Application of microarray-based gene expression technology to neuromuscular disorders

Sterrenburg, P.J.E.

Chapter 1

Introduction

"Transcript abundance doesn't tell us everything, but it tells us a lot more then we knew before". - Patrick O. Brown, Stanford University

1.1 General Introduction

In the past 10 years high throughput techniques to study DNA, RNA and protein alterations have caused a revolution in molecular research. While previously genes, RNAs and proteins were studied one-by-one, it is now possible to get a comprehensive picture of a considerable part of the genome, transcriptome or proteome in a single experiment. Studying a disease, a combination of aforementioned "omes" can provide us with detailed information as to which mechanisms and pathways underlie the disease. In terms of biological relevance, the proteome is the most informative, showing the ultimate products that determine the state of a cell. However, in terms of measurement, it is the most complex one because of the extremely wide abundance range, and because of the tremendous diversity in physical-chemical properties. The transcriptome is the easiest to monitor because of well established protocols mainly due to the relatively homogeneous molecular properties of RNA, leading to relatively simple and low cost experiments. By studying the transcriptome, i.e. looking one step further than the genome, more changes can be found than just the original DNA modification. Expression changes can hint to responsible disease mechanisms, but we must be careful because gene expression levels do not always correspond to protein levels (1-3) and do not represent the protein status (active/inactive). Studies have shown both similarities and discrepancies between transcriptome and proteome but because transcriptome measurement is easier and cheaper and a full proteome measurement is not yet possible, the former is presently more often used to study pathways involved in diseases.

Techniques to measure gene expression include Northern blotting (4), Quantitative RT-PCR (5), Differential Display (6), Serial Analysis of Gene Expression (SAGE) (7) and RNase protection assay (RPA) (8). Northern blotting is a nucleic acid hybridization-based technique developed on the basis of the principles of its predecessor technique, Southern blot (9) and is very sensitive and inexpensive. However, large-scale analysis is difficult and it requires prior knowledge on differentially expressed transcripts. Quantitative Reverse Transcriptase PCR (Q-RT-PCR) is based on the PCR amplification of cDNA and measures the increase in product with each PCR cycle. The technique is very sensitive, highly quantitative and reproducible but too costly as a high-throughput system, and only known or expected changes can be tested. Differential Display is based on quantitative differences in PCR-amplified cDNA products on a gel. Although the oldest and simplest of all transcription profiling techniques, there are some disadvantages to this technique. It is not very quantitative, the sensitivity can be an issue, it is prone to false-positives and it is not easy to automate. The principle of SAGE is to convert every mRNA molecule into a short (10-14 base) tag. It is a very direct and quantitative method with a near infinite dynamic range. The reason why it is not frequently used, is because it is very time-consuming and relatively expensive. Gene expression can also be measured with the RNase protection assay (RPA). The principle of this procedure is an in solution hybridization of labeled antisense riboprobes to the total RNA and quantification on a gel or by running on an automated sequencer. It is a sensitive and reproducible method because the probes are uniformly labelled, but also time consuming.

With new technologies emerging, DNA microarray technology was introduced as a method to measure gene expression of thousands of transcripts at the same time. The array concept was first described by Ed Southern who patented the principle in 1989 (Oxford Gene Technology), but it was not until halfway through the nineties that the technique became popular. In 1994 Pease *et al*. from Affymetrix described the development of photolithographic synthesis of densely packed arrays of oligonucleotide probes (10). These probe arrays, which the authors called DNA chips, could be used for hybridization to target DNA sequences. The authors predicted that the method would be a powerful tool for human genetics, diagnostics, pathogen detection, and DNA molecular recognition. In 1995, a seminal paper by the group of Brown was published on quantitative monitoring of gene expression in Arabidopsis using cDNA microarray (11). With this publication, the microarray technology was born. In ten years time, more than ten thousand papers have appeared using the microarray technology, proving its widespread usefulness. In this thesis, gene expression microarrays are applied to the research of muscle cell differentiation (**Chapter 3**) and two muscular disorders (**Chapters 4-5**). The following paragraphs provide an introduction on these subjects.

1.2 Myogenesis and muscle regeneration

Myogenesis is a multistage process characterized by the specialization, proliferation and terminal differentiation of skeletal muscle cells in the embryo. In adult muscle, a similar mechanism is activated after injury, which makes the muscle a tissue with regenerating capacity. A detailed understanding of the mechanisms that regulate differentiation during myogenesis and regeneration, can provide an essential foundation to investigate whether and how these processes are affected in muscular dystrophies, and how to intervene in these processes to obtain therapeutic effects.

1.2.1 Myogenesis

All vertebrate skeletal muscles (apart from cranial muscles) arise from the somites. Somites are transient epithelial spheres which split off from the paraxial mesoderm lining both sides of the neural tube. Myogenic precursors, expressing Pax3 (12), are first identified in the dermomyotome, an epithelial layer located in the dorsal compartment of the somite, where they receive signals from the Wingless-type MMTV integration site family (Wnts) and Sonic hedgehog (Shh) from surrounding embryonic structures (Figure 1). Specification of the precursor cells to the myogenic lineage (myoblasts) depends on these signaling pathways which induce the expression of the primary myogenic regulatory factors (MRFs), Myf5 and MyoD (13-16). Myf5 expression is induced in the dorsal-medial somites (which give rise to trunk and intercostal muscles), and is followed by expression of MyoD in the dorsal-lateral somites (which give rise to body wall and limb muscles). Expression of MyoD and/or Myf5 is the key step in myoblast commitment since disruption of both genes results in the absence of skeletal myoblasts (17). Commited myoblasts migrate laterally to form the myotome. There, differentiation and fusion of myoblasts to form multinucleated myotubes is initiated. This step is initiated by another MRF, myogenin, as a key regulator. Mice lacking myogenin have very poorly developed skeletal muscle tissue even though myoblasts are present (18,19). The last known MRF, MRF4, shows temporal expression during development and has been shown to act both in muscle determination and terminal differentiation (20,21). The MRFs contain a basic helix-loop-helix (bHLH) motif for heterodimerization with E-proteins. When

heterodimerized, they can bind at sites known as E-boxes (CANNTG) in the promoter and enhancer regions of many skeletal muscle-specific genes (22,23). For effective myogenesis, both the myocyte enhancer factor 2 (Mef2) family and Six family of transcription factors are necessary (24). They bind to Mef2 and Mef3 binding sites of muscle-specific genes, which are in close proximity of E boxes, and together with binding of the MRFs they activate transcription in a cooperative fashion (25). Figure 1 gives a schematic overview of the different stages in maturation of a myofiber.

Inhibiting factors which negatively regulate myogenesis are members of the TGF-β superfamily and helix-loop-helix (HLH) proteins. HLH proteins (eg. Id, Twist, Mist1) can inhibit myogenic transcription, by directly binding to E proteins and/or MyoD family proteins, and blocking their ability to bind E boxes and/or activate transcription at muscle-specific promoters (26,27). TGF-β can inhibit muscle differentiation by silencing the transcription-activating potential of myogenic bHLH proteins without affecting their ability to bind DNA (28).

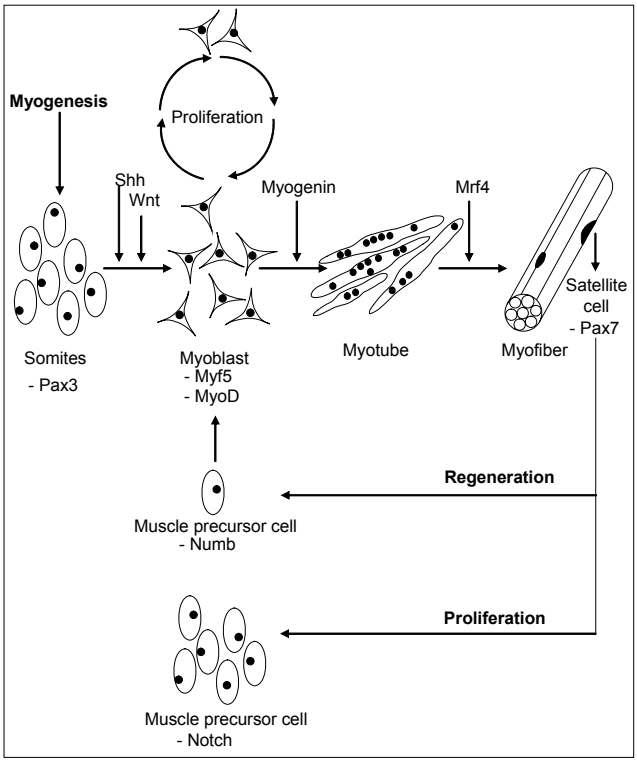

Figure 1. Schematic representation of muscle cell proliferation, differentiation and regeneration.

1.2.2 Muscle regeneration

During muscle development, a distinct subpopulation of muscle precursor cells (mpc), expressing Pax7, fails to differentiate (29). They remain associated with the surface of the developing myofiber, between the basal lamina and the connective tissue, as quiescent muscle satellite cells (30). During muscle regeneration of the mature muscle, a part of the

myogenesis process is mimicked when these satellite cells become activated myoblast precusor cells (mpc) and either fuse with existing myotubes or form new myotubes (31). Activation of satellite cells is mediated by the Notch signaling pathway. By binding of the Notch ligand Delta, Notch is activated; triggering a cascade which leads to transcription of specific genes (32) and thereby promoting myoblast proliferation (33). Proliferation of mpcs is also stimulated by a number of growth factors, including fibroblast growth factor (FGF), insulin-like growth factor (IGF) and hepatocyte growth factor (HGF) (34-36). These growth factors appear to be stored in the basal lamina and extracellular matrix by binding to heparan sulfate proteoglycans, which play an important role in muscle regeneration (37-39). In order to undergo myogenic differentiation, proliferating mpcs must overexpress Numb, which prevents nuclear translocation of Notch and in this way inhibits the Notch signaling pathway. Cells then exit the cell cycle and start to differentiate into myoblasts (Figure 1) (33).
Numb also plays a role in the maintenance of the satellite cell pool. It is asymmetrically expressed in dividing myogenic progenitor cells through which two pools of cells arise. One expressing Numb, where inhibited Notch signalling leads to myogenic differentiation, and the other not expressing Numb, where active Notch signalling promotes undifferentiated proliferation (33).
Again upon activation of satellite cells, Myf5 and MyoD are expressed, which, in their turn, activate the other myogenic regulatory factors and muscle-specific proteins for terminal differentiation of the myoblasts (Figure 1)(40-42).

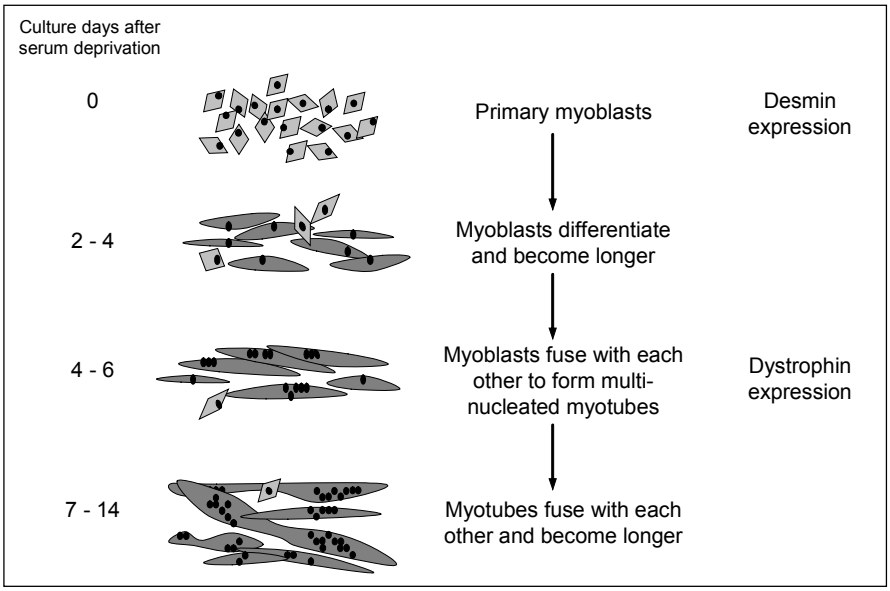

Figure 2. Schematic representation of *in vitro* muscle cell differentiation and fusion of myoblasts to form myotubes.

1.2.3 Model systems

To study myogenesis and / or muscle regeneration different model systems can be used. The muscle injury model is an *in vivo* model system in which by injection of a toxic agent (eg. cardiotoxin) in muscle, specific degeneration of myofibers is induced. After injection, an increase in mononucleated cells is observed which is due to both inflammatory cell

infiltration and proliferation of satellite cells. After three days, new myotubes appear and after two weeks, normal muscle morphology is restored, showing complete regeneration (43). An alternative method is to isolate satellite cells from adult muscle tissue and bring these into culture (44). In high serum conditions, these satellite cells are activated to proliferate. Upon serum withdrawal the activated myoblasts exit the cell cycle and start to differentiate. Cells become longer and eventually fuse with each other to form multinucleated myotubes (Figure 2) (45).
When muscle tissue or myoblasts are not available, non-muscle cells can be used and forced into myogenesis with viral delivery of the myogenic determination gene *MYOD1* (46,47). By isolating RNA from primary myogenic cell cultures, or Myod1-transfected non-muscle cell cultures or toxin-injected muscles at different time points, the different steps in the myogenesis process can be studied by gene expression profiling.

1.3 Muscular Dystrophies

There are more than 20 different types of muscular dystrophy which are characterized by progressive skeletal muscle weakness, defects in muscle proteins, and necrosis of muscle cells and tissue. They can cause shortened life expectancy and there are currently no cures (48). The inheritance pattern, age of onset, initial muscles affected and rate of progression can vary widely within and between the different disorders. In this thesis two very distinct muscular dystrophies are discussed which have been explored using gene expression microarrays; Duchenne muscular dystrophy (DMD) and oculopharyngeal muscular dystrophy (OPMD).

1.3.1 Duchenne muscular dystrophy

Phenotype
Duchenne muscular dystrophy (DMD) is a neuromuscular disorder first described by Edward Meryon in 1851 (49) and later named after Duchenne who gave a detailed description of the disease (50). It is an X-linked recessive disorder with an incidence of 1 in 3300-5000 live male births (48). The first symptoms arise early between 2 and 6 years of age. The disease is characterized by progressive skeletal muscle weakness first affecting the muscles of the hips, pelvic area, thighs and shoulders. Eventually all voluntary muscles will be affected with addition of the heart and respiratory muscles. By the age of ten, most patients are wheelchair bound. Death occurs almost always before the age of 30 due to respiratory problems or heart failure (48).
Structural changes in the muscle include variation in fibre size, centrally located nuclei, degeneration and regeneration of the muscle fibres, fibre necrosis with phagocytosis, and eventually, replacement by fat and connective tissue (49-51). As a result of the fibre necrosis, leakage occurs and muscle proteins can be found in the plasma (e.g. elevated creatine kinase (CK) levels) (52).

Genetic defect
The DMD disease locus was mapped to Xp21 (53-59). The *DMD* gene was first identified by Monaco *et al*. in 1986 (60). Many different mutations have been identified in the *DMD* gene which all cause DMD (http://www.dmd.nl/); nearly every patient has a unique mutation (1/3 de novo). Almost 80% of the mutations are deletions or duplications of one or more exons (61). There is a clear relation between disease severity and preservation or disruption of the

reading frame of the DMD gene. Mutations that disrupt the reading frame (frame-shift) or introduce a stopcodon (nonsense) lead to DMD and mutations that maintain the translational open reading frame lead to the less severe Becker muscular dystrophy (BMD) which had previously been mapped to the same location as DMD (62,63). Also missense mutations that affect the amino and carboxy-termini of the gene, functionally important regions, can cause DMD (64,65).

Disease mechanism

Dystrophin, the protein product of the *DMD* gene (66), is under control of seven (tissue-specific) promoters generating different isoforms of the protein expressed in skeletal muscle, cardiac muscle, brain, retina, kidney, purkinje cells and Schwann cells (60,67-72) (www.dmd.nl). It is a structural protein forming a bridge across the sarcolemma (plasma membrane of cardiac and skeletal muscle) (73), connecting the basal lamina of the extracellular matrix to the inner cytoskeleton (Figure 3). The N-terminal part binds to F-actin and the C-terminal domain interacts with the dystrophin-glycoprotein complex (DGC) located at the sarcolemma (74-76). In this way it stabilizes the sarcolemma and protects muscle fibres from long-term contraction-induced damage and necrosis (76-78). In addition to its mechanical function, the DGC has also been suggested to have a role in cellular communication by acting as a transmembrane signaling complex (76,79).

A frameshift or nonsense mutation in the *DMD* gene disrupts the open reading frame resulting in a premature stopcodon (63). The result is a truncated dystrophin protein, which lacks its C-terminal anchor, is non-functional and unstable and cannot fulfill its bridge function. Furthermore, the absence of dystrophin leads to a dramatic loss of all components of the DGC resulting in membrane instability, abnormal calcium influx, mitochondrial changes and eventually necrosis (80-83).

Table 1. Different theories on defective regeneration in DMD muscle

Replicative senescence		
Blau *et al*. (87)	1983	Satellite cell pool shrinking in DMD muscle
Heslop *et al*. (88)	2000	Satellite cell pool shrinking in DMD muscle Decreased proliferation
Decary *et al*. (94)	2000	Telomere shortening in DMD
Impaired differentiation		
Wakayama *et al*. (89)	1976	Satellite cell pool increasing in DMD muscle
Ishimoto *et al*. (90)	1983	Satellite cell pool increasing in DMD muscle
Maier *et al*. (91)	1999	Satellite cell pool increasing in DMD muscle
Oexle *et al*. (92)	1997	Normal telomere length in DMD
Burr *et al*. (95)	1997	Decreased ratio of Cu/Zn-superoxide dismutase to glutathione peroxidase in DMD

Regeneration in Duchenne muscular dystrophy

At the early stages of the disease process in DMD muscle, regenerating fibres are commonly found. However, the process of regeneration becomes less frequent as the disease progresses and fibres undergoing necrosis become more obvious (48). It is not clear why human DMD muscle is not capable of regenerating completely. Satellite cells of DMD patients have been studied morphologically and with regard to their ability to proliferate and differentiate. One of the possible reasons for inefficient regeneration in DMD muscle is replicative aging, meaning that somatic cells have an upper limit of possible replications (84). The continuous

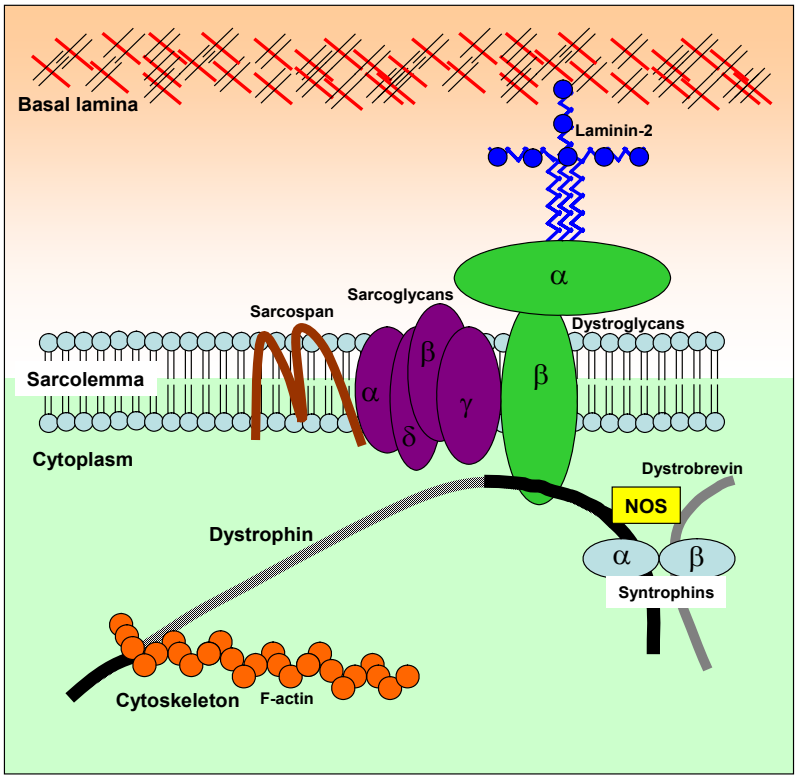

Figure 3. Schematic representation of the localization of the Dystrophin protein in the cell. With its N-terminal part it binds to F-actin and the C-terminal domain interacts with β-dystroglycan located at the sarcolemma.

regeneration of DMD muscle makes the DMD satellite cells age much faster than normal. Ultimately proliferation stops due to replicative aging and if differentiation continues, the satellite cell pool gets exhausted. Satellite cell depletion, telomere shortening (85) and an elevated ratio of Cu/Zn-superoxide dismutase to glutathione peroxidase (86) are all biomarkers for replicative aging. However, in DMD muscle both shrinking (87,88) and increasing (89-91) pools of satellite cells have been reported and both normal (92,93) and shorter (94) telomeres have been found, so there is no consensus about this issue. Burr *et al*. have shown that the ratio of Cu/Zn-superoxide dismutase to glutathione peroxidase is decreased in DMD muscle, which is not in favor of the replicative aging theory (95). In culture, DMD cells show an impaired proliferation in comparison to control cells. This is explained by the presence of the anti-proliferative factor TGF-β1 in the medium of the DMD cells and can be neutralized by adding specific antibodies. This suggests that impaired proliferation of DMD cells in culture is not due to replicative senescence (96).

Another possible reason for inefficient regeneration is impaired differentiation or fusion of satellite cells which has been seen in DMD cell cultures (97-99). In concordance with this, it is reported that macrophages invade DMD muscle (100) which are known to enhance proliferation and delay differentiation (101). Furthermore, histological analysis of DMD muscle shows abnormal, small and short fibers, caliber variations, and multiple fiber branchings which can be explained by incomplete fusion of myotubes and myoblasts (99).

Gene expression profiling of both healthy and DMD myoblasts in culture can provide insight as to how proliferation and differentiation are regulated in these cells and maybe can resolve some of the controversy around this issue.

The second muscular dystrophy explored by gene expression microarray studies in this thesis is Oculopharyngeal muscular dystrophy.

1.3.2 Oculopharyngeal muscular dystrophy

Phenotype
Oculopharyngeal muscular dystrophy (OPMD) was first described by Taylor in 1915 (102), but it was not until 1962 that Victor *et al*. described the disease under its current name (103). The worldwide incidence is 1:200.000 with an increased incidence in some populations; French Canadian 1:1000 (104), Spanish Americans living in New Mexico (105), Bukhara Jews 1:600 (106), and Uruguayans (107)). It shows an autosomal dominant inheritance but an autosomal recessive case has also been described (108). The first symptoms usually appear in the fifth or sixth decade of life and are characterized by eyelid ptosis and progressive swallowing difficulties (dysphagia) (103). During progression proximal limb weakness occurs, and recently this has been reported as one of the first major complaints in a population of Dutch OPMD patients (109). Other, later occuring symptoms may include a nasal voice, weakness and atrophy of the tongue, external ophthalmoplegia, mild facial temporal and masseter muscle involvement (110-113).
A pathological hallmark of the disease is the presence of intranuclear inclusions (INIs) comprising clusters of tubular filaments in 2-5% of skeletal muscle cells (114,115). These INIs are not seen in other cells, not even in muscle satellite cells (114,116,117). Other characteristics found in muscle cells of OPMD patients are small angulated fibers that often react strongly for oxidative enzymes and are more frequently type 1 then type 2, as seen with ATP-ase reactions (118,119). Rimmed vacuoles can be seen within muscle fibers and consist of irregularly round or polygonal clear spaces lined by a ring of material that is basophilic (also found in inclusion body myositis (120)). They vary in size and shape and contain membranous and myeloid structures, a few glycogen granules, and debris arising from the breakdown of unidentified muscle components. Mitochondrial changes have also been reported as a phenomenon of OPMD. Cytochrome-c-oxidase-negative fibers occur and aggregates of mitochondria containing paracrystalline inclusions can be seen with electron microscopy (121-123). Muscular changes that are common to other muscular dystrophies also occur in OPMD, including loss of muscle fibers, abnormal variation in fiber size, increase in the number of nuclei and the presence of internal nuclei (111,118,119).

Genetic defect
The disease locus was mapped to chromosome 14q11.1 by completing a linkage analysis using large French Canadian families (124). Positional cloning led to the discovery of a short GCG expansion leading to the lengthening of a polyalanine tract in the first exon of the poly(A) binding protein nuclear 1 (*PABPN1*) gene (108). Dominant mutations consist of the normal $(GCG)_6$ stretch elongated by two to seven (GCG) repeats whereas the only reported recessive form has a $(GCG)_7$ repeat on both chromosomes (108). Some OPMD cases have been described without a GCG expansion, but having three to five extra nonidentical polyalanine coding triplets (125-127), which favors the theory of unequal crossing-over over

the slippage model as the causative mechanism of PABPN1 mutations (table 2). Autosomal dominant OPMD is thus caused by an expansion of a non homogenous GCG repeat beyond the threshold of 12 alanines. No clear correlation between the length of the repeat and phenotypic expression of the disease has been reported (126,128).

Table 2. Different alanine expansions in the *PABPN1* gene

Normal PABPN1	**Alanine repeats**	**OPMD**
(GCG)6 (GCA)3 (GCG)	10	---
Mutated PABPN1		
(GCG)6 **(GCG)** (GCA)3 (GCG)	11	recessive
(GCG)6 **(GCG)2-7** (GCA)3 (GCG)	12-17	dominant
(GCG)6 **(GCA) (GCG)2** (GCA)3 (GCG)	13	dominant
(GCG)6 **(GCA) (GCG)3** (GCA)3 (GCG)	14	dominant
(GCG)6 **(GCA)2 (GCG)2** (GCA)3 (GCG)	14	dominant
(GCG)6 **(GCA)3 (GCG)2** (GCA)3 (GCG)	15	dominant

Gene dosage has a clear effect on the age of onset and the severity of the OPMD phenotype. The most severe OPMD phenotype is reported in individuals carying a homozygous dominant OPMD mutation (108,117,129). In addition, among severely affected OPMD patients in a French Canadian cohort, 20% are compound heterozygotes for the dominant mutation and have a $(GCG)_7$ polymorphism in their other copy of the *PABPN1* gene (108).

Disease mechanism

Poly(A) binding protein nuclear 1 is a ubiquitously expressed protein primarily present in the nuclear speckles and functionally involved in polyadenylation. More specifically, it binds with high affinity to nascent poly(A) tails and thereby stimulates its elongation by poly(A) polymerase (Figure 4, PAP, (130)). PAP catalyzes the polyadenylation reaction but without PABPN1 it is almost inactive due to a low affinity for the RNA (130). PABPN1 is also involved in the mechanism that regulates the length of the poly(A) tail. It has been proposed that the number of PABPN1 molecules incorporated into the elongating complex is involved in a counting mechanism that determines the length of the tail (131). In addition, PABPN1 may contribute to export of mRNA from the nucleus to the cytoplasm (132-134). Although the gene is ubiquitously expressed, the phenotype is muscle-specific and polyadenylation seems to be unaffected by the repeat expansion (116). Because the intranuclear inclusions (INI) are the most specific hallmark of the disease and only occur in the muscle cells, most studies on the disease mechanism have focussed on this phenomenon.

PABPN1 appears to locate to the INIs that further include poly(A) RNA, ubiquitin, heat-shock proteins and proteasome subunits (116,135,136). The percentage of cells containing INIs correlates with the severity of the disease. This is shown by seven severe homozygote cases of a dominant OPMD mutation, which have almost twice as many inclusions in comparison with heterozygotes of the dominant mutation (9.4% vs. 4.9% respectively) (129). Furthermore, upon treatment with doxycycline, the inclusion percentage is reduced in a transgenic mouse model for OPMD and their muscle condition is improved (137). This and the presence of ubiquitin and proteasome subunits in the inclusions, suggest that INIs are toxic to the cell (138,139). If these INIs are so important to the disease phenotype, two questions emerge; how is the muscle-specific formation of INIs initiated, and which proteins and mRNAs are trapped in the INIs by which their function is hampered and a phenotypic effect can be seen? Cellular and animal models of OPMD have been established to attempt to answer these questions.

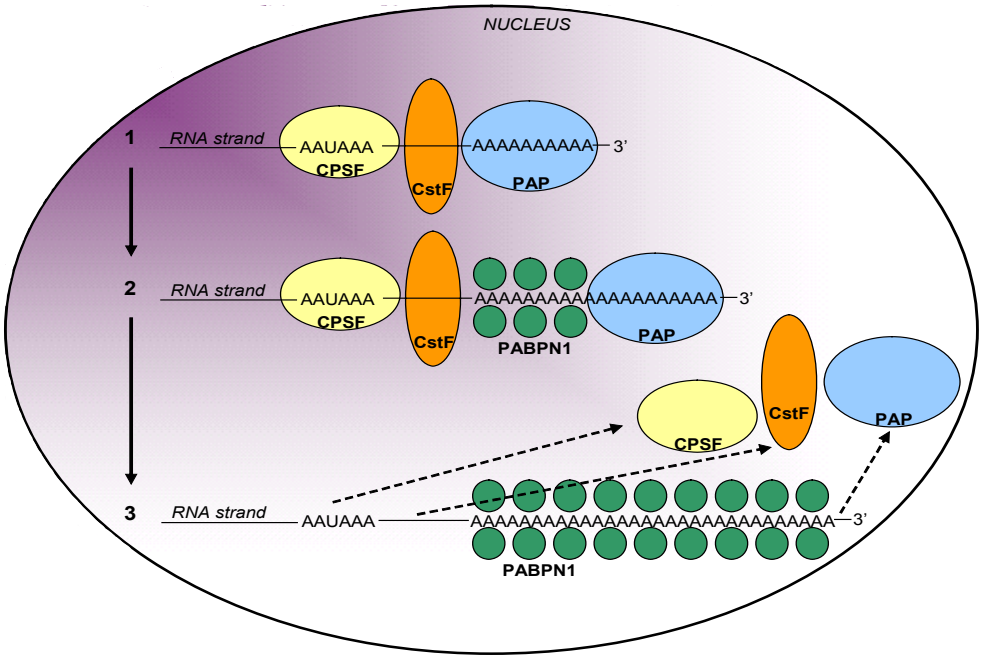

Figure 4. Schematic representation of the polyadenylation process.
1. Cleavage and polyadenylation specific factor (CPSF) and Cleavage stimulation factor (CstF) bind to the pre-mRNA. Polyadenylate polymerase (PAP) starts writing the polyadenosine tail.
2. Poly(A) binding protein 1 (PABPN1) immediately binds to the new polyadenosine sequence, stabilizing the poly(A) tail.
3. PABPN1 specifies when polyadenylation should stop (200 residues). CPSF, CstF and PAP dissociate.

Cellular and animal models of OPMD

It has been suggested that the polyalanine expansion in PABPN1 induces or facilitates formation of INIs. *In vitro* studies show however, that formation of PABPN1 aggregates is not dependent on the elongated repeat, but on protein domains in PABPN1 which are functionally required for polyadenylation (138,140). Furthermore, different cellular models show that overexpression of wild-type PABPN1 has the same effect on INI formation as an expansion of the alanine repeat (140,141). Given the muscle-specificity of the intranuclear inclusions, other additional factors could be necessary for inclusion formation, possibly a muscle-specific protein or mRNA. The fact that the INI formation is restricted to the nucleus (142) and the observation that the inclusion formation in muscle cells only occurs in differentiated cells (116, Chapter 5) further reduces the list of possible candidates. Kim *et al*. showed that PABPN1 interacts with Ski-interacting protein (SKIP), a transcription factor, to stimulate muscle-specific gene expression through MyoD and myogenin increased expression (143). This would explain the muscle-specific phenotype of the disease. However, SKIP is also highly expressed in proliferating myoblasts, which cannot explain why the INIs are not formed in these cells. To date, other candidates have not been proposed.

Cellular processes can be affected by the presence of intranuclear inclusions. Not only because of their toxicity, but also by entrapping other proteins that normally function in the nucleus. Using large-scale gene expression microarraying, Corbeil-Girard *et al.* found

pathways which were up- and downregulated in cells containing INIs. Several nuclear genes appeared to be upregulated in their cell model with 60-80 % inclusions. Following up on this, they revealed the entrapment of these nuclear proteins in the inclusions, of which none could explain the (muscle-specific) phenotype of the disease (141). However, a non-muscle cell line was used for the experiments, so that essential interactions with muscle-specific proteins or mRNAs may well have been missed.
Some mouse models for OPMD have been generated in which mutated PABPN1 is overexpressed (137,144,145). Hino *et al.* demonstrate that a threshold level of mutant PABPN1 overexpression is necessary to form intranuclear inclusions and an early myopathic phenotype (144). Furthermore, northern blot analysis in different tissues of the transgenic mice revealed ubiquitous expression of the transgene with only INI formation in muscle and a muscle-specific phenotype. These results confirm the theory that probably an additional muscle-specific factor is necessary to form intranuclear inclusions.
Dion *et al*. generated an animal model of OPMD with no specific muscle phenotype but a neuronal phenotype and the presence of INIs in neurons (145). With this model it is not possible to study the muscle-specific INI formation, but it does provide evidence that INI formation requires a postmitotic, stable nuclear environment allowing inclusions to be established and persist in the cells (145).
Expression profiling of a muscle cell model of OPMD, an animal model of OPMD or muscle tissue derived from OPMD patients could provide more insight in the disease mechanism of Oculopharyngeal muscular dystrophy.

1.4 Gene expression profiling using DNA microarrays

1.4.1 Principle and platforms

The principle of a DNA microarray experiment, is opposite to the classical northern blotting analysis where a specific probe is labelled and hybridized to immobilized and separated RNA molecules; mRNA from a given cell line or tissue is fluorescently labeled (target), and hybridized to thousands of DNA sequences (probes), immobilized on a solid surface in an ordered array format (11).
The most commonly used microarray systems can be divided into two groups, according to the material present on the array (probes): complementary DNA (cDNA) and oligonucleotide arrays. Probes for cDNA arrays are usually PCR products generated from cDNA libraries or clone collections, using vector-specific primers. The length of the probes can vary from 200-2000 base pairs. They are printed on a solid surface (eg. glass slide) at defined locations, using robotics spotting. The spot diameter can be between 75 and 200 μm and the number of spots can vary from a few hundred to 45,000. For oligonucleotide arrays, a collection of oligonucleotide probes (usually 20- to 80-mer) are synthesized either on-chip or prefabricated by conventional synthesis followed by direct printing on the platform. The photolithographic technique using photolabile protecting groups utilized by Affymetrix is an example of on-chip synthesis. Ink-jet technology is another method of on-chip synthesis used by Agilent Technologies. Because of high costs of commercially available, in situ-synthesized oligonucleotide arrays and synthetic oligonucleotide libraries, initially cDNA array technology was the most frequently used technique in academic labs. However, with prices falling over time, spotted long-oligonucleotide arrays and commercial arrays have become the preferred choice in recent years.

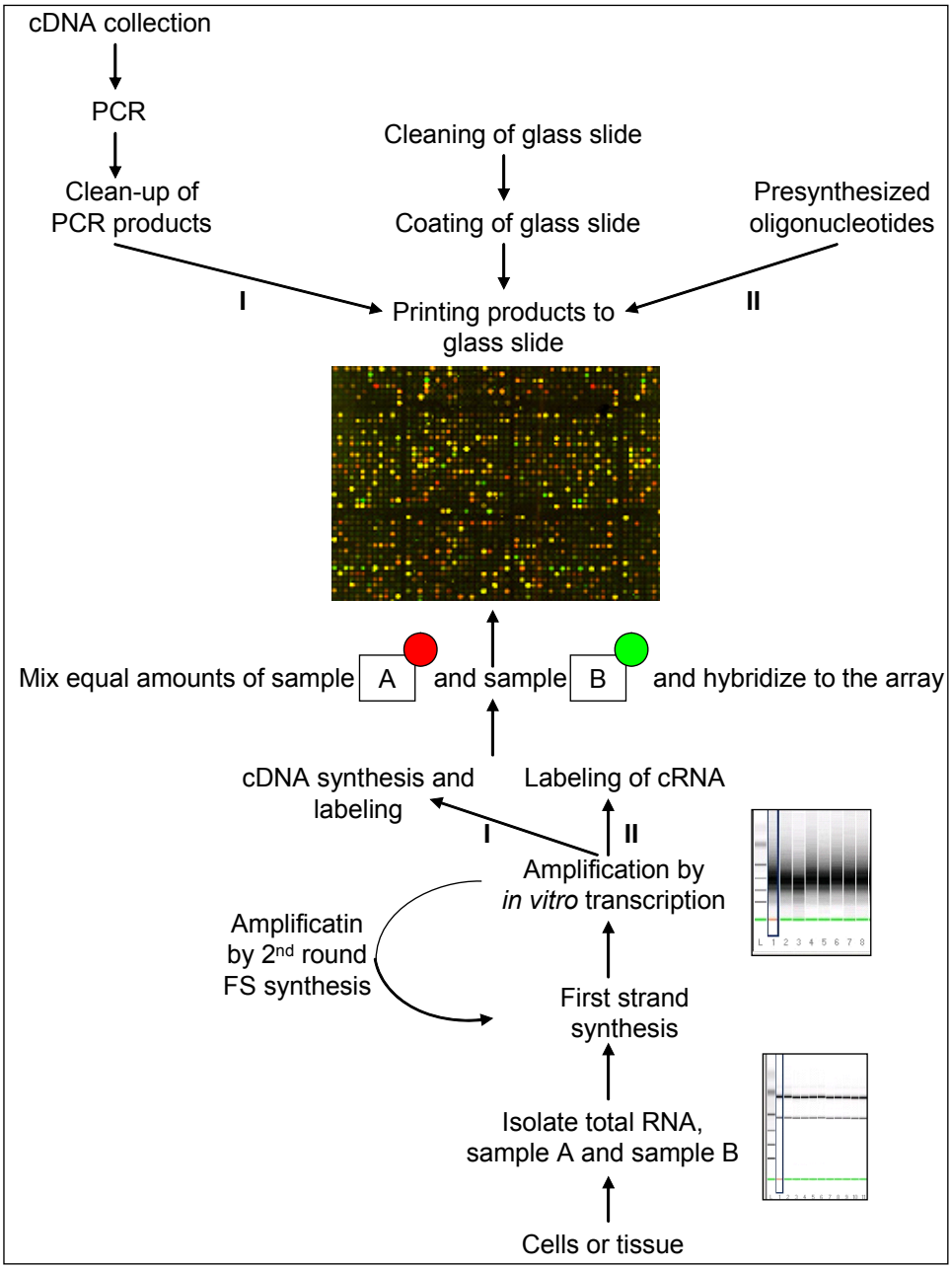

Figure 5a. Preparation of probes and targets for microarray expression studies as used for the experiments described in chapter 2-4 (**I**) and 3,5 (**II**)

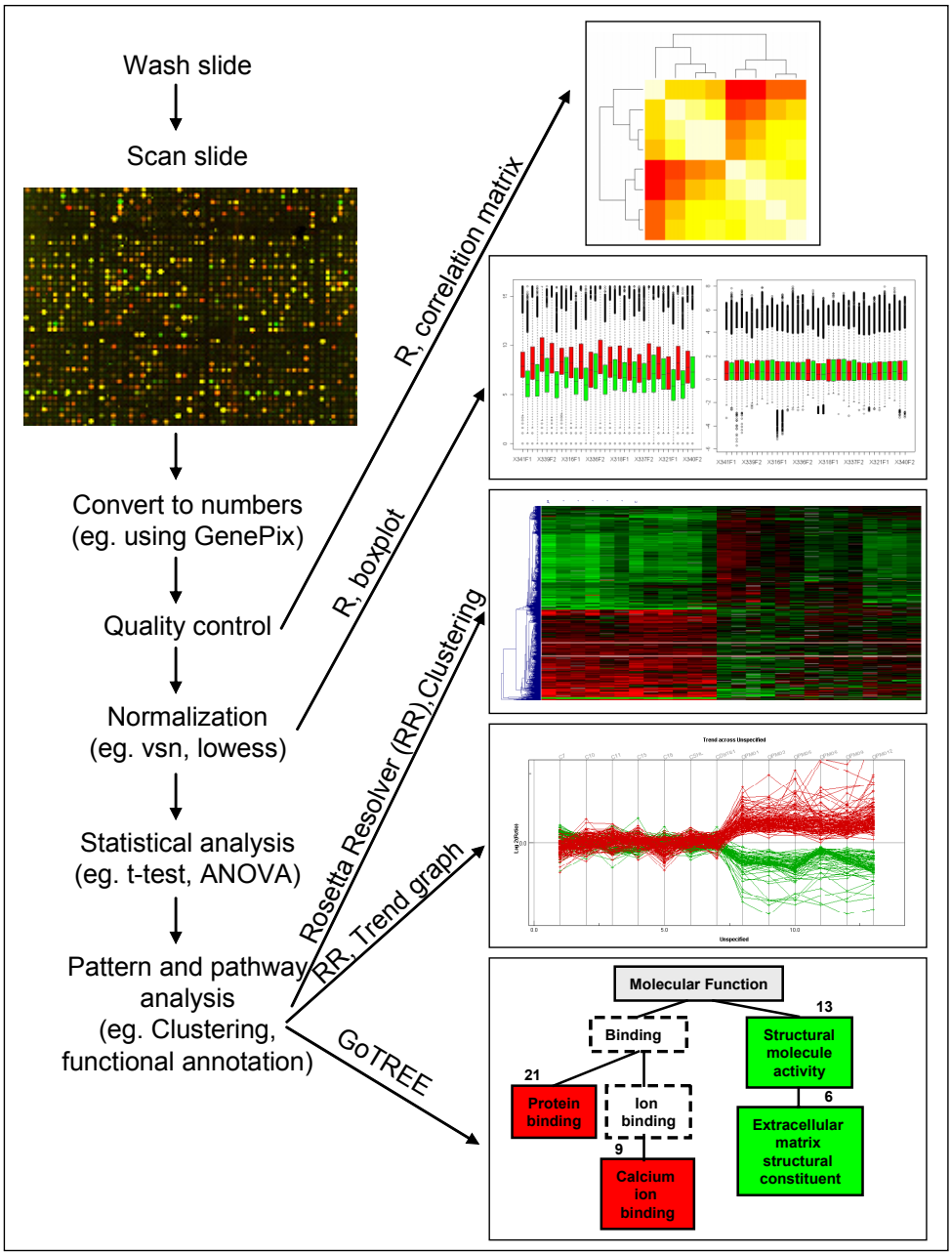

Figure 5b. Data analysis steps for microarray expression studies as used for the experiments described in Chapter 2-5.

1.4.2 Target preparation

In microarray expression profiling, the target is the mRNA extracted from cells or the tissue under study which will be hybridized to the probes on the array. The RNA can directly be labeled with fluorescent dyes in a chemical reaction (eg. ULS (146)) or is reverse-transcribed into cDNA, which is labeled by fluorescent dyes, such as Cy3 and Cy5. When the amount of starting RNA is limited, an amplification of the RNA is possible through in vitro transcription (147,148) or PCR based methods (149) and subsequent labeling of the resulting products (150). The labeled target is hybridized to the DNA probes on the array, and after stringent washing, the spots containing fluorescent dye are detected by laser scanning. mRNA from two different cell populations or tissues, labeled with two different fluorescent dyes, can be hybridized to the same array. When the slide is scanned using two different wavelengths, corresponding to the dyes used, the intensity of the same spot in both channels can be compared. This results in a measurement of the ratio of transcript levels for each gene represented on the array. Figure 5a gives an overview of the preparation of a two-color microarray experiment. To be able to compare the expression levels of many different samples, a large batch of the same reference RNA can be used. A reference sample should provide a signal in as many spots as possible to facilitate the calculation of ratios. We designed a common reference for cDNA arrays, based on PCR amplification and labeling of the probes present on the array (**Chapter 2**). On Affymetrix Gene chips only a single sample is hybridized but it does allow comparison of signals generated by samples hybridized to separate arrays.

1.4.3 Experimental design

Before starting a microarray experiment a good experimental design is crucial. Different experimental designs have been discussed by Churchill (151). Mainly there are two different types of design. The first is the reference design in which all samples are hybridized to a single common reference sample (Figure 6a). With a reference sample, it is relatively easy to statistically analyze the results and it is possible to account for variation between arrays. Also, an additional sample can easily be added to the design in later stages as long as the reference sample is available. A drawback however is the use of twice as many arrays due to the reference sample occupying the second channel, with the consequence of rising costs. Furthermore, a reference sample is finite in shelf-life and amount. The second hybridization design is the loop or cross-over design (Figure 6b). Less hybridizations are neccesary, but if ratios are taken into account, statistical analysis will be much more complicated. When only intensities are considered, analysis is straightforward, but results are less reliable because of the heterogeneity between different arrays. Irrespective of whether a loop-design or a reference design is used, results have shown that technical replication (eg. same sample on two or more different arrays) markedly reduces the number of false positives. In addition, the efficiency of incorporation of nucleotides labeled with different fluorescent dyes during preparation of the target may not be equal, and the fluorescent properties of the different dyes differ. Reciprocal labeling with swapped colors is recommended to correct for this bias. Another important issue is the sample size, which should be large enough to account for biological and technical variation (152). In practice however, some experimental (human) tissue samples are difficult to obtain and costs are high, meaning that not all requirements can be fulfilled and concessions have to be made towards the number of technical and biological replicates that one can afford to analyse.

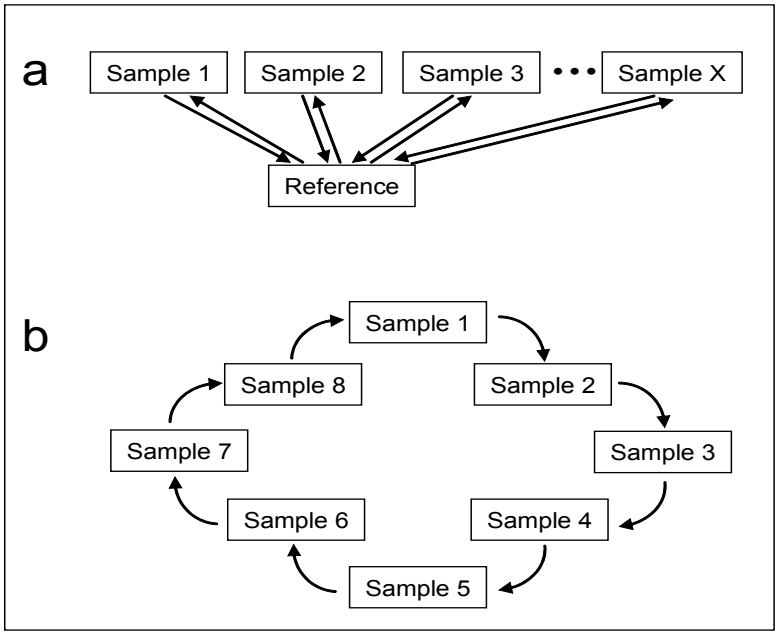

Figure 6. Different hybridization designs for microarray experiments.
a) Reference design with the same reference in each hybridization (boxes represent RNA sample and double arrow is a dye-swap hybridization).
b) Loop design, each sample is hybridized two times in a loop (boxes repre sent RNA samles and single arrow is one hybridization).

1.4.4 Statistical analysis

Microarray results typically consist of long lists of spot intensities, background intensities, ratios and standard deviations. The challenge is to identify whether or not a gene is detected at a given spot and to determine if it is differentially expressed with as few false positives as possible. Before the 'large-scale-gene-expression-era', only a few genes were studied in multiple samples, now thousands of genes in only a small number of samples are studied in a single experiment. Problems arose with normal statistical methods and new analytical methods had to be developed including multiple test correction. Apart from the hundreds of papers on the technique itself, a substantial number of papers on statistical analysis of microaray data were published in the past ten years. To date, there is no consensus on how to statistically analyze microarray data, and thus array results are analyzed by a variety of approaches. Stepwize, the following analyses have to be done and for each, tens to hundreds of solutions exist. Quality control, filtering, normalization (153), determine differentially expressed genes (including multiple test correction) (154), find co-regulated genes (clustering) (155) and finally identify patterns and pathways (156). Figure 5b gives an overview of the steps involved after scanning of the microarray slide, including examples of different methods that can be used in these steps.

Because of the complexity of the data sets generated by microarray experiments, the use of powerful computers for data-analysis is essential. Since the early days of microarraying, R,

a language and environment for statistical computing, has been used to analyze microarray data (http://lib.stat.cmu.edu/R/CRAN/). Bioconductor is an open source, user-based webpage providing scripts necessary for performing the different statistical tests in R (157). Lately, commercial data-analysis tools have been developed and proven their usefulness (for example, Rosetta Resolver, GeneSpring, Spotfire).

1.4.5 Pathway analysis

Making sense out of gene-lists containing hundreds (to thousands) of interesting genes which appear to be differentially expressed in the experiment is the next hurdle to take. It would be easy if every gene had one unique identifier and all functional information was readily available and could be linked to other databases like PubMed or SWISS-PROT. Although bioinformatics is evolving rapidly, it is still not possible to summarize all the available information on a large scale into something that makes sense.
The Gene Ontology Consortium has followed one route to get more meaningful data of long gene lists by producing a database in which genes are placed in functional categories on the level of molecular function, biological process and cellular location (158,159). Several pathway analysis programs use this database and are well suited for a first global screening providing statistical tests to determine if a certain functional group is over- or under represented in your gene-list. However, problems occur after (or during) this initial functional screening. In some databases the data are retrieved based on the gene name. A significant difficulty is that huge homonym and synonym problems exists in gene names (one gene abbreviation can refer to different genes and one gene can have multiple abbreviations)(160). Thus, unintentionally different functions can be mixed. Furthermore, functional annotation is incomplete for most genes of interest. To overcome this, hundreds of papers may have to be read to verify or explore the function of certain proteins. This number exponentially increases with the number of differentially expressed genes in the list. Moreover, when a new, not annotated pathway is involved, it is difficult or impossible to uncover this with the existing pathway databases. New tools are currently being developed which align genes not on their sequences but on the textual information that is available in genomics databases and the biomedical literature (161). This so-called biosemantics approach intelligently summarizes literature and other web-published, human annotated information on a large scale, by text mining and meta-analysis. The power of this analysis is that two genes can have no overlap in papers (i.e. no co-occurrence in one paper), but have an almost identical fingerprint (basically a list of key words referring to the gene, ranked in order of importance) and therefore may have a relationship and are interesting to follow up (discovery of implicit knowledge). In this way, new hypotheses on the biological processes involved in a certain experimental situation can be generated and tested.

1.4.6 Confirmation of gene expression data and correlation to protein levels

When data analysis is completed and results are interpretable, independent confirmation of the results is a prerequisite, especially when a conclusion is based on individual genes, instead of a whole pathway which is up- or downregulated. A direct confirmation of the microarray expression results can be made using quantitative RT-PCR. Although often used, it is difficult to multiplex this method and thus costly. In **Chapter 4** RT-MLPA is used to confirm microarray results. This is a new technique which allows the rapid and simultaneous

quantification of up to 40 transcripts in multiple samples in a single assay (162). The best confirmation can be made by using an independent set of new samples which show the same result in any of the discussed methods.
In the end, gene expression level changes in combination with knowledge on the protein level will provide a more complete picture of the experimental situation. Western blotting and immunohistochemical staining are used to check protein levels of differentially expressed genes, extrapolating the results to the protein level. For large-scale analysis of differential protein levels, protein microarrays can be used.

1.4.7 Application

Gene expression microarrays are used in many different applications. Because tumor samples are relatively easy to obtain, the most progress has been made in cancer research. It has been used for diagnostics (163,164), prognostic predictions (165,166) and to assist the choice for therapeutic intervention (167,168). Other applications are in the analysis of complex diseases (169), discovery of gene functions (170), biological pathway dissection (171) and biomarker identification (172,173). Also in pharmacogenomics it is a widely used tool for drug-target validation (174,175) and in toxicogenomics it is used in compound toxicity studies (176,177).

1.5 Outline of this thesis

With large-scale gene expression profiling different questions can be answered. For this thesis, gene expression microarrays have been applied to the research of muscle cell differentiation and cellular pathologies in two muscular disorders, Duchenne muscular dystrophy and Oculopharyngeal muscular dystrophy.
Before starting, the technique was optimized and a common reference for two-color cDNA gene expression experiments was developed (**Chapter 2**). The reference sample represents all of the genes spotted on the array and makes uncomplicated ratio data-analysis possible.
Chapter 3 describes a study on gene expression changes during myoblast differentiation. Here, primary myogenic cell cultures have been used as a model system and cell fusion is monitored in time by isolating RNA at different time points. On basis of these results and results of other studies, the relation between myogenesis and muscle regeneration is discussed. In Duchenne Muscular Dystrophy, the regeneration process is hampered and in **Chapter 4** this process is studied in a cellular model using primary myogenic cell cultures of DMD patients. The results show differences in gene expression pattern with the healthy situation which points to a possible cause of the inefficient regeneration in DMD patients. Interference in this affected pathway could be useful in treatment of the disease.
In **Chapter 5** oligonucleotide arrays have been used to study gene expression in a cellular model for Oculopharyngeal muscular dystrophy. The results show that as a result of intranuclear inclusion formation, a functional pathway, which involves collagen synthesis and extracellular matrix components, is clearly affected. This is the first time this pathway is associated with the disease and it provides a better understanding of the molecular mechanisms involved in OPMD.
Finally, the results presented in the Chapters 2-5 are discussed in **Chapter 6** with future perspectives.

Reference List

1. Gygi,S.P., Rochon,Y., Franza,B.R., and Aebersold,R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol.Cell Biol.*, **19**, 1720-1730.
2. Lee,P.S., Shaw,L.B., Choe,L.H., Mehra,A., Hatzimanikatis,V., and Lee,K.H. (2003) Insights into the relation between mrna and protein expression patterns: II. Experimental observations in Escherichia coli. *Biotechnol.Bioeng.*, **84**, 834-841.
3. Griffin,T.J., Gygi,S.P., Ideker,T., Rist,B., Eng,J., Hood,L., and Aebersold,R. (2002) Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. *Mol.Cell Proteomics.*, **1**, 323-333.
4. Alwine,J.C., Kemp,D.J., and Stark,G.R. (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc.Natl.Acad.Sci.U.S.A*, **74**, 5350-5354.
5. Bustin,S.A. (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J.Mol.Endocrinol.*, **29**, 23-39.
6. Liang,P., Pardee,A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**, 967-971.
7. Velculescu,V.E., Zhang,L., Vogelstein,B., and Kinzler,K.W. (1995) Serial analysis of gene expression. *Science*, **270**, 484-487.
8. Prediger,E.A. (2001) Detection and quantitation of mRNAs using ribonuclease protection assays. *Methods Mol.Biol.*, **160**, 495-505.
9. Southern,E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J.Mol.Biol.*, **98**, 503-517.
10. Pease,A.C., Solas,D., Sullivan,E.J., Cronin,M.T., Holmes,C.P., and Fodor,S.P. (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc.Natl.Acad.Sci. U.S.A*, **91**, 5022-5026.
11. Schena,M., Shalon,D., Davis,R.W., and Brown,P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467-470.
12. Maroto,M., Reshef,R., Munsterberg,A.E., Koester,S., Goulding,M., and Lassar,A.B. (1997) Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell*, **89**, 139-148.
13. Tajbakhsh,S., Borello,U., Vivarelli,E., Kelly,R., Papkoff,J., Duprez,D., Buckingham,M., and Cossu,G. (1998) Differential activation of Myf5 and MyoD by different Wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of Myf5. *Development*, **125**, 4155-4162.
14. Buckingham,M. (2001) Skeletal muscle formation in vertebrates. *Curr.Opin.Genet.Dev.*, **11**, 440-448.
15. Pownall,M.E., Gustafsson,M.K., and Emerson,C.P., Jr. (2002) Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu.Rev.Cell Dev.Biol.*, **18**, 747-783.
16. Gustafsson,M.K., Pan,H., Pinney,D.F., Liu,Y., Lewandowski,A., Epstein,D.J., and Emerson,C.P., Jr. (2002) Myf5 is a direct target of long-range Shh signaling and Gli regulation for muscle specification. *Genes Dev.*, **16**, 114-126.
17. Rudnicki,M.A., Schnegelsberg,P.N., Stead,R.H., Braun,T., Arnold,H.H., and Jaenisch,R. (1993) MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell*, **75**, 1351-1359.
18. Hasty,P., Bradley,A., Morris,J.H., Edmondson,D.G., Venuti,J.M., Olson,E.N., and Klein,W. H. (1993) Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature*, **364**, 501-506.
19. Nabeshima,Y., Hanaoka,K., Hayasaka,M., Esumi,E., Li,S., Nonaka,I., and Nabeshima,Y.

(1993) Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature*, **364**, 532-535.

20. Zhu,Z., Miller,J.B. (1997) MRF4 can substitute for myogenin during early stages of myogenesis. *Dev.Dyn.*, **209**, 233-241.
21. Kassar-Duchossoy,L., Gayraud-Morel,B., Gomes,D., Rocancourt,D., Buckingham,M., Shinin,V., and Tajbakhsh,S. (2004) Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature*, **431**, 466-471.
22. Blackwell,T.K., Weintraub,H. (1990) Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science*, **250**, 1104-1110.
23. Lassar,A.B., Davis,R.L., Wright,W.E., Kadesch,T., Murre,C., Voronova,A., Baltimore,D., and Weintraub,H. (1991) Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. *Cell*, **66**, 305-315.
24. Black,B.L., Olson,E.N. (1998) Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu.Rev.Cell Dev.Biol.*, **14**, 167-196.
25. Fickett,J.W. (1996) Coordinate positioning of MEF2 and myogenin binding sites. *Gene*, **172**, GC19-GC32.
26. Benezra,R., Davis,R.L., Lockshon,D., Turner,D.L., and Weintraub,H. (1990) The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell*, **61**, 49-59.
27. Arnold,H.H., Winter,B. (1998) Muscle differentiation: more complexity to the network of myogenic regulators. *Curr.Opin.Genet.Dev.*, **8**, 539-544.
28. Martin,J.F., Li,L., and Olson,E.N. (1992) Repression of myogenin function by TGF-beta 1 is targeted at the basic helix-loop-helix motif and is independent of E2A products. *J.Biol. Chem.*, **267**, 10956-10960.
29. Seale,P., Sabourin,L.A., Girgis-Gabardo,A., Mansouri,A., Gruss,P., and Rudnicki,M.A. (2000) Pax7 is required for the specification of myogenic satellite cells. *Cell*, **102**, 777-786.
30. Mauro,A. (1961) Satellite cell of skeletal muscle fibers. *J.Biophys.Biochem.Cytol.*, **9**, 493-495.
31. Anderson,J.E. (1998) Murray L. Barr Award Lecture. Studies of the dynamics of skeletal muscle regeneration: the mouse came back! *Biochem.Cell Biol.*, **76**, 13-26.
32. Artavanis-Tsakonas,S., Matsuno,K., and Fortini,M.E. (1995) Notch signaling. *Science*, **268**, 225-232.
33. Conboy,I.M., Rando,T.A. (2002) The regulation of Notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis. *Dev.Cell*, **3**, 397-409.
34. Tatsumi,R., Anderson,J.E., Nevoret,C.J., Halevy,O., and Allen,R.E. (1998) HGF/SF is present in normal adult skeletal muscle and is capable of activating satellite cells. *Dev.Biol.*, **194**, 114-128.
35. Machida,S., Booth,F.W. (2004) Insulin-like growth factor 1 and muscle growth: implication for satellite cell proliferation. *Proc.Nutr.Soc.*, **63**, 337-340.
36. Yablonka-Reuveni,Z., Seger,R., and Rivera,A.J. (1999) Fibroblast growth factor promotes recruitment of skeletal muscle satellite cells in young and old rats. *J.Histochem.Cytochem.*, **47**, 23-42.
37. Carrino,D.A. (1998) Dynamic expression of proteoglycans during skeletal muscle development. *Basic Appl.Myol.*, **8**, 95-106.
38. Brandan,E., Larrain,J. (1998) Heparan sulfate proteoglycans during terminal skeletal muscle cell differentiation: possible functions and regulations of their expression. *Basic Appl.Myol.*, **8**, 107-113.
39. Velleman,S.G., Mcfarland,D.C., and Coy,C.S. (1998) Skeletal muscle satellite cells: identification of a heparan sulfate proteoglycan. *Basic Appl.Myol.*, **8**, 115-120.

40. Yablonka-Reuveni,Z., Rivera,A.J. (1994) Temporal expression of regulatory and structural muscle proteins during myogenesis of satellite cells on isolated adult rat fibers. *Dev.Biol.*, **164**, 588-603.
41. Cornelison,D.D., Wold,B.J. (1997) Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. *Dev.Biol.*, **191**, 270-283.
42. Megeney,L.A., Kablar,B., Garrett,K., Anderson,J.E., and Rudnicki,M.A. (1996) MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev.*, **10**, 1173-1183.
43. Yan,Z., Choi,S., Liu,X., Zhang,M., Schageman,J.J., Lee,S.Y., Hart,R., Lin,L., Thurmond,F. A., and Williams,R.S. (2003) Highly coordinated gene regulation in mouse skeletal muscle regeneration. *J.Biol.Chem.*, **278**, 8826-8836.
44. Bischoff,R., Holtzer,H. (1969) Mitosis and the processes of differentiation of myogenic cells in vitro. *J.Cell Biol.*, **41**, 188-200.
45. Andres,V., Walsh,K. (1996) Myogenin expression, cell cycle withdrawal, and phenotypic differentiation are temporally separable events that precede cell fusion upon myogenesis. *J.Cell Biol.*, **132**, 657-666.
46. Weintraub,H., Tapscott,S.J., Davis,R.L., Thayer,M.J., Adam,M.A., Lassar,A.B., and Miller,A.D. (1989) Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc.Natl.Acad.Sci.U.S.A*, **86**, 5434-5438.
47. Roest,P.A., van der Tuijn,A.C., Ginjaar,H.B., Hoeben,R.C., Hoger-Vorst,F.B., Bakker,E., den Dunnen,J.T., and van Ommen,G.J. (1996) Application of in vitro Myo-differentiation of non-muscle cells to enhance gene expression and facilitate analysis of muscle proteins. *Neuromuscul.Disord.*, **6**, 195-202.
48. Emery,A., Muntoni,F. (2003) *Duchenne Muscular Dystrophy*. Oxford university press.
49. Meryon,E. (1852) On granular and fatty degeneration of the voluntary muscles. *Medico-Chirurgical Transactions (London)*, **35**, 73-84.
50. Duchenne,G. (1868) Recherches sur la paralysie musculaire pseudohypertrophique ou paralysie myo-sclérosique. *Archives Générales de Médecine*, **11**, 5-25.
51. Erb,W.H. (1891) Dystrophia muscularis progressiva: Klinische und pathologisch-anatomische Studien. *Dtsch Z Nervenheilk*, **1**, 173.
52. Ebashi,S., Toyokura,Y., Momoi,H., and Sugita,H. (1959) High creatine phospokinase activity of sera of progressive muscular dystrophy. *Journal of Biochemistry*, **46**, 103-104.
53. Jacobs,P.A., Hunt,P.A., Mayer,M., and Bart,R.D. (1981) Duchenne muscular dystrophy (DMD) in a female with an X/autosome translocation: further evidence that the DMD locus is at Xp21. *Am.J.Hum.Genet.*, **33**, 513-518.
54. Murray,J.M., Davies,K.E., Harper,P.S., Meredith,L., Mueller,C.R., and Williamson,R. (1982) Linkage relationship of a cloned DNA sequence on the short arm of the X chromosome to Duchenne muscular dystrophy. *Nature*, **300**, 69-71.
55. Davies,K.E., Pearson,P.L., Harper,P.S., Murray,J.M., O'Brien,T., Sarfarazi,M., and Williamson,R. (1983) Linkage analysis of two cloned DNA sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome. *Nucleic Acids Res.*, **11**, 2303-2312.
56. Francke,U., Ochs,H.D., de Martinville,B., Giacalone,J., Lindgren,V., Disteche,C., Pagon,R. A., Hofker,M.H., van Ommen,G.J., Pearson,P.L., and . (1985) Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy, chronic granulomatous disease, retinitis pigmentosa, and McLeod syndrome. *Am.J.Hum.Genet.*, **37**, 250-267.
57. Ray,P.N., Belfall,B., Duff,C., Logan,C., Kean,V., Thompson,M.W., Sylvester,J.E., Gorski,J. L., Schmickel,R.D., and Worton,R.G. (1985) Cloning of the breakpoint of an X;21

translocation associated with Duchenne muscular dystrophy. *Nature*, **318**, 672-675.

58. Kunkel,L.M., Hejtmancik,J.F., Caskey,C.T., Speer,A., Monaco,A.P., Middlesworth,W., Colletti,C.A., Bertelson,C., Muller,U., Bresnan,M., Shapiro,F., Tantravahi,U., Speer,J., Latt,S.A., Bartlett,R., Pericak-Vance,M.A., Roses,A.D., Thompson,M.W., Ray,P.N., Worton,R.G., Fischbeck,K.H., Gallano,P., Coulon,M., Duros,C., Boue,J., Junien,C., Chelly,J., Hamard,G., Jeanpierre,M., Lambert,M., Kaplan,J.C., Emery,A., Dorkins,H., McGlade,S., Davies,K.E., Boehm,C., Arveiler,B., Lemaire,C., Morgan,G.J., Denton,M. J., Amos,J., Bobrow,M., Benham,F., Boswinkel,E., Cole,C., Dubowitz,V., Hart,K., Hodgson,S., Johnson,L., Walker,A., Roncuzzi,L., Ferlini,A., Nobile,C., Romeo,G., Wilcox,D.E., Affara,N.A., Ferguson-Smith,M.A., Lindolf,M., Kaariainen,H., de la,C. A., Ionasescu,V., Searby,C., Ionasescu,R., Bakker,E., van Ommen,G.J., Pearson,P.L., Greenberg,C.R., Hamerton,J.L., Wrogemann,K., Doherty,R.A., Polakowska,R., Hyser,C., Quirk,S., Thomas,N., Harper,J.F., Darras,B.T., and Francke,U. (1986) Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. *Nature*, **322**, 73-77.
59. van Ommen,G.J., Verkerk,J.M., Hofker,M.H., Monaco,A.P., Kunkel,L.M., Ray,P., Worton,R., Wieringa,B., Bakker,E., and Pearson,P.L. (1986) A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X-chromosome. *Cell*, **47**, 499-504.
60. Monaco,A.P., Neve,R.L., Colletti-Feener,C., Bertelson,C.J., Kurnit,D.M., and Kunkel,L. M. (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature*, **323**, 646-650.
61. Aartsma-Rus,A., Van Deutekom,J.C., Fokkema,I.F., van Ommen,G.J., and den Dunnen,J.T. (2006) Entries in the Leiden Duchenne muscular dystrophy mutation database: An overview of mutation types and paradoxical cases that confirm the reading-frame rule. *Muscle Nerve*.
62. Kingston,H.M., Sarfarazi,M., Thomas,N.S., and Harper,P.S. (1984) Localisation of the Becker muscular dystrophy gene on the short arm of the X chromosome by linkage to cloned DNA sequences. *Hum.Genet.*, **67**, 6-17.
63. Monaco,A.P., Bertelson,C.J., Liechti-Gallati,S., Moser,H., and Kunkel,L.M. (1988) An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics*, **2**, 90-95.
64. Muntoni,F., Gobbi,P., Sewry,C., Sherratt,T., Taylor,J., Sandhu,S.K., Abbs,S., Roberts,R., Hodgson,S.V., Bobrow,M., and . (1994) Deletions in the 5' region of dystrophin and resulting phenotypes. *J.Med.Genet.*, **31**, 843-847.
65. Nevo,Y., Muntoni,F., Sewry,C., Legum,C., Kutai,M., Harel,S., and Dubowitz,V. (2003) Large in-frame deletions of the rod-shaped domain of the dystrophin gene resulting in severe phenotype. *Isr.Med.Assoc.J.*, **5**, 94-97.
66. Hoffman,E.P., Brown,R.H., Jr., and Kunkel,L.M. (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell*, **51**, 919-928.
67. Byers,T.J., Lidov,H.G., and Kunkel,L.M. (1993) An alternative dystrophin transcript specific to peripheral nerve. *Nat.Genet.*, **4**, 77-81.
68. Nudel,U., Zuk,D., Einat,P., Zeelon,E., Levy,Z., Neuman,S., and Yaffe,D. (1989) Duchenne muscular dystrophy gene product is not identical in muscle and brain. *Nature*, **337**, 76-78.
69. Gorecki,D.C., Monaco,A.P., Derry,J.M., Walker,A.P., Barnard,E.A., and Barnard,P.J. (1992) Expression of four alternative dystrophin transcripts in brain regions regulated by different promoters. *Hum.Mol.Genet.*, **1**, 505-510.
70. Pillers,D.A., Bulman,D.E., Weleber,R.G., Sigesmund,D.A., Musarella,M.A., Powell,B.R., Murphey,W.H., Westall,C., Panton,C., Becker,L.E., and . (1993) Dystrophin expression in the human retina is required for normal function as defined by electroretinography. *Nat. Genet.*, **4**, 82-86.
71. Lidov,H.G., Selig,S., and Kunkel,L.M. (1995) Dp140: a novel 140 kDa CNS transcript from

the dystrophin locus. *Hum.Mol.Genet.*, **4**, 329-335.

72. Bar,S., Barnea,E., Levy,Z., Neuman,S., Yaffe,D., and Nudel,U. (1990) A novel product of the Duchenne muscular dystrophy gene which greatly differs from the known isoforms in its structure and tissue distribution. *Biochem.J.*, **272**, 557-560.
73. Zubrzycka-Gaarn,E.E., Bulman,D.E., Karpati,G., Burghes,A.H., Belfall,B., Klamut,H. J., Talbot,J., Hodges,R.S., Ray,P.N., and Worton,R.G. (1988) The Duchenne muscular dystrophy gene product is localized in sarcolemma of human skeletal muscle. *Nature*, **333**, 466-469.
74. Campbell,K.P., Kahl,S.D. (1989) Association of dystrophin and an integral membrane glycoprotein. *Nature*, **338**, 259-262.
75. Ervasti,J.M., Campbell,K.P. (1991) Membrane organization of the dystrophin-glycoprotein complex. *Cell*, **66**, 1121-1131.
76. Rando,T.A. (2001) The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. *Muscle Nerve*, **24**, 1575-1594.
77. Michalak,M., Opas,M. (1997) Functions of dystrophin and dystrophin associated proteins. *Curr.Opin.Neurol.*, **10**, 436-442.
78. Straub,V., Campbell,K.P. (1997) Muscular dystrophies and the dystrophin-glycoprotein complex. *Curr.Opin.Neurol.*, **10**, 168-175.
79. Petrof,B.J. (2002) Molecular pathophysiology of myofiber injury in deficiencies of the dystrophin-glycoprotein complex. *Am.J.Phys.Med.Rehabil.*, **81**, S162-S174.
80. Carpenter,S., Karpati,G. (1979) Duchenne muscular dystrophy: plasma membrane loss initiates muscle cell necrosis unless it is repaired. *Brain*, **102**, 147-161.
81. Ervasti,J.M., Ohlendieck,K., Kahl,S.D., Gaver,M.G., and Campbell,K.P. (1990) Deficiency of a glycoprotein component of the dystrophin complex in dystrophic muscle. *Nature*, **345**, 315-319.
82. Ohlendieck,K., Campbell,K.P. (1991) Dystrophin-associated proteins are greatly reduced in skeletal muscle from mdx mice. *J.Cell Biol.*, **115**, 1685-1694.
83. Aartsma-Rus, A. Development of an antisense-mediated exon skipping therapy for duchenne muscular dystrophy. 18. 2005.
    Ref Type: Thesis/Dissertation
84. Rubin,H. (1997) Cell aging in vivo and in vitro. *Mech.Ageing Dev.*, **98**, 1-35.
85. Allsopp,R.C., Chang,E., Kashefi-Aazam,M., Rogaev,E.I., Piatyszek,M.A., Shay,J.W., and Harley,C.B. (1995) Telomere shortening is associated with cell division in vitro and in vivo. *Exp.Cell Res.*, **220**, 194-200.
86. de Haan,J.B., Cristiano,F., Iannello,R., Bladier,C., Kelner,M.J., and Kola,I. (1996) Elevation in the ratio of Cu/Zn-superoxide dismutase to glutathione peroxidase activity induces features of cellular senescence and this effect is mediated by hydrogen peroxide. *Hum.Mol. Genet.*, **5**, 283-292.
87. Blau,H.M., Webster,C., and Pavlath,G.K. (1983) Defective myoblasts identified in Duchenne muscular dystrophy. *Proc.Natl.Acad.Sci.U.S.A*, **80**, 4856-4860.
88. Heslop,L., Morgan,J.E., and Partridge,T.A. (2000) Evidence for a myogenic stem cell that is exhausted in dystrophic muscle. *J.Cell Sci.*, **113**, 2299-2308.
89. Wakayama,Y. (1976) Electron microscopic study on the satellite cell in the muscle of Duchenne muscular dystrophy. *J.Neuropathol.Exp.Neurol.*, **35**, 532-540.
90. Ishimoto,S., Goto,I., Ohta,M., and Kuroiwa,Y. (1983) A quantitative study of the muscle satellite cells in various neuromuscular disorders. *J.Neurol.Sci.*, **62**, 303-314.
91. Maier,F., Bornemann,A. (1999) Comparison of the muscle fiber diameter and satellite cell frequency in human muscle biopsies. *Muscle Nerve*, **22**, 578-583.
92. Oexle,K., Zwirner,A., Freudenberg,K., Kohlschutter,A., and Speer,A. (1997) Examination

of telomere lengths in muscle tissue casts doubt on replicative aging as cause of progression in Duchenne muscular dystrophy. *Pediatr.Res.*, **42**, 226-231.

93. Oexle,K., Kohlschutter,A. (2001) Cause of progression in Duchenne muscular dystrophy: impaired differentiation more probable than replicative aging. *Neuropediatrics*, **32**, 123-129.
94. Decary,S., Hamida,C.B., Mouly,V., Barbet,J.P., Hentati,F., and Butler-Browne,G.S. (2000) Shorter telomeres in dystrophic muscle consistent with extensive regeneration in young children. *Neuromuscul.Disord.*, **10**, 113-120.
95. Burr,I.M., Asayama,K., and Fenichel,G.M. (1987) Superoxide dismutases, glutathione peroxidase, and catalase in neuromuscular disease. *Muscle Nerve*, **10**, 150-154.
96. Melone,M.A., Peluso,G., Petillo,O., Galderisi,U., and Cotrufo,R. (1999) Defective growth in vitro of Duchenne Muscular Dystrophy myoblasts: the molecular and biochemical basis. *J.Cell Biochem.*, **76**, 118-132.
97. Delaporte,C., Dehaupas,M., and Fardeau,M. (1984) Comparison between the growth pattern of cell cultures from normal and Duchenne dystrophy muscle. *J.Neurol.Sci.*, **64**, 149-160.
98. Jasmin,G., Tautu,C., Vanasse,M., Brochu,P., and Simoneau,R. (1984) Impaired muscle differentiation in explant cultures of Duchenne muscular dystrophy. *Lab Invest*, **50**, 197-207.
99. Schmalbruch,H. (1984) Regenerated muscle fibers in Duchenne muscular dystrophy: a serial section study. *Neurology*, **34**, 60-65.
100. Arahata,K., Engel,A.G. (1984) Monoclonal antibody analysis of mononuclear cells in myopathies. I: Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. *Ann.Neurol.*, **16**, 193-208.
101. Merly,F., Lescaudron,L., Rouaud,T., Crossin,F., and Gardahaut,M.F. (1999) Macrophages enhance muscle satellite cell proliferation and delay their differentiation. *Muscle Nerve*, **22**, 724-732.
102. Taylor,E.W. (1915) Progressive vagus-glossopharyngeal paralysis with ptosis: a contribution to the group of family diseases. *J.Nerv.Ment.Dis.*, **42**, 129-139.
103. Victor,M., Hayes,R., and Adams,R.D. (1962) Oculopharyngeal muscular dystrophy. A familial disease of late life characterized by dysphagia and progressive ptosis of the evelids. *N.Engl.J.Med.*, **267**, 1267-1272.
104. Bouchard,J.P. (1997) Andre Barbeau and the oculopharyngeal muscular dystrophy in French Canada and North America. *Neuromuscul.Disord.*, **7 Suppl 1**, S5-11.
105. Becher,M.W., Morrison,L., Davis,L.E., Maki,W.C., King,M.K., Bicknell,J.M., Reinert,B.L., Bartolo,C., and Bear,D.G. (2001) Oculopharyngeal muscular dystrophy in Hispanic New Mexicans. *JAMA*, **286**, 2437-2440.
106. Blumen,S.C., Nisipeanu,P., Sadeh,M., Asherov,A., Blumen,N., Wirguin,Y., Khilkevich,O., Carasso,R.L., and Korczyn,A.D. (1997) Epidemiology and inheritance of oculopharyngeal muscular dystrophy in Israel. *Neuromuscul.Disord.*, **7 Suppl 1**, S38-S40.
107. Rodriguez,M., Camejo,C., Bertoni,B., Braida,C., Rodriguez,M.M., Brais,B., Medici,M., and Roche,L. (2005) (GCG)11 founder mutation in the PABPN1 gene of OPMD Uruguayan families. *Neuromuscul.Disord.*, **15**, 185-190.
108. Brais,B., Bouchard,J.P., Xie,Y.G., Rochefort,D.L., Chretien,N., Tome,F.M., Lafreniere,R.G., Rommens,J.M., Uyama,E., Nohira,O., Blumen,S., Korczyn,A.D., Heutink,P., Mathieu,J., Duranceau,A., Codere,F., Fardeau,M., and Rouleau,G.A. (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat.Genet.*, **18**, 164-167.
109. Van Der Sluijs,B.M., Hoefsloot,L.H., Padberg,G.W., Van Der Maarel,S.M., and Van Engelen,B.G. (2003) Oculopharyngeal muscular dystrophy with limb girdle weakness as major complaint. *J.Neurol.*, **250**, 1307-1312.
110. Bouchard,J.P., Brais,B., Brunet,D., Gould,P.V., and Rouleau,G.A. (1997) Recent studies on

oculopharyngeal muscular dystrophy in Quebec. *Neuromuscul.Disord.*, **7 Suppl 1**, S22-S29.

111. Mirabella,M., Silvestri,G., de Rosa,G., Di Giovanni,S., Di Muzio,A., Uncini,A., Tonali,P., and Servidei,S. (2000) GCG genetic expansions in Italian patients with oculopharyngeal muscular dystrophy. *Neurology*, **54**, 608-614.
112. Hill,M.E., Creed,G.A., McMullan,T.F., Tyers,A.G., Hilton-Jones,D., Robinson,D.O., and Hammans,S.R. (2001) Oculopharyngeal muscular dystrophy: phenotypic and genotypic studies in a UK population. *Brain*, **124**, 522-526.
113. Wong,K.T., Dick,D., and Anderson,J.R. (1996) Mitochondrial abnormalities in oculopharyngeal muscular dystrophy. *Neuromuscul.Disord.*, **6**, 163-166.
114. Tome,F.M., Fardeau,M. (1980) Nuclear inclusions in oculopharyngeal dystrophy. *Acta Neuropathol.(Berl)*, **49**, 85-87.
115. Tome,F.M., Chateau,D., Helbling-Leclerc,A., and Fardeau,M. (1997) Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul.Disord.*, **7 Suppl 1**, S63-S69.
116. Calado,A., Tome,F.M., Brais,B., Rouleau,G.A., Kuhn,U., Wahle,E., and Carmo-Fonseca,M. (2000) Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum.Mol.Genet.*, **9**, 2321-2328.
117. Blumen,S.C., Sadeh,M., Korczyn,A.D., Rouche,A., Nisipeanu,P., Asherov,A., and Tome,F. M. (1996) Intranuclear inclusions in oculopharyngeal muscular dystrophy among Bukhara Jews. *Neurology*, **46**, 1324-1328.
118. Schmitt,H.P., Krause,K.H. (1981) An autopsy study of a familial oculopharyngeal muscular dystrophy (OPMD) with distal spread and neurogenic involvement. *Muscle Nerve*, **4**, 296-305.
119. Little,B.W., Perl,D.P. (1982) Oculopharyngeal muscular dystrophy. An autopsied case from the French-Canadian kindred. *J.Neurol.Sci.*, **53**, 145-158.
120. Lotz,B.P., Engel,A.G., Nishino,H., Stevens,J.C., and Litchy,W.J. (1989) Inclusion body myositis. Observations in 40 patients. *Brain*, **112**, 727-747.
121. Pratt,M.F., Meyers,P.K. (1986) Oculopharyngeal muscular dystrophy: recent ultrastructural evidence for mitochondrial abnormalities. *Laryngoscope*, **96**, 368-373.
122. Pauzner,R., Blatt,I., Mouallem,M., Ben David,E., Farfel,Z., and Sadeh,M. (1991) Mitochondrial abnormalities in oculopharyngeal muscular dystrophy. *Muscle Nerve*, **14**, 947-952.
123. Gambelli,S., Malandrini,A., Ginanneschi,F., Berti,G., Cardaioli,E., De Stefano,R., Franci,M., Salvadori,C., Mari,F., Bruttini,M., Rossi,A., Federico,A., and Renieri,A. (2004) Mitochondrial abnormalities in genetically assessed oculopharyngeal muscular dystrophy. *Eur.Neurol.*, **51**, 144-147.
124. Brais,B., Xie,Y.G., Sanson,M., Morgan,K., Weissenbach,J., Korczyn,A.D., Blumen,S. C., Fardeau,M., Tome,F.M., Bouchard,J.P., and . (1995) The oculopharyngeal muscular dystrophy locus maps to the region of the cardiac alpha and beta myosin heavy chain genes on chromosome 14q11.2-q13. *Hum.Mol.Genet.*, **4**, 429-434.
125. Scacheri,P.C., Garcia,C., Hebert,R., and Hoffman,E.P. (1999) Unique PABP2 mutations in "Cajuns" suggest multiple founders of oculopharyngeal muscular dystrophy in populations with French ancestry. *Am.J.Med.Genet.*, **86**, 477-481.
126. Nakamoto,M., Nakano,S., Kawashima,S., Ihara,M., Nishimura,Y., Shinde,A., and Kakizuka,A. (2002) Unequal crossing-over in unique PABP2 mutations in Japanese patients: a possible cause of oculopharyngeal muscular dystrophy. *Arch.Neurol.*, **59**, 474-477.
127. Van Der Sluijs,B.M., Van Engelen,B.G., and Hoefsloot,L.H. (2003) Oculopharyngeal muscular dystrophy (OPMD) due to a small duplication in the PABPN1 gene. *Hum.Mutat.*,

21, 553.

128. Muller,T., Schroder,R., and Zierz,S. (2001) GCG repeats and phenotype in oculopharyngeal muscular dystrophy. *Muscle Nerve*, **24**, 120-122.
129. Blumen,S.C., Brais,B., Korczyn,A.D., Medinsky,S., Chapman,J., Asherov,A., Nisipeanu,P., Codere,F., Bouchard,J.P., Fardeau,M., Tome,F.M., and Rouleau,G.A. (1999) Homozygotes for oculopharyngeal muscular dystrophy have a severe form of the disease. *Ann.Neurol.*, **46**, 115-118.
130. Wahle,E. (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell*, **66**, 759-768.
131. Wahle,E. (1995) Poly(A) tail length control is caused by termination of processive synthesis. *J.Biol.Chem.*, **270**, 2800-2808.
132. Chen,Z., Li,Y., and Krug,R.M. (1999) Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *EMBO J.*, **18**, 2273-2283.
133. Bear,D.G., Fomproix,N., Soop,T., Bjorkroth,B., Masich,S., and Daneholt,B. (2003) Nuclear poly(A)-binding protein PABPN1 is associated with RNA polymerase II during transcription and accompanies the released transcript to the nuclear pore. *Exp.Cell Res.*, **286**, 332-344.
134. Calado,A., Kutay,U., Kuhn,U., Wahle,E., and Carmo-Fonseca,M. (2000) Deciphering the cellular pathway for transport of poly(A)-binding protein II. *RNA.*, **6**, 245-256.
135. Fan,X., Messaed,C., Dion,P., Laganiere,J., Brais,B., Karpati,G., and Rouleau,G.A. (2003) HnRNP A1 and A/B interaction with PABPN1 in oculopharyngeal muscular dystrophy. *Can.J.Neurol.Sci.*, **30**, 244-251.
136. Abu-Baker,A., Messaed,C., Laganiere,J., Gaspar,C., Brais,B., and Rouleau,G.A. (2003) Involvement of the ubiquitin-proteasome pathway and molecular chaperones in oculopharyngeal muscular dystrophy. *Hum.Mol.Genet.*, **12**, 2609-2623.
137. Davies,J.E., Wang,L., Garcia-Oroz,L., Cook,L.J., Vacher,C., O'Donovan,D.G., and Rubinsztein,D.C. (2005) Doxycycline attenuates and delays toxicity of the oculopharyngeal muscular dystrophy mutation in transgenic mice. *Nat.Med.*, **11**, 672-677.
138. Fan,X., Dion,P., Laganiere,J., Brais,B., and Rouleau,G.A. (2001) Oligomerization of polyalanine expanded PABPN1 facilitates nuclear protein aggregation that is associated with cell death. *Hum.Mol.Genet.*, **10**, 2341-2351.
139. Bao,Y.P., Cook,L.J., O'Donovan,D., Uyama,E., and Rubinsztein,D.C. (2002) Mammalian, yeast, bacterial, and chemical chaperones reduce aggregate formation and death in a cell model of oculopharyngeal muscular dystrophy. *J.Biol.Chem.*, **277**, 12263-12269.
140. Tavanez,J.P., Calado,P., Braga,J., Lafarga,M., and Carmo-Fonseca,M. (2005) In vivo aggregation properties of the nuclear poly(A)-binding protein PABPN1. *RNA.*, **11**, 752-762.
141. Corbeil-Girard,L.P., Klein,A.F., Sasseville,A.M., Lavoie,H., Dicaire,M.J., Saint-Denis,A., Page,M., Duranceau,A., Codere,F., Bouchard,J.P., Karpati,G., Rouleau,G.A., Massie,B., Langelier,Y., and Brais,B. (2005) PABPN1 overexpression leads to upregulation of genes encoding nuclear proteins that are sequestered in oculopharyngeal muscular dystrophy nuclear inclusions. *Neurobiol.Dis.*, **18**, 551-567.
142. Abu-Baker,A., Laganiere,S., Fan,X., Laganiere,J., Brais,B., and Rouleau,G.A. (2005) Cytoplasmic targeting of mutant poly(A)-binding protein nuclear 1 suppresses protein aggregation and toxicity in oculopharyngeal muscular dystrophy. *Traffic.*, **6**, 766-779.
143. Kim,Y.J., Noguchi,S., Hayashi,Y.K., Tsukahara,T., Shimizu,T., and Arahata,K. (2001) The product of an oculopharyngeal muscular dystrophy gene, poly(A)-binding protein 2, interacts with SKIP and stimulates muscle-specific gene expression. *Hum.Mol.Genet.*, **10**, 1129-1139.
144. Hino,H., Araki,K., Uyama,E., Takeya,M., Araki,M., Yoshinobu,K., Miike,K., Kawazoe,Y., Maeda,Y., Uchino,M., and Yamamura,K. (2004) Myopathy phenotype in transgenic mice

expressing mutated PABPN1 as a model of oculopharyngeal muscular dystrophy. *Hum.Mol. Genet.*, **13**, 181-190.

145. Dion,P., Shanmugam,V., Gaspar,C., Messaed,C., Meijer,I., Toulouse,A., Laganiere,J., Roussel,J., Rochefort,D., Laganiere,S., Allen,C., Karpati,G., Bouchard,J.P., Brais,B., and Rouleau,G.A. (2005) Transgenic expression of an expanded (GCG)13 repeat PABPN1 leads to weakness and coordination defects in mice. *Neurobiol.Dis.*, **18**, 528-536.
146. van Gijlswijk,R.P., Talman,E.G., Janssen,P.J., Snoeijers,S.S., Killian,J., Tanke,H.J., and Heetebrij,R.J. (2001) Universal Linkage System: versatile nucleic acid labeling technique. *Expert.Rev.Mol.Diagn.*, **1**, 81-91.
147. Van Gelder,R.N., von Zastrow,M.E., Yool,A., Dement,W.C., Barchas,J.D., and Eberwine,J. H. (1990) Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc.Natl.Acad.Sci.U.S.A*, **87**, 1663-1667.
148. Pabon,C., Modrusan,Z., Ruvolo,M.V., Coleman,I.M., Daniel,S., Yue,H., and Arnold,L.J., Jr. (2001) Optimized T7 amplification system for microarray analysis. *Biotechniques*, **31**, 874-879.
149. Klur,S., Toy,K., Williams,M.P., and Certa,U. (2004) Evaluation of procedures for amplification of small-size samples for hybridization on microarrays. *Genomics*, **83**, 508-517.
150. 't Hoen,P.A., de Kort,F., van Ommen,G.J., and den Dunnen,J.T. (2003) Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.*, **31**, e20.
151. Churchill,G.A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.*, **32 Suppl**, 490-495.
152. Wei,C., Li,J., and Bumgarner,R.E. (2004) Sample size for detecting differentially expressed genes in microarray experiments. *BMC.Genomics*, **5**, 87.
153. Quackenbush,J. (2002) Microarray data normalization and transformation. *Nat.Genet.*, **32 Suppl**, 496-501.
154. Reiner,A., Yekutieli,D., and Benjamini,Y. (2003) Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics.*, **19**, 368-375.
155. Eisen,M.B., Spellman,P.T., Brown,P.O., and Botstein,D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc.Natl.Acad.Sci.U.S.A*, **95**, 14863-14868.
156. Slonim,D.K. (2002) From patterns to pathways: gene expression data analysis comes of age. *Nat.Genet.*, **32 Suppl**, 502-508.
157. Gentleman,R.C., Carey,V.J., Bates,D.M., Bolstad,B., Dettling,M., Dudoit,S., Ellis,B., Gautier,L., Ge,Y., Gentry,J., Hornik,K., Hothorn,T., Huber,W., Iacus,S., Irizarry,R., Leisch,F., Li,C., Maechler,M., Rossini,A.J., Sawitzki,G., Smith,C., Smyth,G., Tierney,L., Yang,J.Y., and Zhang,J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.
158. The Gene Ontology Consortium (2001) Creating the gene ontology resource: design and implementation. *Genome Res.*, **11**, 1425-1433.
159. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T., Harris,M.A., Hill,D.P., Issel-Tarver,L., Kasarskis,A., Lewis,S., Matese,J.C., Richardson,J.E., Ringwald,M., Rubin,G.M., and Sherlock,G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25-29.
160. Mons,B. (2005) Which gene did you mean? *BMC.Bioinformatics.*, **6**, 142.
161. Weeber,M., Kors,J.A., and Mons,B. (2005) Online tools to support literature-based discovery in the life sciences. *Brief.Bioinform.*, **6**, 277-286.
162. Eldering,E., Spek,C.A., Aberson,H.L., Grummels,A., Derks,I.A., de Vos,A.F., McElgunn,C. J., and Schouten,J.P. (2003) Expression profiling via novel multiplex assay allows rapid assessment of gene regulation in defined signalling pathways. *Nucleic Acids Res.*, **31**, e153.

163. Lin,Y.M., Furukawa,Y., Tsunoda,T., Yue,C.T., Yang,K.C., and Nakamura,Y. (2002) Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. *Oncogene*, **21**, 4120-4128.

164. Ramaswamy,S., Tamayo,P., Rifkin,R., Mukherjee,S., Yeang,C.H., Angelo,M., Ladd,C., Reich,M., Latulippe,E., Mesirov,J.P., Poggio,T., Gerald,W., Loda,M., Lander,E.S., and Golub,T.R. (2001) Multiclass cancer diagnosis using tumor gene expression signatures. *Proc.Natl.Acad.Sci.U.S.A*, **98**, 15149-15154.

165. 't Veer,L.J., Dai,H., van de Vijver,M.J., He,Y.D., Hart,A.A., Mao,M., Peterse,H.L., van der,K.K., Marton,M.J., Witteveen,A.T., Schreiber,G.J., Kerkhoven,R.M., Roberts,C., Linsley,P.S., Bernards,R., and Friend,S.H. (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, **415**, 530-536.

166. Pomeroy,S.L., Tamayo,P., Gaasenbeek,M., Sturla,L.M., Angelo,M., McLaughlin,M. E., Kim,J.Y., Goumnerova,L.C., Black,P.M., Lau,C., Allen,J.C., Zagzag,D., Olson,J.M., Curran,T., Wetmore,C., Biegel,J.A., Poggio,T., Mukherjee,S., Rifkin,R., Califano,A., Stolovitzky,G., Louis,D.N., Mesirov,J.P., Lander,E.S., and Golub,T.R. (2002) Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, **415**, 436-442.

167. Welsh,J.B., Sapinoso,L.M., Su,A.I., Kern,S.G., Wang-Rodriguez,J., Moskaluk,C.A., Frierson,H.F., Jr., and Hampton,G.M. (2001) Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res.*, **61**, 5974-5978.

168. Rosenwald,A., Wright,G., Chan,W.C., Connors,J.M., Campo,E., Fisher,R.I., Gascoyne,R. D., Muller-Hermelink,H.K., Smeland,E.B., Giltnane,J.M., Hurt,E.M., Zhao,H., Averett,L., Yang,L., Wilson,W.H., Jaffe,E.S., Simon,R., Klausner,R.D., Powell,J., Duffey,P.L., Longo,D.L., Greiner,T.C., Weisenburger,D.D., Sanger,W.G., Dave,B.J., Lynch,J.C., Vose,J., Armitage,J.O., Montserrat,E., Lopez-Guillermo,A., Grogan,T.M., Miller,T.P., LeBlanc,M., Ott,G., Kvaloy,S., Delabie,J., Holte,H., Krajci,P., Stokke,T., and Staudt,L.M. (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N.Engl.J.Med.*, **346**, 1937-1947.

169. Shai,R.M. (2006) Microarray tools for deciphering complex diseases. *Front Biosci.*, **11**, 1414-1424.

170. Kainou,T., Shinzato,T., Sasaki,K., Mitsui,Y., Giga-Hama,Y., Kumagai,H., and Uemura,H. (2006) Spsgt1, a new essential gene of Schizosaccharomyces pombe, is involved in carbohydrate metabolism. *Yeast*, **23**, 35-53.

171. King,J.Y., Ferrara,R., Tabibiazar,R., Spin,J.M., Chen,M.M., Kuchinsky,A., Vailaya,A., Kincaid,R., Tsalenko,A., Deng,D.X., Connolly,A., Zhang,P., Yang,E., Watt,C., Yakhini,Z., Ben Dor,A., Adler,A., Bruhn,L., Tsao,P., Quertermous,T., and Ashley,E.A. (2005) Pathway analysis of coronary atherosclerosis. *Physiol Genomics*, **23**, 103-118.

172. Nagaraja,G.M., Othman,M., Fox,B.P., Alsaber,R., Pellegrino,C.M., Zeng,Y., Khanna,R., Tamburini,P., Swaroop,A., and Kandpal,R.P. (2006) Gene expression signatures and biomarkers of noninvasive and invasive breast cancer cells: comprehensive profiles by representational difference analysis, microarrays and proteomics. *Oncogene*, **25**, 2328-2338.

173. Turk,R., Sterrenburg,E., van der Wees,C.G., de Meijer,E.J., de Menezes,R.X., Groh,S., Campbell,K.P., Noguchi,S., van Ommen,G.J., den Dunnen,J.T., and 't Hoen,P.A. (2006) Common pathological mechanisms in mouse models for muscular dystrophies. *FASEB J.*, **20**, 127-129.

174. van Duin,M., Woolson,H., Mallinson,D., and Black,D. (2003) Genomics in target and drug discovery. *Biochem.Soc.Trans.*, **31**, 429-432.

175. Imoto,S., Savoie,C.J., Aburatani,S., Kim,S., Tashiro,K., Kuhara,S., and Miyano,S. (2003) Use of gene networks for identifying and validating drug targets. *J.Bioinform.Comput.Biol.*, **1**, 459-474.

176. Searfoss,G.H., Ryan,T.P., and Jolly,R.A. (2005) The role of transcriptome analysis in pre-clinical toxicology. *Curr.Mol.Med.*, **5**, 53-64.

177. Narayanan,B.A., Narayanan,N.K., Desai,D., Pittman,B., and Reddy,B.S. (2004) Effects of a combination of docosahexaenoic acid and 1,4-phenylene bis(methylene) selenocyanate on cyclooxygenase 2, inducible nitric oxide synthase and beta-catenin pathways in colon cancer cells. *Carcinogenesis*, **25**, 2443-2449.