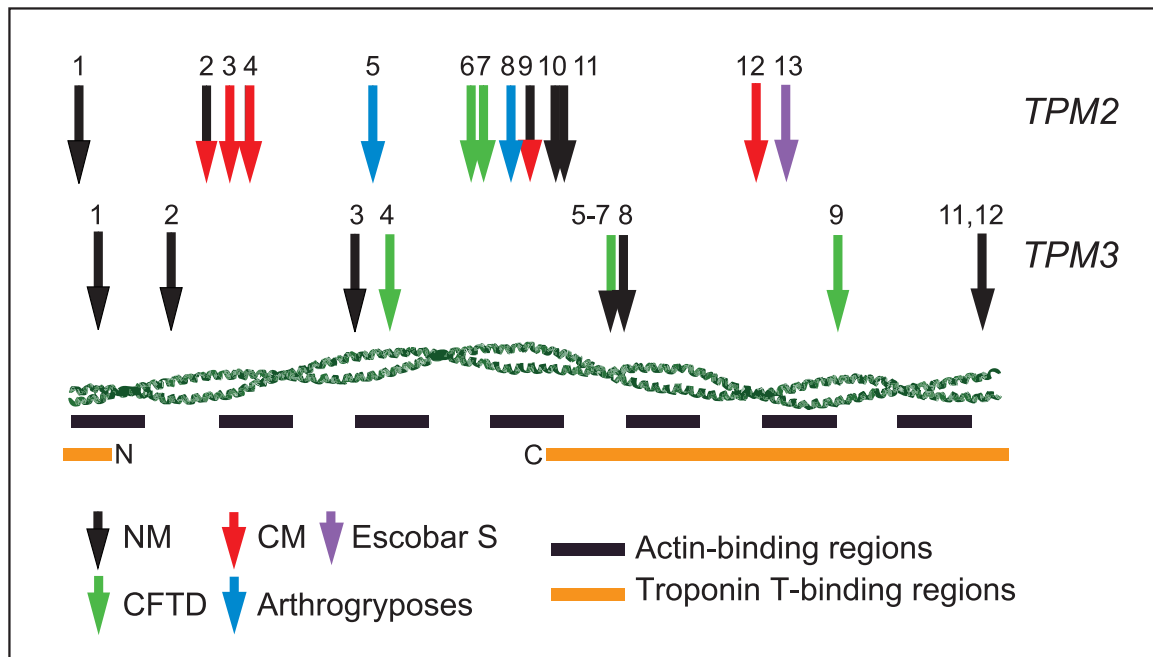
Figure 19. The locations of the alterations identified in the α(slow)- and β-tropomyosins caused by mutations in *TPM3* and *TPM2* respectively. The locations are estimations. The splice-site mutation (*TPM3* n:o 10) is not shown. The alterations are numbered according to *Tables 4* and *5*. (Kindly provided by and modified with permission of Professor Anders Oldfors, University of Göteborg, Sweden)

Conclusions and future prospects

In this PhD study, *NEB*, *TPM2* and *TPM3* mutations were identified in patients with NM, cap myopathy or distal nebulin myopathy - a previously unidentified myopathy.

At the time this project was launched, 18 mutations had been identified in 18 NM families; all had been found in the last 40 of the 183 *NEB* exons using SSCP. The mutation analysis of *NEB* was started at the 3' end of the gene, because the corresponding region of the protein binds to the Z disc, from which nemaline bodies are derived. Previous linkage analyses, and the identification of only one AR mutation in *NEB* in eight families indicated that possibly tens of families within our sample series might have undiscovered mutations in *NEB*. It became obvious that more effective methods were needed in order to analyse all its 183 exons.

Optimisation of the dHPLC method and analysis of samples from 45 NM families revealed that NM-causing mutations could be found along the whole length of the gene, and that the vast majority of the patients have a unique combination of two compound heterozygous mutations. To date, we have analysed 96 NM families and identified 115

different AR mutations in *NEB,* showing that *NEB* mutations truly are the most common cause of AR NM.

In addition to these mutations, deletion of the exon 55 was identified in 2004 in the Ashkenazi population by Anderson's group. During the course of the present study, the occurrence of this mutation was studied and estimated to be 2% among families with NM world-wide. In addition, it was shown that this mutation segregated with a haplotype shared by all the patients carrying the mutation. Although as many as 115 mutations have been identified in *NEB*, this is not enough to discern any reliable correlations between the severity of NM and the type of the *NEB* mutation.

The identification of *NEB* mutations causing a novel myopathy, DNM, and of those causing core-rod myopathy, indicated that mutations in *NEB* might underlie other entities in addition to NM. It may well be that different types of mutations in *NEB* cause different entities. The pathological mechanisms which give rise to these different entities need to be elucidated in future studies. More mutations have to be identified and functional analyses performed to elucidate the possible molecular mechanisms leading to a variety of NM forms and other disorders. Novel mutation analysis methods, such as MLPA and microarrays will be required to detect all types of mutations in *NEB,* since changes in methylation patterns or large alterations such as deletions covering a whole exon, or several exons, cannot be detected using dHPLC.

The previous and ongoing linkage and mutation analyses of known NM genes have show that there should be at least a seventh NM gene yet to be identified. A candidate gene as well as a genome-wide linkage approach was applied in this project in attempts to identify the seventh NM gene and to identify genes underlying those myopathies related to NM which have no genetic cause previously known. The sample cohort of the project included DNA samples from such patients with cap myopathy and core-rod myopathy. A systematic classification of these samples was conducted according to specific, often unusual, histological and/or clinical features and linkage analyses. Sequencing of the genes known to cause similar features in patients with other muscle disorders was then performed (published case reports or personal communications). This led to the identification of the first genetic cause, an AD *de novo* mutation in *TPM2,* underlying cap myopathy.

At the time samples were chosen for the genome-wide linkage analysis, some twenty families showed no linkage to any of the known NM genes. Most of these families originated from Turkey, and therefore all the Turkish samples in our sample cohort were included in this study, which was performed using 300 microsatellite markers (the distance between the two markers being approximately 10 cM). Even the preliminary results of the scan showed large homozygous regions in several chromosomes in some of the families. Analysis of these homozygous regions using more markers, showed that many of the regions had given false

positive results in the genome-wide scan, and further analyses showed some of the regions in fact, to be, heterozygous. No homozygous haplotypes or linkage were seen in most of the families, or the subsequent analyses "broke" the linkage/haplotype. In addition, only two of the families clearly shared the same haplotype in one of the chromosomes, i.e. chromosome 1, and the shared homozygous region in the patients of these two families was the region where *TPM3* is located. Sequencing of *TPM3* revealed a shared homozygous mutation disrupting the termination codon of the muscle-specific isoform of the gene. This is a likely founder mutation within the Turkish population. In order to identify the seventh gene using the samples from the families included in the microsatellite scan, a SNP scan would be the method of choice.

The underlying reasons for the clinical and histological variability in patients with mutations in the same gene or even with exactly the same mutation, remain unclear. Examples are the families with cap and nemaline myopathy patients, the patients with homozygous deletion of the whole nebulin exon 55, or the patients from the two different Turkish families sharing the same homozygous mutation. On the other hand, many of the patients with mutations in different genes (in this context encoding different proteins in the muscle sarcomere) are clinically very similar. This means that examination of the phenotype does not alone permit identification of the causative gene. An example is the case with Laing distal myopathy and nebulin distal myopathy. A significant component of this variability may result from differences in splicing of alternative exons between different individuals, since creation of different isoforms through alternative splicing is extensive in both nebulin and the tropomyosins. Variability in expression of the different tropomyosin genes in different individuals, as well as the effects of other muscle genes, modifier genes and/or normal or abnormal variability in the genome may all influence the patients' outcome. Moreover, even though the basic molecular structure and the functional principles of the skeletal muscle sarcomere have been known for decades, detailed molecular knowledge of both the structure and the events leading to muscle contraction remains to be elucidated.

The exact molecular mechanisms behind the disorders caused by mutations in nebulin and the tropomyosins remain to be elucidated, but it is possible that some of the overlapping clinical features may be explained by shared pathogenetic pathways. Based on the present study, these kinds of mechanisms might cause aberrant interactions of the abnormal proteins with their binding partners within the sarcomere.

Table 7. Summary of the molecular genetic analyses performed and mutations identified during this study.

Gene / Disorder	NEB	TPM2	TPM3
NM Families included (tot) → mutations identified in Mutations	Studies I & II ~ 150 96 115 (AR)	Studies IV & V 45 1 (unpublished) 1	Studies IV & V 12 (+ 59 unpubl) 2 (+ 2 unpubl) 1 (AR) + (2 AD)
DNM Families included (tot) → mutations identified in Mutations	Study III 24 4 2 (AR)	-	-
CAP Families included (tot) → mutations identified in Mutations	-	Study V 7 1 1 (AD)	Study V 7 - -
CRM Families Mutations	Unpublished 1 2 (AR)	-	-
Arthrogryposis Families Mutations	-	Unpublished 1 1 (AD)	-
CFTD Families Mutations		Unpublished 1 1 (AD)	

Identification of the genes and mutations in great number of patients and correlating the clinical and the histological pictures to the mutations identified, is a prerequisite for understanding the pathogenesis of inherited disorders. This is essential for developing specific modes of treatments for these disorders. Mutation identification is often important for the families and patients in order to have a correct diagnosis and makes the prenatal diagnostic possible.

Acknowledgements

vähän pinkkiä ja oranssia!), Paula, Minnamari, Hanna O, Outi, Merja ja Ann-Liz ja kaikki Folkin työkaverit!

Elämäni eläimille, poneille Tambourinille ja Intolle sekä mäyräkoirille Harmille ja Oikulle, kuuluu kiitos siitä, että muutaman kerran (mm yöllä pieneläinklinikalla) olen unohtanut koko väitöskirjan olemassaolon. Lisäksi kiitos lukemattomista vastustuskykyä boostavista kannustuspusuistanne! Kiitän kaikkia eläimeni elämään ja hoitoon osallistuvia ihmisiä eläinlääkäreistä Nelliin ja Marjosta Hennaan ja Helmiin. Erityiskiitos Susalle, maailman ihanimmalle hepoisenhoitajalle. Tallikavereilleni sekä Kilosta että Marjon tallilta kiitos hulvattomasta seurasta ja ystävyydestänne. On ilo viettää aikaa kanssanne tilanteessa kuin tilanteessa.

Kaisa, Nina, Riikka, Laura, Raija-"täti" ja Mikko; kiitos henkireikinäni olemisesta. Uimarpojjiille Ristolle kiitos mutkattomasta, ja Antille uudelleen löytyneestä ystävyydestä.

Haluan lisäksi kiittää kotiväkeäni isää, Iiviä ja Hannaa, sekä sukuani sen massiivisessa laajuudessaan vaarini, pappani, mummoni, rakkaat vanhat vapisevat tätini, setäni, enoni sekä erityisesti serkkuni Annan, Jennin ja Mikin sekä Lehden serkut mukaan lukien. Olette olleet korvaamaton tukiverkko läpi elämäni. Serkkuni Juho ansaitsee hirmusuuren lisäkiitoksen avuliaisuudesta lukuisissa tilanteissa tänä keväänä, ja Venla-serkkua puolestaan kiitän lähes letaaleista tahattomasti aiheutetuista naurukohtauksista sekä mielentilakansioni ahkerasta päivittämisestä.

Äidilleni lähtee kiitos pilven reunalle opinnoissani ja harrastuksissani eteenpäin kannustamisesta sekä mäyräkoirien ymppäämisestä elämääni. Isälle lisäkiitos istuttaneestasi yleisestä kiinnostuksesta maailmanmenoa kohtaan, sekä luonnon ja kaikenkarvaisten/sulkaisten/suomuisten/kuorellisten/nahkaisten hengittävien ja yhteyttävien eläväisten kunnioittamisesta; hakeutuminen luonnontieteiden pariin oli helppo valinta. Lopuksi kiitän Nikoa valtavan suuresta sydämestä, sekä tilannekomiikan- ja huumorintajusta, joka on pelastanut monen monta tilannetta elämässäni.

Helsingissä 27.3.2009   Vilma-Lotta Lehtokari

References

1. **Linnaeus C**. Systema naturae per regna tria naturae: Secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Editio decima, reformata ed. Holmiae (manuscript at Linnean Society, London, United Kingdom): Impensis Direct. Laurentii Salvii; 1758-1759.
2. **Jablokow VR**. Carl von Linne. Can Med Assoc J 1956 Jun 15;74(12):1009-10.
3. **Darwin C**. On the origin of species. London, United Kingdom: John Murray; 1859.
4. **Mendel G.** Experiments in plant hybridization. Read at the February 8th, and March 8th, 1865, meeting sof the Brünn Natural History Society; 1865.
5. **Barnett CF,Jr.** Gregor Johann Mendel - geneticist. New Physician 1964 Feb;13:A88-9.
6. **Sorsby A**. Gregor Mendel. Br Med J 1965 Feb 6;1(5431):333-8.
7. **Gouyon PH, Henry JP, Arnould J.** Gene avatars: The neo-Darwinian theory of evolution. Springer; 2002.
8. **Avery OT**, **MacLeod CM, McCarty** M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J Exp Med 1979 Feb 1;149(2):297-326.
9. **Watson JD, Crick FH.** Genetical implications of the structure of deoxyribonucleic acid. Nature 1953 May 30;171(4361):964-7.
10. **Tijo JH, Levan A.** The chromosome number of man. Hereditas 1956;42:1–6.
11. **Gartler SM.** The chromosome number in humans: A brief history. Nat Rev Genet 2006 Aug;7(8):655-60.
12. **Dave BJ, Sanger WG.** Role of cytogenetics and molecular cytogenetics in the diagnosis of genetic imbalances. Semin Pediatr Neurol 2007 Mar;14(1):2-6.
13. **Sanger F, Coulson AR**. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 1975 May 25;94(3):441-8.
14. **Sanger F, Nicklen S, Coulson AR**. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977 Dec;74(12):5463-7.
15. **Maxam AM, Gilbert W**. A new method for sequencing DNA. Proc Natl Acad Sci U S A 1977 Feb;74(2):560-4.
16. **Mullis KB, Faloona FA**. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol 1987;155:335-50.
17. **Winnacker E.** From genes to clones: Introduction to gene technology. New York, U.S.: VHC; 1987.
18. **Glick B, Pasternak J.** Molecular biotechnology: Principles and applications of recombinant DNA technology. 2nd ed. Washington, U.S.: American Society for Microbiology; 1998.
19. **Bud R**. The uses of life: A history of biotechnology. Cambridge, United Kingdom: Cambridge University Press; 1993.
20. http://www.ncbi.nlm.nih.gov/sites/entrez/ [Internet].
21. **Lander ES**, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton

RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001 Feb 15;409(6822):860-921.

22. **International Human Genome Sequencing Consortium.** Finishing the euchromatic sequence of the human genome. Nature 2004 Oct 21;431(7011):931-45.

23. **Alberts B,** Bray D, Lewis J, Raff M, Roberts K, Watson JD. Molecular biology of the cell. 3rd ed. ed. London, U.K.: Garland Publishing; 1994.

24. **Mayr E.** The growth of biological thought: Diversity, evolution, and inheritance. Cambridge, U.S.: The Belknap Press of Harvard University Press; 1982.

25. **Strachan T, Read A.** Human molecular genetic. 2nd ed. Oxford, UK: BIOS Scientific Publishers Ltd; 1999.

26. **Lander ES, Green P.** Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci U S A 1987 Apr;84(8):2363-7.

27. **Terwilliger JD, Ding Y, Ott J.** On the relative importance of marker heterozygosity and intermarker distance in gene mapping. Genomics 1992 Aug;13(4):951-6.

28. **Terwilliger JD, Ott J.** Handbook of human genetic linkage. Baltimore, U.S.: The John Hopkins University Press; 1994.

29. **International HapMap Consortium.** A haplotype map of the human genome. Nature 2005 Oct 27;437(7063):1299-320.

30. **Matsuzaki H,** Dong S, Loi H, Di X, Liu G, Hubbell E, Law J, Berntsen T, Chadha M, Hui H, Yang G, Kennedy GC, Webster TA, Cawley S, Walsh PS, Jones KW, Fodor SP, Mei R. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods 2004 Nov;1(2):109-11.

31. **International HapMap Consortium**, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt

DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007 Oct 18;449(7164):851-61.

32. **Pennisi E.** Breakthrough of the year. Human genetic variation. Science 2007 Dec 21;318(5858):1842-3.

33. **Chang YH,** Su WH, Lee TC, Sun HF, Chen CH, Pan WH, Tsai SF, Jou YS. TPMD: A database and resources of microsatellite marker genotyped in taiwanese populations. Nucleic Acids Res 2005 Jan 1;33(Database issue):D174-7.

34. **Freeman JL,** Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. Copy number variation: New insights in genome diversity. Genome Res 2006 Aug;16(8):949-61.

35. **Liang Q,** Conte N, Skarnes WC, Bradley A. Extensive genomic copy number variation in embryonic stem cells. Proc Natl Acad Sci U S A 2008.

36. **Pennisi E.** Genomics. DNA study forces rethink of what it means to be a gene. Science 2007 Jun 15;316(5831):1556-7.

37. **Woolfe A,** Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SF, North P, Callaway H, Kelly K, Walter K, Abnizova I, Gilks W, Edwards YJ, Cooke JE, Elgar G. Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol 2005 Jan;3(1):e7.

38. http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml [Internet].

39. **Wheeler DA,** Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM. The complete genome of an individual by massively parallel DNA sequencing. Nature 2008 Apr 17;452(7189):872-6.

40. **Kidd JM,** Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F, Haugen E, Zerr T, Yamada NA, Tsang P, Newman TL, Tuzun E, Cheng Z, Ebling HM, Tusneem N, David R, Gillett W, Phelps KA, Weaver M, Saranga D, Brand A, Tao W, Gustafson E, McKernan K, Chen L, Malig M, Smith JD, Korn JM, McCarroll SA, Altshuler DA, Peiffer DA, Dorschner M, Stamatoyannopoulos J, Schwartz D, Nickerson DA, Mullikin JC, Wilson RK, Bruhn L, Olson MV, Kaul R, Smith DR, Eichler EE. Mapping and sequencing of structural variation from eight human genomes. Nature 2008 May 1;453(7191):56-64.

41. http://www.ncbi.nlm.nih.gov/sites/entrez?db=taxonomy [Internet].
42. www.jax.org [Internet].
43. **Pasternak J.** An introduction to human molecular genetics: Mechanisms of inherited diseases. New Jersey, U.S.: A John Wiley & Sons; 2005.
44. **Happe F, Ronald A.** The 'fractionable autism triad': A review of evidence from behavioural, genetic, cognitive and neural research. Neuropsychol Rev 2008.
45. **Vasudevan** R, Ismail P, Stanslas J, Shamsudin N, Ali AB. Association of insertion/deletion polymorphism of alpha-adrenoceptor gene in essential hypertension with or without type 2 diabetes mellitus in malaysian subjects. Int J Biol Sci 2008;4(6):362-7.
46. **Botstein D, Risch N.** Discovering genotypes underlying human phenotypes: Past successes for mendelian disease, future approaches for complex disease. Nat Genet 2003 Mar;33 Suppl:228-37.
47. **Finsterer J.** Hematological manifestations of primary mitochondrial disorders. Acta Haematol 2007;118(2):88-98.
48. **Verny C,** Amati-Bonneau P, Letournel F, Person B, Dib N, Malinge MC, Slama A, Le Marechal C, Ferec C, Procaccio V, Reynier P, Bonneau D. Mitochondrial DNA A3243G mutation involved in familial diabetes, chronic intestinal pseudo-obstruction and recurrent pancreatitis. Diabetes Metab 2008.
49. **Lee T, Luo L.** Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron 1999 Mar;22(3):451-61.
50. **De Marchi M,** Carbonara AO, Carozzi F, Massara F, Belforte L, Molinatti GM, Bisbocci D, Passarino MP, Palestro G. True hermaphroditism with XX/XY sex chromosome mosaicism: Report of a case. Clin Genet 1976 Nov;10(5):265-72
51. **Hall JG.** Genomic imprinting. Arch Dis Child 1990 Oct;65(10 Spec No):1013-5.
52. **Hall JG.** Genomic imprinting: Review and relevance to human diseases. Am J Hum Genet 1990 May;46(5):857-73.
53. **Roizen NJ, Patterson D.** Down's syndrome. Lancet 2003 Apr 12;361(9365):1281-9.
54. **Hossjer O.** Modeling the effect of inbreeding among founders in linkage analysis. Theor Popul Biol 2006 Sep;70(2):146-63.
55. **Gasbarra D,** Pirinen M, Sillanpaa MJ, Arjas E. Estimating genealogies from linked marker data: A bayesian approach. BMC Bioinformatics 2007 Oct 25;8:411.
56. **Wijsman EM.** A deductive method of haplotype analysis in pedigrees. Am J Hum Genet 1987 Sep;41(3):356-73.
57. **Sobel E, Lange K.** Descent graphs in pedigree analysis: Applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996 Jun;58(6):1323-37.
58. **Gao G,** Hoeschele I, Sorensen P, Du F. Conditional probability methods for haplotyping in pedigrees. Genetics 2004 Aug;167(4):2055-65.
59. **John S,** Shephard N, Liu G, Zeggini E, Cao M, Chen W, Vasavda N, Mills T, Barton A, Hinks A, Eyre S, Jones KW, Ollier W, Silman A, Gibson N, Worthington J, Kennedy GC. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: Comparison with microsatellites. Am J Hum Genet 2004 Jul;75(1):54-64.
60. **Siintola E,** Topcu M, Aula N, Lohi H, Minassian BA, Paterson AD, Liu XQ, Wilson C, Lahtinen U, Anttonen AK, Lehesjoki AE. The novel neuronal ceroid lipofuscinosis gene MFSD8 encodes a putative lysosomal transporter. Am J Hum Genet 2007 Jul;81(1):136-46.
61. **Lathrop GM, Lalouel JM.** Efficiency of recombination estimates using two-and three-point linkage data. Prog Clin Biol Res 1985;194:97-102.
62. **Kruglyak L,** Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: A unified multipoint approach. Am J Hum Genet 1996 Jun;58(6):1347-63.
63. **Sullivan PF,** Neale BM, Neale MC, van den Oord E, Kendler KS. Multipoint and single point non-parametric linkage analysis with imperfect data. Am J Med Genet B Neuropsychiatr Genet 2003 Aug 15;121B(1):89-94.

64. **Hirschhorn JN, Daly MJ.** Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005 Feb;6(2):95-108.

65. **Lathrop GM, Lalouel JM**. Estimation of recombination and genetic risks using several markers. Prog Clin Biol Res 1984;147:267-9.

66. **Lathrop GM,** Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 1984 Jun;81(11):3443-6.

67. **Kruglyak L**. Thresholds and sample sizes. Nat Genet 1996 Oct;14(2):132-3.

68. **Abecasis GR,** Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002 Jan;30(1):97-101.

69. **Smithies O.** Animal models of human genetic diseases. Trends Genet 1993 Apr;9(4):112-6.

70. **Nguyen MA, Hardeman EC**. Mouse models for thin filament disease. Adv Exp Med Biol 2008;642:66-77.

71. **McKeown T.** The origins of human disease. Oxford, United Kingdom: Basil Blackwell; 1988.

72. **ten Kate LP**. Completeness of catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes. Am J Med Genet 1992 Jun 1;43(3):606-8.

73. **McKusick VA.** Mendelian inheritance in man: Catalogs of autosomal dominant, autosomal recessive, and X linked phenotypes. 10th ed ed. Baltimore, U.S.: Johns Hopkins University Press; 1992:xxi.

74. **Wilkie AO.** The molecular basis of genetic dominance. J Med Genet 1994 Feb;31(2):89-98.

75. **Cartegni L, Chew SL, Krainer AR**. Listening to silence and understanding nonsense: Exonic mutations that affect splicing. Nat Rev Genet 2002 Apr;3(4):285-98.

76. **Goren A**, Ram O, Amit M, Keren H, Lev-Maor G, Vig I, Pupko T, Ast G. Comparative analysis identifies exonic splicing regulatory sequences--the complex definition of enhancers and silencers. Mol Cell 2006 Jun 23;22(6):769-81.

77. **Dietz HC,** Valle D, Francomano CA, Kendzior RJ,Jr, Pyeritz RE, Cutting GR. The skipping of constitutive exons in vivo induced by nonsense mutations. Science 1993 Jan 29;259(5095):680-3.

78. **McIntosh I, Hamosh A, Dietz HC.** Nonsense mutations and diminished mRNA levels. Nat Genet 1993 Jul;4(3):219.

79. **Miriami E, Margalit H, Sperling R.** Conserved sequence elements associated with exon skipping. Nucleic Acids Res 2003 Apr 1;31(7):1974-83.

80. **Sheffield VC**, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. Genomics 1993 May;16(2):325-32.

81. **Jones AC,** Austin J, Hansen N, Hoogendoorn B, Oefner PJ, Cheadle JP, O'Donovan MC. Optimal temperature selection for mutation detection by denaturing HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis. Clin Chem 1999 Aug;45(8 Pt 1):1133-40.

82. **Oefner PJ, Huber CG.** A decade of high-resolution liquid chromatography of nucleic acids on styrene-divinylbenzene copolymers. J Chromatogr B Analyt Technol Biomed Life Sci 2002 Dec 25;782(1-2):27-55.

83. **Cremonesi L,** Stenirri S, Fermo I, Paroni R, Ferrari M, Cazzola M, Arosio P. Denaturing HPLC analysis of DNA deletions and insertions. Hum Mutat 2003 Jul;22(1):98-102.

84. **Kosaki K, Udaka T, Okuyama T.** DHPLC in clinical molecular diagnostic services. Mol Genet Metab 2005 Sep-Oct;86(1-2):117-23.

85. **Ballana E,** Govea N, de Cid R, Garcia C, Arribas C, Rosell J, Estivill X. Detection of unrecognized low-level mtDNA heteroplasmy may explain the variable phenotypic expressivity of apparently homoplasmic mtDNA mutations. Hum Mutat 2008 Feb;29(2):248-57.

86. **Southern EM**. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 1975 Nov 5;98(3):503-17.

87. **Southern E.** Southern blotting. Nat Protoc 2006;1(2):518-25.

88. **Schouten JP,** McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 2002 Jun 15;30(12):e57.

89. **Nygren AO,** Ameziane N, Duarte HM, Vijzelaar RN, Waisfisz Q, Hess CJ, Schouten JP, Errami A. Methylation-specific MLPA (MS-MLPA): Simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res 2005 Aug 16;33(14):e128.

90. **Tang J, Xiao P.** Polymerizing immobilization of acrylamide-modified nucleic acids and its application. Biosens Bioelectron 2008

91. **Barrett MT,** Scheffer A, Ben-Dor A, Sampas N, Lipson D, Kincaid R, Tsang P, Curry B, Baird K, Meltzer PS, Yakhini Z, Bruhn L, Laderman S. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. Proc Natl Acad Sci U S A 2004 Dec 21;101(51):17765-70.

92. **Hegde MR,** Chin EL, Mulle JG, Okou DT, Warren ST, Zwick ME. Microarray-based mutation detection in the dystrophin gene. Hum Mutat 2008 Sep;29(9):1091-9.

93. http://blast.ncbi.nlm.nih.gov/Blast.cgi [Internet].

94. **Heid CA,** Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. Genome Res 1996 Oct;6(10):986-94.

95. **Geller B,** Badner JA, Tillman R, Christian SL, Bolhofner K, Cook EH,Jr. Linkage disequilibrium of the brain-derived neurotrophic factor Val66Met polymorphism in children with a prepubertal and early adolescent bipolar disorder phenotype. Am J Psychiatry 2004 Sep;161(9):1698-700.

96. **Martin MW,** Grazhdankin DV, Bowring SA, Evans DA, Fedonkin MA, Kirschvink JL. Age of neoproterozoic bilatarian body and trace fossils, White Sea, Russia: Implications for metazoan evolution. Science 2000 May 5;288(5467):841-5.

97. **OOta S, Saitou N.** Phylogenetic relationship of muscle tissues deduced from superimposition of gene trees. Mol Biol Evol 1999 Jun;16(6):856-67.

98. **Stone R, Stone J.** Atlas of skeletal muscles. McGraw Hill book Co; 2006.

99. **Dubowitz V, Sewry C,** editors. Muscle biopsy - A practical approach 3rd edition ed. Elsevier Health Sciences; 2007.

100. **Spangenburg EE, Booth FW.** Molecular regulation of individual skeletal muscle fibre types. Acta Physiol Scand 2003 Aug;178(4):413-24.

101. **Campanaro S,** Romualdi C, Fanin M, Celegato B, Pacchioni B, Trevisan S, Laveder P, De Pitta C, Pegoraro E, Hayashi YK, Valle G, Angelini C, Lanfranchi G. Gene expression profiling in dysferlinopathies using a dedicated muscle microarray. Hum Mol Genet 2002 Dec 15;11(26):3283-98.

102. **Chen YW,** Zhao P, Borup R, Hoffman EP. Expression profiling in the muscular dystrophies: Identification of novel aspects of molecular pathophysiology. J Cell Biol 2000 Dec 11;151(6):1321-36.

103. **Lanfranchi G,** Muraro T, Caldara F, Pacchioni B, Pallavicini A, Pandolfo D, Toppo S, Trevisan S, Scarso S, Valle G. Identification of 4370 expressed sequence tags from a 3'-end-specific cDNA library of human skeletal muscle by DNA sequencing and filter hybridization. Genome Res 1996 Jan;6(1):35-42.

104. http://www.ncbi.nlm.nih.gov/geo/ [Internet].

105. **Pietu G,** Eveno E, Soury-Segurens B, Fayein NA, Mariage-Samson R, Matingou C, Leroy E, Dechesne C, Krieger S, Ansorge W, Reguigne-Arnould I, Cox D, Dehejia A, Polymeropoulos MH, Devignes MD, Auffray C. The genexpress IMAGE knowledge base of the human muscle transcriptome: A resource of structural, functional, and positional candidate genes for muscle physiology and pathologies. Genome Res 1999 Dec;9(12):1313-20.

106. **Huxley AF, Niedergerke R.** Structural changes in muscle during contraction; interference microscopy of living muscle fibres. Nature 1954 May 22;173(4412):971-3.

107. **Spiro D.** The ultrastructure of striated muscle at various sarcomere lengths. J Biophys Biochem Cytol 1956 Jul 25;2(4, Suppl):157-62.

108. **Lange S, Ehler E, Gautel M.** From A to Z and back? multicompartment proteins in the sarcomere. Trends Cell Biol 2006 Jan;16(1):11-8.

109. **Squire JM.** Architecture and function in the muscle sarcomere. Curr Opin Struct Biol 1997 Apr;7(2):247-57.

110. **Luther PK,** Padron R, Ritter S, Craig R, Squire JM. Heterogeneity of Z-band structure within a single muscle sarcomere: Implications for sarcomere assembly. J Mol Biol 2003 Sep 5;332(1):161-9.

111. **Millevoi S,** Trombitas K, Kolmerer B, Kostin S, Schaper J, Pelin K, Granzier H, Labeit S. Characterization of nebulette and nebulin and emerging concepts of their roles for vertebrate Z-discs. J Mol Biol 1998 Sep 11;282(1):111-23.

112. **Witt CC,** Burkart C, Labeit D, McNabb M, Wu Y, Granzier H, Labeit S. Nebulin regulates thin filament length, contractility, and Z-disk structure in vivo. EMBO J 2006 Aug 23;25(16):3843-55.

113. **Zhou Q,** Ruiz-Lozano P, Martone ME, Chen J. Cypher, a striated muscle-restricted PDZ and LIM domain-containing protein, binds to alpha-actinin-2 and protein kinase C. J Biol Chem 1999 Jul 9;274(28):19807-13.

114. **Otey CA,** Rachlin A, Moza M, Arneman D, Carpen O. The palladin/myotilin/myopalladin family of actin-associated scaffolds. Int Rev Cytol 2005;246:31-58.

115. **Hance JE,** Fu SY, Watkins SC, Beggs AH, Michalak M. Alpha-actinin-2 is a new component of the dystrophin-glycoprotein complex. Arch Biochem Biophys 1999 May 15;365(2):216-22.

116. **Sorimachi H,** Freiburg A, Kolmerer B, Ishiura S, Stier G, Gregorio CC, Labeit D, Linke WA, Suzuki K, Labeit S. Tissue-specific expression and alpha-actinin binding properties of the Z-disc titin: Implications for the nature of vertebrate Z-discs. J Mol Biol 1997 Aug 1;270(5):688-95.

117. **Pavalko FM,** Otey CA, Simon KO, Burridge K. Alpha-actinin: A direct link between actin and integrins. Biochem Soc Trans 1991 Nov;19(4):1065-9.

118. **Kontrogianni-Konstantopoulos A,** Catino DH, Strong JC, Bloch RJ. De novo myofibrillogenesis in C2C12 cells: Evidence for the independent assembly of M bands and Z disks. Am J Physiol Cell Physiol 2006 Feb;290(2):C626-37.

119. **Luther PK.** Three-dimensional structure of a vertebrate muscle Z-band: Implications for titin and alpha-actinin binding. J Struct Biol 2000 Feb;129(1):1-16.

120. **Gregorio CC,** Trombitas K, Centner T, Kolmerer B, Stier G, Kunke K, Suzuki K, Obermayr F, Herrmann B, Granzier H, Sorimachi H, Labeit S. The NH2 terminus of titin spans the Z-disc: Its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity. J Cell Biol 1998 Nov 16;143(4):1013-27.

121. **Politou AS**, Spadaccini R, Joseph C, Brannetti B, Guerrini R, Helmer-Citterich M, Salvadori S, Temussi PA, Pastore A. The SH3 domain of nebulin binds selectively to type II peptides: Theoretical prediction and experimental validation. J Mol Biol 2002 Feb 15;316(2):305-15.

122. **Wang K,** Knipfer M, Huang QQ, van Heerden A, Hsu LC, Gutierrez G, Quian XL, Stedman H. Human skeletal muscle nebulin sequence encodes a blueprint for thin filament architecture. sequence motifs and affinity profiles of tandem repeats and terminal SH3. J Biol Chem 1996 Feb 23;271(8):4304-14.

123. **Horowits R.** Nebulin regulation of actin filament lengths: New angles. Trends Cell Biol 2006 Mar;16(3):121-4.

124. **Vandekerckhove J,** Weber K. At least six different actins are expressed in a higher mammal: An analysis based on the amino acid sequence of the amino-terminal tryptic peptide. J Mol Biol 1978 Dec 25;126(4):783-802.

125. **Gordon AM, Homsher E, Regnier M**. Regulation of contraction in striated muscle. Physiol Rev 2000 Apr;80(2):853-924.

126. **Lukoyanova N,** VanLoock MS, Orlova A, Galkin VE, Wang K, Egelman EH. Each actin subunit has three nebulin binding sites: Implications for steric blocking. Curr Biol 2002 Mar 5;12(5):383-8.

127. **Perry SV.** Vertebrate tropomyosin: Distribution, properties and function. J Muscle Res Cell Motil 2001;22(1):5-49.

128. **Kruger M, Wright J, Wang K.** Nebulin as a length regulator of thin filaments of vertebrate skeletal muscles: Correlation of thin filament length, nebulin size, and epitope profile. J Cell Biol 1991 Oct;115(1):97-107.

129. **Labeit S,** Gibson T, Lakey A, Leonard K, Zeviani M, Knight P, Wardale J, Trinick J. Evidence that nebulin is a protein-ruler in muscle thin filaments. FEBS Lett 1991 May 6;282(2):313-6.

130. **McElhinny AS,** Schwach C, Valichnac M, Mount-Patrick S, Gregorio CC. Nebulin regulates the assembly and lengths of the thin filaments in striated muscle. J Cell Biol 2005 Sep 12;170(6):947-57.

131. **Ottenheijm CA**, Fong C, Vangheluwe P, Wuytack F, Babu GJ, Periasamy M, Witt CC, Labeit S, Granzier H. Sarcoplasmic reticulum calcium uptake and speed of relaxation are depressed in nebulin-free skeletal muscle. FASEB J 2008 Aug;22(8):2912-9.

132. **Pfuhl M,** Winder SJ, Castiglione Morelli MA, Labeit S, Pastore A. Correlation between conformational and binding properties of nebulin repeats. J Mol Biol 1996 Mar 29;257(2):367-84.

133. **Wang K, Wright J.** Architecture of the sarcomere matrix of skeletal muscle: Immunoelectron microscopic evidence that suggests a set of parallel inextensible nebulin filaments anchored at the Z line. J Cell Biol 1988 Dec;107(6 Pt 1):2199-212.

134. **Schafer DA, Cooper JA.** Control of actin assembly at filament ends. Annu Rev Cell Dev Biol 1995;11:497-518.

135. **Politou AS,** Millevoi S, Gautel M, Kolmerer B, Pastore A. SH3 in muscles: Solution structure of the SH3 domain from nebulin. J Mol Biol 1998 Feb 13;276(1):189-202.

136. **Bang ML,** Mudry RE, McElhinny AS, Trombitas K, Geach AJ, Yamasaki R, Sorimachi H, Granzier H, Gregorio CC, Labeit S. Myopalladin, a novel 145-kilodalton sarcomeric protein with multiple roles in Z-disc and I-band protein assemblies. J Cell Biol 2001 Apr 16;153(2):413-27.

137. **Ma K, Wang K.** Interaction of nebulin SH3 domain with titin PEVK and myopalladin: Implications for the signaling and assembly role of titin and nebulin. FEBS Lett 2002 Dec 18;532(3):273-8.

138. **Pappas CT,** Bhattacharya N, Cooper JA, Gregorio CC. Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc. Mol Biol Cell 2008 May;19(5):1837-47.

139. **Bang ML, Gregorio C, Labeit S.** Molecular dissection of the interaction of desmin with the C-terminal region of nebulin. J Struct Biol 2002 Jan-Feb;137(1-2):119-27.

140. **Costa ML,** Escaleira R, Cataldo A, Oliveira F, Mermelstein CS. Desmin: Molecular interactions and putative functions of the muscle intermediate filament protein. Braz J Med Biol Res 2004 Dec;37(12):1819-30.

141. **Moncman CL, Wang K.** Architecture of the thin filament-Z-line junction: Lessons from nebulette and nebulin homologies. J Muscle Res Cell Motil 2000 Feb;21(2):153-69.

142. **Labeit S, Kolmerer B.** The complete primary structure of human nebulin and its correlation to muscle structure. J Mol Biol 1995 Apr 28;248(2):308-15.

143. **Root DD, Wang K.** Calmodulin-sensitive interaction of human nebulin fragments with actin and myosin. Biochemistry 1994 Oct 25;33(42):12581-91.

144. **Gregorio CC,** Weber A, Bondad M, Pennise CR, Fowler VM. Requirement of pointed-end capping by tropomodulin to maintain actin filament length in embryonic chick cardiac myocytes. Nature 1995 Sep 7;377(6544):83-6.

145. **Fowler VM.** Capping actin filament growth: Tropomodulin in muscle and nonmuscle cells. Soc Gen Physiol Ser 1997;52:79-89.

146. **McElhinny AS,** Kolmerer B, Fowler VM, Labeit S, Gregorio CC. The N-terminal end of nebulin interacts with tropomodulin at the pointed ends of the thin filaments. J Biol Chem 2001 Jan 5;276(1):583-92.

147. **Bang ML,** Li X, Littlefield R, Bremner S, Thor A, Knowlton KU, Lieber RL, Chen J. Nebulin-deficient mice exhibit shorter thin filament lengths and reduced contractile function in skeletal muscle. J Cell Biol 2006 Jun 19;173(6):905-16.

148. **Wallgren-Pettersson C,** Donner K, Sewry C, Bijlsma E, Lammens M, Bushby K, Giovannucci Uzielli ML, Lapi E, Odent S, Akcoren Z, Topaloglu H, Pelin K. Mutations in the nebulin gene can cause severe congenital nemaline myopathy. Neuromuscul Disord 2002 Oct;12(7-8):674-9.

149. **Pelin K,** Ridanpaa M, Donner K, Wilton S, Krishnarajah J, Laing N, Kolmerer B, Millevoi S, Labeit S, de la Chapelle A, Wallgren-Petterson C. Refined localisation of the genes for nebulin and titin on chromosome 2q allows the assignment of nebulin as a candidate gene for autosomal recessive nemaline myopathy. Eur J Hum Genet 1997 Jul-Aug;5(4):229-34.

150. **Pelin K, Wallgren-Pettersson C.** Nebulin - A giant chameleon. In: Nigel G. Laing, editor. The sarcomere and skeletal muscle disease. Landes Bioscience; 2008. ID: 258.

151. **Donner K,** Sandbacka M, Lehtokari VL, Wallgren-Pettersson C, Pelin K. Complete genomic structure of the human nebulin gene and identification of alternatively spliced transcripts. Eur J Hum Genet 2004 Sep;12(9):744-51.

152. **Kazmierski ST,** Antin PB, Witt CC, Huebner N, McElhinny AS, Labeit S, Gregorio CC. The complete mouse nebulin gene sequence and the identification of cardiac nebulin. J Mol Biol 2003 May 9;328(4):835-46.

153. **McElhinny AS,** Kazmierski ST, Labeit S, Gregorio CC. Nebulin: The nebulous, multifunctional giant of striated muscle. Trends Cardiovasc Med 2003 Jul;13(5):195-201.

154. **Wang K, Williamson CL**. Identification of an N2 line protein of striated muscle. Proc Natl Acad Sci U S A 1980 Jun;77(6):3254-8.

155. **Joo YM,** Lee MA, Lee YM, Kim MS, Kim SY, Jeon EH, Choi JK, Kim WH, Lee HC, Min BI, Kang HS, Kim CR. Identification of chicken nebulin isoforms of the 31-residue motifs and non-muscle nebulin. Biochem Biophys Res Commun 2004 Dec 24;325(4):1286-91.

156. **Ryadnov MG.** Peptide alpha-helices for synthetic nanostructures. Biochem Soc Trans 2007 Jun;35(Pt 3):487-91.

157. **Caspar DL, Cohen C, Longley W.** Tropomyosin: Crystal structure, polymorphism and molecular interactions. J Mol Biol 1969 Apr 14;41(1):87-107.

158. **Phillips GN,Jr,** Lattman EE, Cummins P, Lee KY, Cohen C. Crystal structure and molecular interactions of tropomyosin. Nature 1979 Mar 29;278(5703):413-7.

159. **Matsumura F, Yamashiro-Matsumura S, Lin JJ. I**solation and characterization of tropomyosin-containing microfilaments from cultured cells. J Biol Chem 1983 May 25;258(10):6636-44.

160. **Holmes KC,** Popp D, Gebhard W, Kabsch W. Atomic model of the actin filament. Nature 1990 Sep 6;347(6288):44-9.

161. **Cooper JA.** Actin dynamics: Tropomyosin provides stability. Curr Biol 2002 Aug 6;12(15):R523-5.

162. **Tiso N,** Rampoldi L, Pallavicini A, Zimbello R, Pandolfo D, Valle G, Lanfranchi G, Danieli GA. Fine mapping of five human skeletal muscle genes: Alpha-tropomyosin, beta-tropomyosin, troponin-I slow-twitch, troponin-I fast-twitch, and troponin-C fast. Biochem Biophys Res Commun 1997 Jan 13;230(2):347-50.

163. **Pittenger MF, Kazzaz JA, Helfman DM.** Functional properties of non-muscle tropomyosin isoforms. Curr Opin Cell Biol 1994 Feb;6(1):96-104.

164. **Dufour C**, Weinberger RP, Schevzov G, Jeffrey PL, Gunning P. Splicing of two internal and four carboxyl-terminal alternative exons in nonmuscle tropomyosin 5 pre-mRNA is independently regulated during development. J Biol Chem 1998 Jul 17;273(29):18547-55.

165. **Cooley BC, Bergtrom G.** Multiple combinations of alternatively spliced exons in rat tropomyosin-alpha gene mRNA: Evidence for 20 new isoforms in adult tissues and cultured cells. Arch Biochem Biophys 2001 Jun 1;390(1):71-7.

166. **Thirion C,** Stucka R, Mendel B, Gruhler A, Jaksch M, Nowak KJ, Binz N, Laing NG, Lochmuller H. Characterization of human muscle type cofilin (CFL2) in normal and regenerating muscle. Eur J Biochem 2001 Jun;268(12):3473-82.

167. **Schultheiss T,** Lin ZX, Lu MH, Murray J, Fischman DA, Weber K, Masaki T, Imamura M, Holtzer H. Differential distribution of subsets of myofibrillar proteins in cardiac nonstriated and striated myofibrils. J Cell Biol 1990 Apr;110(4):1159-72.

168. **Berg JS, Powell BC, Cheney RE.** A millennial myosin census. Mol Biol Cell 2001 Apr;12(4):780-94.

169. **McElhinny AS,** Kakinuma K, Sorimachi H, Labeit S, Gregorio CC. Muscle-specific RING finger-1 interacts with titin to regulate sarcomeric M-line and thick filament structure and may have nuclear functions via its interaction with glucocorticoid modulatory element binding protein-1. J Cell Biol 2002 Apr 1;157(1):125-36.

170. **Pinotsis N,** Lange S, Perriard JC, Svergun DI, Wilmanns M. Molecular basis of the C-terminal tail-to-tail assembly of the sarcomeric filament protein myomesin. EMBO J 2008 Jan 9;27(1):253-64.

171. **Wang K.** Titin/connectin and nebulin: Giant protein rulers of muscle structure and function. Adv Biophys 1996;33:123-34.

172. **Labeit S, Kolmerer B, Linke WA**. The giant protein titin. emerging roles in physiology and pathophysiology. Circ Res 1997 Feb;80(2):290-4.

173. **Granzier HL, Labeit S.** Titin and its associated proteins: The third myofilament system of the sarcomere. Adv Protein Chem 2005;71:89-119.

174. **Granzier HL, Labeit S.** The giant muscle protein titin is an adjustable molecular spring. Exerc Sport Sci Rev 2006 Apr;34(2):50-3.

175. **Zou P,** Pinotsis N, Lange S, Song YH, Popov A, Mavridis I, Mayans OM, Gautel M, Wilmanns M. Palindromic assembly of the giant muscle protein titin in the sarcomeric Z-disk. Nature 2006 Jan 12;439(7073):229-33.

176. **Knupp C, Luther PK, Squire JM.** Titin organisation and the 3D architecture of the vertebrate-striated muscle I-band. J Mol Biol 2002 Sep 27;322(4):731-9.

177. **Lin Z,** Lu MH, Schultheiss T, Choi J, Holtzer S, DiLullo C, Fischman DA, Holtzer H. Sequential appearance of muscle-specific proteins in myoblasts as a function of time after cell division: Evidence for a conserved myoblast differentiation program in skeletal muscle. Cell Motil Cytoskeleton 1994;29(1):1-19.

178. **Lange S,** Xiang F, Yakovenko A, Vihola A, Hackman P, Rostkova E, Kristensen J, Brandmeier B, Franzen G, Hedberg B, Gunnarsson LG, Hughes SM, Marchand S, Sejersen T, Richard I, Edstrom L, Ehler E, Udd B, Gautel M. The kinase domain of titin controls muscle gene expression and protein turnover. Science 2005 Jun 10;308(5728):1599-603.

179. **Weinert S,** Bergmann N, Luo X, Erdmann B, Gotthardt M. M line-deficient titin causes cardiac lethality through impaired maturation of the sarcomere. J Cell Biol 2006 **May 22;173(4):559-70.**

180. **Beckmann JS, Spencer M.** Calpain 3, the "gatekeeper" of proper sarcomere assembly, turnover and maintenance. Neuromuscul Disord 2008 Dec;18(12):913-21.

181. **Granzier H,** Helmes M, Cazorla O, McNabb M, Labeit D, Wu Y, Yamasaki R, Redkar A, Kellermayer M, Labeit S, Trombitas K. Mechanical properties of titin isoforms. Adv Exp Med Biol 2000;481:283,300; discussion 300-4.

182. **Bang ML,** Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H, Labeit S. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. Circ Res 2001 Nov 23;89(11):1065-72.

183. **Freiburg A,** Trombitas K, Hell W, Cazorla O, Fougerousse F, Centner T, Kolmerer B, Witt C, Beckmann JS, Gregorio CC, Granzier H, Labeit S. Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. Circ Res 2000 Jun 9;86(11):1114-21.

184. **Behr T,** Fischer P, Muller-Felber W, Schmidt-Achert M, Pongratz D. Myofibrillogenesis in primary tissue cultures of adult human skeletal muscle: Expression of desmin, titin, and nebulin. Clin Investig 1994 Jan;72(2):150-5.

185. **Gordon AM, Regnier M, Homsher E.** Skeletal and cardiac muscle contractile activation: Tropomyosin "rocks and rolls". News Physiol Sci 2001 Apr;16:49-55.

186. **Shah SB, Lieber RL.** Simultaneous imaging and functional assessment of cytoskeletal protein connections in passively loaded single muscle cells. J Histochem Cytochem 2003 Jan;51(1):19-29.

187. **Shy GM,** Engel WK, Somers JE, Wanko T. Nemaline Myopathy. A New Congenital Myopathy. Brain 1963 Dec;86:793-810.

188. **Conen PE, Murphy EG, Donohue WI.** Light And Electron Microscopic Studies Of "Myogranules" In A Child With Hypotonia And Muscle Weakness. Can Med Assoc J 1963 Nov 9;89:983-6.

189. **Fidzianska A,** Badurska B, Ryniewicz B, Dembek I. "Cap disease": New congenital myopathy. Neurology 1981 Sep;31(9):1113-20.

190. **Magee KR, Shy GM.** A new congenital non-progressive myopathy. Brain 1956 Dec;79(4):610-21.

191. **Laing NG.** The sarcomere and skeletal muscle disease. Landes Bioscence Books; 2008.

192. **Schnell C,** Kan A, North KN. 'An artefact gone awry': Identification of the first case of nemaline myopathy by dr R.D.K. reye. Neuromuscul Disord 2000 Jun;10(4-5):307-12.

193. **Wallgren-Pettersson C,** Pelin K, Hilpela P, Donner K, Porfirio B, Graziano C, Swoboda KJ, Fardeau M, Urtizberea JA, Muntoni F, Sewry C, Dubowitz V, Iannaccone S, Minetti C, Pedemonte M, Seri M, Cusano R, Lammens M, Castagna-Sloane A, Beggs AH, Laing NG, de la Chapelle A. Clinical and genetic heterogeneity in autosomal recessive nemaline myopathy. Neuromuscul Disord 1999 Dec;9(8):564-72.

194. **Wallgren-Pettersson C,** Pelin K, Nowak KJ, Muntoni F, Romero NB, Goebel HH, North KN, Beggs AH, Laing NG, ENMC International Consortium On Nemaline Myopathy. Genotype-phenotype correlations in nemaline myopathy caused by mutations in the genes for nebulin and skeletal muscle alpha-actin. Neuromuscul Disord 2004 Sep;14(8-9):461-70.

195. **Agrawal PB,** Strickland CD, Midgett C, Morales A, Newburger DE, Poulos MA, Tomczak KK, Ryan MM, Iannaccone ST, Crawford TO, Laing NG, Beggs AH. Heterogeneity of nemaline myopathy cases with skeletal muscle alpha-actin gene mutations. Ann Neurol 2004 Jul;56(1):86-96.

196. **Wallgren-Pettersson C,** Jungbluth H. The congenital (structural) myopathies. Emery & Rimoin's Principles and Practice of Medical Genetics 2006.

197. **Wallgren-Pettersson C.** Congenital nemaline myopathy: A longitudal study. Helsinki, Finland; 1990.

198. **Laing NG,** Wilton SD, Akkari PA, Dorosz S, Boundy K, Kneebone C, Blumbergs P, White S, Watkins H, Love DR. A mutation in the alpha tropomyosin gene TPM3 associated with autosomal dominant nemaline myopathy NEM1. Nat Genet 1995 Jun;10(2):249.

199. **Pelin K,** Hilpela P, Donner K, Sewry C, Akkari PA, Wilton SD, Wattanasirichaigoon D, Bang ML, Centner T, Hanefeld F, Odent S, Fardeau M, Urtizberea JA, Muntoni F, Dubowitz V, Beggs AH, Laing NG, Labeit S, de la Chapelle A, Wallgren-Pettersson C. Mutations in the nebulin gene associated with autosomal recessive nemaline myopathy. Proc Natl Acad Sci U S A 1999 Mar 2;96(5):2305-10.

200. **Nowak KJ,** Wattanasirichaigoon D, Goebel HH, Wilce M, Pelin K, Donner K, Jacob RL, Hubner C, Oexle K, Anderson JR, Verity CM, North KN, Iannaccone ST, Muller CR, Nurnberg P, Muntoni F, Sewry C, Hughes I, Sutphen R, Lacson AG, Swoboda KJ, Vigneron J, Wallgren-Pettersson C, Beggs AH, Laing NG. Mutations in the skeletal muscle alpha-actin gene in patients with actin myopathy and nemaline myopathy. Nat Genet 1999 Oct;23(2):208-12.

201. **Simpson DM, Bender AN.** Human immunodeficiency virus-associated myopathy: Analysis of 11 patients. Ann Neurol 1988 Jul;24(1):79-84.

202. **Wallgren-Pettersson C, Laing NG.** Report of the 70th ENMC international workshop: Nemaline myopathy, 11-13 June 1999, Naarden, the Netherlands. Neuromuscul Disord 2000 Jun;10(4-5):299-306.

203. **Rifai Z,** Kazee AM, Kamp C, Griggs RC. Intranuclear rods in severe congenital nemaline myopathy. Neurology 1993 Nov;43(11):2372-7.

204. **Johnston JJ,** Kelley RI, Crawford TO, Morton DH, Agarwala R, Koch T, Schaffer AA, Francomano CA, Biesecker LG. A novel nemaline myopathy in the Amish caused by a mutation in troponin T1. Am J Hum Genet 2000 Oct;67(4):814-21.

205. **Donner K,** Ollikainen M, Ridanpaa M, Christen HJ, Goebel HH, de Visser M, Pelin K, Wallgren-Pettersson C. Mutations in the beta-tropomyosin (TPM2) gene - a rare cause of nemaline myopathy. Neuromuscul Disord 2002 Feb;12(2):151-8.
206. **Agrawal PB,** Greenleaf RS, Tomczak KK, Lehtokari VL, Wallgren-Pettersson C, Wallefeld W, Laing NG, Darras BT, Maciver SK, Dormitzer PR, Beggs AH. Nemaline myopathy with minicores caused by mutation of the CFL2 gene encoding the skeletal muscle actin-binding protein, cofilin-2. Am J Hum Genet 2007 Jan;80(1):162-7.
207. **Pelin K,** Donner K, Holmberg M, Jungbluth H, Muntoni F, Wallgren-Pettersson C. Nebulin mutations in autosomal recessive nemaline myopathy: An update. Neuromuscul Disord 2002 Oct;12(7-8):680-6.
208. **Anderson SL,** Ekstein J, Donnelly MC, Keefe EM, Toto NR, LeVoci LA, Rubin BY. Nemaline myopathy in the Ashkenazi Jewish population is caused by a deletion in the nebulin gene. Hum Genet 2004 Aug;115(3):185-90.
209. **Feng JJ, Marston S.** Genotype-phenotype correlations in ACTA1 mutations that cause congenital myopathies. Neuromuscul Disord 2009 Jan;19(1):6-16.
210. **Sparrow JC,** Nowak KJ, Durling HJ, Beggs AH, Wallgren-Pettersson C, Romero N, Nonaka I, Laing NG. Muscle disease caused by mutations in the skeletal muscle alpha-actin gene (ACTA1). Neuromuscul Disord 2003 Sep;13(7-8):519-31.
211. **Goebel HH,** Anderson JR, Hubner C, Oexle K, Warlo I. Congenital myopathy with excess of thin myofilaments. Neuromuscul Disord 1997 May;7(3):160-8.
212. **Jenis EH, Lindquist RR, Lister RC.** New congenital myopathy with crystalline intranuclear inclusions. Arch Neurol 1969 Mar;20(3):281-7.
213. **Laing NG,** Clarke NF, Dye DE, Liyanage K, Walker KR, Kobayashi Y, Shimakawa S, Hagiwara T, Ouvrier R, Sparrow JC, Nishino I, North KN, Nonaka I. Actin mutations are one cause of congenital fibre type disproportion. Ann Neurol 2004 Nov;56(5):689-94.
214. **Jungbluth H,** Sewry CA, Brown SC, Nowak KJ, Laing NG, Wallgren-Pettersson C, Pelin K, Manzur AY, Mercuri E, Dubowitz V, Muntoni F. Mild phenotype of nemaline myopathy with sleep hypoventilation due to a mutation in the skeletal muscle alpha-actin (ACTA1) gene. Neuromuscul Disord 2001 Jan;11(1):35-40.
215. **Ilkovski B,** Cooper ST, Nowak K, Ryan MM, Yang N, Schnell C, Durling HJ, Roddick LG, Wilkinson I, Kornberg AJ, Collins KJ, Wallace G, Gunning P, Hardeman EC, Laing NG, North KN. Nemaline myopathy caused by mutations in the muscle alpha-skeletal-actin gene. Am J Hum Genet 2001 Jun;68(6):1333-43.
216. **North KN, Laing NG.** Skeletal muscle alpha-actin diseases. Adv Exp Med Biol 2008;642:15-27.
217. **Brandis A, Aronica E, Goebel HH.** TPM2 mutation. Neuromuscul Disord 2008 Dec;18(12):1005.
218. **Sung SS,** Brassington AM, Grannatt K, Rutherford A, Whitby FG, Krakowiak PA, Jorde LB, Carey JC, Bamshad M. Mutations in genes encoding fast-twitch contractile proteins cause distal arthrogryposis syndromes. Am J Hum Genet 2003 Mar;72(3):681-90.
219. **Kee AJ, Hardeman EC.** Tropomyosins in skeletal muscle diseases. Adv Exp Med Biol 2008;644:143-57.
220. **Brooke MH, Engel WK.** The histographic analysis of human muscle biopsies with regard to fiber types. 4. Children's biopsies. Neurology 1969 Jun;19(6):591-605.
221. **Clarke NF,** Kolski H, Dye DE, Lim E, Smith RL, Patel R, Fahey MC, Bellance R, Romero NB, Johnson ES, Labarre-Vila A, Monnier N, Laing NG, North KN. Mutations in TPM3 are a common cause of congenital fiber type disproportion. Ann Neurol 2008 Mar;63(3):329-37.
222. **Wallgren-Pettersson C,** Rapola J, Donner M. Pathology of congenital nemaline myopathy. A follow-up study. J Neurol Sci 1988 Feb;83(2-3):243-57.
223. **Clarke NF, North KN.** Congenital fiber type disproportion--30 years on. J Neuropathol Exp Neurol 2003 Oct;62(10):977-89.
224. **Ryan MM,** Ilkovski B, Strickland CD, Schnell C, Sanoudou D, Midgett C, Houston R, Muirhead D, Dennett X, Shield LK, De Girolami U, Iannaccone ST, Laing NG,

North KN, Beggs AH. Clinical course correlates poorly with muscle pathology in nemaline myopathy. Neurology 2003 Feb 25;60(4):665-73.

225. **Domazetovska A,** Ilkovski B, Kumar V, Valova VA, Vandebrouck A, Hutchinson DO, Robinson PJ, Cooper ST, Sparrow JC, Peckham M, North KN. Intranuclear rod myopathy: Molecular pathogenesis and mechanisms of weakness. Ann Neurol 2007 Dec;62(6):597-608.

226. **Kaimaktchiev V,** Goebel H, Laing N, Narus M, Weeks D, Nixon R. Intranuclear nemaline rod myopathy. Muscle Nerve 2006 Sep;34(3):369-72.

227. **Fidzianska A.** "Cap disease" - a failure in the correct muscle fibre formation. J Neurol Sci 2002 Sep 15;201(1-2):27-31.

228. **Gibbels E,** Kellermann K, Schadlich HJ, Adams R, Haupt WF. Follow-up studies in a case of unusual congenital myopathy, suggestive of nemaline type. Acta Neuropathol 1992;83(4):371-8.

229. **Martland T,** Allibone R, Lowe J, Mellor D. Cap disease myopathy: A rare type of congenital myopathy. 1996; Suppl:S59.

230. **Cuisset JM**, Maurage CA, Pellissier JF, Barois A, Urtizberea JA, Laing N, Tajsharghi H, Vallee L. 'Cap myopathy': Case report of a family. Neuromuscul Disord 2006 Apr;16(4):277-81.

231. **Munoz-Jareno N,** Lopez-Martinez A, Martin Fernandez-Mayoralas D, Meizoso-Latova T, Cabello A. Cap myopathy: A case report. Rev Neurol 2007 Dec 1-15;45(11):669-71.

232. **Tajsharghi H,** Ohlsson M, Lindberg C, Oldfors A. Congenital myopathy with nemaline rods and cap structures caused by a mutation in the beta-tropomyosin gene (TPM2). Arch Neurol 2007 Sep;64(9):1334-8.

233. **Ohlsson M,** Quijano-Roy S, Darin N, Brochier G, Lacene E, Avila-Smirnow D, Fardeau M, Oldfors A, Tajsharghi H. New morphologic and genetic findings in cap disease associated with beta-tropomyosin (TPM2) mutations. Neurology 2008 Dec 2;71(23):1896-901.

234. **Goebel HH.** Cap disease uncapped. Neuromuscul Disord 2007 Jun;17(6):429-32.

235. **Oldfors A.** Hereditary myosin myopathies. Neuromuscul Disord 2007 May;17(5):355-67.

236. **Oldfors A, Lamont PJ.** Thick filament diseases. In: Nigel G. Laing, editor. The sarcomere and skeletal muscle disease. Landes Bioscience; 2008. ID: 258.

237. **Laing NG,** Laing BA, Meredith C, Wilton SD, Robbins P, Honeyman K, Dorosz S, Kozman H, Mastaglia FL, Kakulas BA. Autosomal dominant distal myopathy: Linkage to chromosome 14. Am J Hum Genet 1995 Feb;56(2):422-7.

238. **Meredith C,** Herrmann R, Parry C, Liyanage K, Dye DE, Durling HJ, Duff RM, Beckman K, de Visser M, van der Graaff MM, Hedera P, Fink JK, Petty EM, Lamont P, Fabian V, Bridges L, Voit T, Mastaglia FL, Laing NG. Mutations in the slow skeletal muscle fiber myosin heavy chain gene (MYH7) cause Laing early-onset distal myopathy (MPD1). Am J Hum Genet 2004 Oct;75(4):703-8.

239. **Darin N,** Tajsharghi H, Ostman-Smith I, Gilljam T, Oldfors A. New skeletal myopathy and cardiomyopathy associated with a missense mutation in MYH7. Neurology 2007 Jun 5;68(23):2041-2.

240. **Overeem S,** Schelhaas HJ, Blijham PJ, Grootscholten MI, ter Laak HJ, Timmermans J, van den Wijngaard A, Zwarts MJ. Symptomatic distal myopathy with cardiomyopathy due to a MYH7 mutation. Neuromuscul Disord 2007 Jun;17(6):490-3.

241. **Lamont PJ,** Udd B, Mastaglia FL, de Visser M, Hedera P, Voit T, Bridges LR, Fabian V, Rozemuller A, Laing NG. Laing early onset distal myopathy: Slow myosin defect with variable abnormalities on muscle biopsy. J Neurol Neurosurg Psychiatry 2006 Feb;77(2):208-15.

242. **Cancilla PA,** Kalyanaraman K, Verity MA, Munsat T, Pearson CM. Familial myopathy with probable lysis of myofibrils in type I fibers. Neurology 1971 Jun;21(6):579-85.

243. **Tajsharghi H,** Oldfors A, Macleod DP, Swash M. Homozygous mutation in MYH7 in myosin storage myopathy and cardiomyopathy. Neurology 2007 Mar 20;68(12):962.

244. **Richard I,** Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, Bourg N, Brenguier L, Devaud C, Pasturaud P, Roudaut C. Mutations in the proteolytic enzyme

calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 1995 Apr 7;81(1):27-40.

245. **Nigro V.** Molecular bases of autosomal recessive limb-girdle muscular dystrophies. Acta Myol 2003 Sep;22(2):35-42.

246. **Penisson-Besnier I,** Talvinen K, Dumez C, Vihola A, Dubas F, Fardeau M, Hackman P, Carpen O, Udd B. Myotilinopathy in a family with late onset myopathy. Neuromuscul Disord 2006 Jul;16(7):427-31.

247. **Siu BL,** Niimura H, Osborne JA, Fatkin D, MacRae C, Solomon S, Benson DW, Seidman JG, Seidman CE. Familial dilated cardiomyopathy locus maps to chromosome 2q31. Circulation 1999 Mar 2;99(8):1022-6.

248. **Bos JM, Poley RN,** Ny M, Tester DJ, Xu X, Vatta M, Towbin JA, Gersh BJ, Ommen SR, Ackerman MJ. Genotype-phenotype relationships involving hypertrophic cardiomyopathy-associated mutations in titin, muscle LIM protein, and telethonin. Mol Genet Metab 2006 May;88(1):78-85.

249. **Udd B,** Partanen J, Halonen P, Falck B, Hakamies L, Heikkila H, Ingo S, Kalimo H, Kaariainen H, Laulumaa V. Tibial muscular dystrophy. Late adult-onset distal myopathy in 66 Finnish patients. Arch Neurol 1993 Jun;50(6):604-8.

250. **Hackman P,** Vihola A, Haravuori H, Marchand S, Sarparanta J, De Seze J, Labeit S, Witt C, Peltonen L, Richard I, Udd B. Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. Am J Hum Genet 2002 Sep;71(3):492-500.

251. **Edstrom L,** Thornell LE, Albo J, Landin S, Samuelsson M. Myopathy with respiratory failure and typical myofibrillar lesions. J Neurol Sci 1990 May;96(2-3):211-28.

252. **Milic A,** Daniele N, Lochmuller H, Mora M, Comi GP, Moggio M, Noulet F, Walter MC, Morandi L, Poupiot J, Roudaut C, Bittner RE, Bartoli M, Richard I. A third of LGMD2A biopsies have normal calpain 3 proteolytic activity as determined by an in vitro assay. Neuromuscul Disord 2007 Feb;17(2):148-56.

253. **Van den Bergh PY,** Bouquiaux O, Verellen C, Marchand S, Richard I, Hackman P, Udd B. Tibial muscular dystrophy in a Belgian family. Ann Neurol 2003 Aug;54(2):248-51.

254. **Hackman P,** Marchand S, Sarparanta J, Vihola A, Penisson-Besnier I, Eymard B, Pardal-Fernandez JM, Hammouda e, Richard I, Illa I, Udd B. Truncating mutations in C-terminal titin may cause more severe tibial muscular dystrophy (TMD). Neuromuscul Disord 2008 Dec;18(12):922-8.

255. **Sewry CA,** Brown SC, Pelin K, Jungbluth H, Wallgren-Pettersson C, Labeit S, Manzur A, Muntoni F. Abnormalities in the expression of nebulin in chromosome-2 linked nemaline myopathy. Neuromuscul Disord 2001 Mar;11(2):146-53.

256. **Gurgel-Giannetti J,** Reed U, Bang ML, Pelin K, Donner K, Marie SK, Carvalho M, Fireman MA, Zanoteli E, Oliveira AS, Zatz M, Wallgren-Pettersson C, Labeit S, Vainzof M. Nebulin expression in patients with nemaline myopathy. Neuromuscul Disord 2001 Mar;11(2):154-62.

257. **Gurgel-Giannetti J,** Bang ML, Reed U, Marie S, Zatz M, Labeit S, Vainzof M. Lack of the C-terminal domain of nebulin in a patient with nemaline myopathy. Muscle Nerve 2002 May;25(5):747-52.

258. **McCaughan KK,** Brown CM, Dalphin ME, Berry MJ, Tate WP. Translational termination efficiency in mammals is influenced by the base following the stop codon. Proc Natl Acad Sci U S A 1995 Jun 6;92(12):5431-5.

259. **Littlefield RS, Fowler VM.** Thin filament length regulation in striated muscle sarcomeres: Pointed-end dynamics go beyond a nebulin ruler. Semin Cell Dev Biol 2008 Dec;19(6):511-9.

260. **Wattanasirichaigoon D,** Swoboda KJ, Takada F, Tong HQ, Lip V, Iannaccone ST, Wallgren-Pettersson C, Laing NG, Beggs AH. Mutations of the slow muscle alpha-tropomyosin gene, TPM3, are a rare cause of nemaline myopathy. Neurology 2002 Aug 27;59(4):613-7.

261. **Dabrowska R, Sosinski J, Drabikowski W.** Dimerization of the polypeptide chains of skeletal muscle tropomyosin. Biochim Biophys Acta 1983 Mar 30;743(3):331-7.

262. **Morais AC, Ferreira ST.** Folding and stability of a coiled-coil investigated using chemical and physical denaturing agents: Comparative analysis of polymerized

and non-polymerized forms of alpha-tropomyosin. Int J Biochem Cell Biol 2005 Jul;37(7):1386-95.

263. **Tan P,** Briner J, Boltshauser E, Davis MR, Wilton SD, North K, Wallgren-Pettersson C, Laing NG. Homozygosity for a nonsense mutation in the alpha-tropomyosin slow gene TPM3 in a patient with severe infantile nemaline myopathy. Neuromuscul Disord 1999 Dec;9(8):573-9.

264. **Kohn WD, Kay CM, Hodges RS.** Orientation, positional, additivity, and oligomerization-state effects of interhelical ion pairs in alpha-helical coiled-coils. J Mol Biol 1998 Nov 13;283(5):993-1012.

265. **Brown JH,** Zhou Z, Reshetnikova L, Robinson H, Yammani RD, Tobacman LS, Cohen C. Structure of the mid-region of tropomyosin: Bending and binding sites for actin. Proc Natl Acad Sci U S A 2005 Dec 27;102(52):18878-83.

266. **Greenfield NJ, Fowler VM.** Tropomyosin requires an intact N-terminal coiled coil to interact with tropomodulin. Biophys J 2002 May;82(5):2580-91.

267. **Thierfelder L,** Watkins H, MacRae C, Lamas R, McKenna W, Vosberg HP, Seidman JG, Seidman CE. Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: A disease of the sarcomere. Cell 1994 Jun 3;77(5):701-12.

268. **Watkins H,** Conner D, Thierfelder L, Jarcho JA, MacRae C, McKenna WJ, Maron BJ, Seidman JG, Seidman CE. Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. Nat Genet 1995 Dec;11(4):434-7.

269. **Karibe A,** Tobacman LS, Strand J, Butters C, Back N, Bachinski LL, Arai AE, Ortiz A, Roberts R, Homsher E, Fananapazir L. Hypertrophic cardiomyopathy caused by a novel alpha-tropomyosin mutation (V95A) is associated with mild cardiac phenotype, abnormal calcium binding to troponin, abnormal myosin cycling, and poor prognosis. Circulation 2001 Jan 2;103(1):65-71.

270. **Tajsharghi H,** Kimber E, Holmgren D, Tulinius M, Oldfors A. Distal arthrogryposis and muscle weakness associated with a beta-tropomyosin mutation. Neurology 2007 Mar 6;68(10):772-5.

271. **Monnier N,** Lunardi J, Marty I, Mezin P, Labarre-Vila A, Dieterich K, Jouk PS. Absence of beta-tropomyosin is a new cause of escobar syndrome associated with nemaline myopathy. Neuromuscul Disord 2009 Feb;19(2):118-23.

272. **Escobar V,** Bixler D, Gleiser S, Weaver DD, Gibbs T. Multiple pterygium syndrome. Am J Dis Child 1978 Jun;132(6):609-11.

273. **Durling HJ,** Reilich P, Muller-Hocker J, Mendel B, Pongratz D, Wallgren-Pettersson C, Gunning P, Lochmuller H, Laing NG. De novo missense mutation in a constitutively expressed exon of the slow alpha-tropomyosin gene TPM3 associated with an atypical, sporadic case of nemaline myopathy. Neuromuscul Disord 2002 Dec;12(10):947-51.

274. **Penisson-Besnier I,** Monnier N, Toutain A, Dubas F, Laing N. A second pedigree with autosomal dominant nemaline myopathy caused by TPM3 mutation: A clinical and pathological study. Neuromuscul Disord 2007 Apr;17(4):330-7.

Appendix 1: Codons and amino acids encoded

	T	C	A	G	
T	TTT **Phenylalanine** (Phe)	TCT **Serine** (Ser)	TAT **Tyrosine** (Tyr)	TGT **Cysteine** (Cys)	T
	TTC Phe	TCC Ser	TAC Tyr	TGC Cys	C
	TTA **Leucine** (LeT)	TCA Ser	TAA **STOP**	TGA **STOP**	A
	TTG Leu	TCG Ser	TAG **STOP**	TGG **Tryptophan** (Trp)	G
C	CTT **Leucine** (LeT)	CCT **Proline** (Pro)	CAT **Histidine** (His)	CGT **Arginine** (Arg)	T
	CTC Leu	CCC Pro	CAC His	CGC Arg	C
	CTA Leu	CCA Pro	CAA **Glutamine** (Gln)	CGA Arg	A
	CTG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	ATT **Isoleucine** (Ile)	ACT **Threonine** (Thr)	AAT **Asparagine** (Asn)	AGT **Serine** (Ser)	T
	ATC Ile	ACC Thr	AAC Asn	AGC Ser	C
	ATA Ile	ACA Thr	AAA **Lysine** (Lys)	AGA **Arginine** (Arg)	A
	ATG **Methionine** (Met) or **START**	ACG Thr	AAG Lys	AGG Arg	G
G	GTT **Valine** Val	GCT **Alanine** (Ala)	GAT **Aspartic acid** (Asp)	GGT **Glycine** (Gly)	T
	GTC (Val)	GCC Ala	GAC Asp	GGC Gly	C
	GTA Val	GCA Ala	GAA **Glutamic acid** (Glu)	GGA Gly	A
	GTG Val	GCG Ala	GAG Glu	GGG Gly	G

Abbreviations of the amino acids:

Ala = A, Arg = R, Asn = N, Asp = D, Cys = C, Gln = Q, Glu = E, Gly = G, His = H, Ile = I, Leu = L,
Lys = K, Met = M, Phe = F, Pro = P, Ser = S, Thr = T, Trp = W, Tyr = Y, Val = V