doi:10.1093/brain/awl023 *Brain* (2006), **129,** 996–1013

Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb–MyoD pathways in muscle regeneration

Marina Bakay,[1,*] Zuyi Wang,[1,2,*] Gisela Melcon,[1,*] Louis Schiltz,[4] Jianhua Xuan,[2] Po Zhao,[1] Vittorio Sartorelli,[4] Jinwook Seo,[1,5] Elena Pegoraro,[9] Corrado Angelini,[9] Ben Shneiderman,[5] Diana Escolar,[1] Yi-Wen Chen,[1] Sara T. Winokur,[6] Lauren M. Pachman,[7] Chenguang Fan,[1] Raul Mandler,[3] Yoram Nevo,[10] Erynn Gordon,[1] Yitan Zhu,[8] Yibin Dong,[8] Yue Wang[8] and Eric P. Hoffman[1]

[1]Research Center for Genetic Medicine, Children's National Medical Center, [2]Department of Electrical Engineering and Computer Science, The Catholic University of America, [3]Department of Neurology, George Washington University School of Medicine, Washington, DC, [4]Muscle Development Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, [5]Department of Computer Science, University of Maryland, College Park, MD, [6]Department of Biological Chemistry, University of California, Irvine, CA, [7]Children's Memorial Research Center, Northwestern University's Feinberg School of Medicine, Chicago, IL, [8]Department of Electrical, Computer and Biomedical Engineering, Virginia Polytechnic Institute and State University, Arlington, VA, USA, [9]Department of Neurosciences, University of Padova, Padova, Italy and [10]Pediatric Neurology, Hadassah Medical Center, Jerusalem, Israel

Correspondence to: Dr Eric Hoffman, PhD, Director, Research Center for Genetic Medicine, Children's National Medical Center, 111 Michigan Ave NW, Washington DC 20010, USA
E-mail: ehoffman@cnmcresearch.org
*These authors contributed equally to this work

Mutations of lamin A/C (LMNA) cause a wide range of human disorders, including progeria, lipodystrophy, neuropathies and autosomal dominant Emery–Dreifuss muscular dystrophy (EDMD). EDMD is also caused by X-linked recessive loss-of-function mutations of emerin, another component of the inner nuclear lamina that directly interacts with LMNA. One model for disease pathogenesis of LMNA and emerin mutations is cell-specific perturbations of the mRNA transcriptome in terminally differentiated cells. To test this model, we studied 125 human muscle biopsies from 13 diagnostic groups (125 U133A, 125 U133B microarrays), including EDMD patients with LMNA and emerin mutations. A Visual and Statistical Data Analyzer (VISDA) algorithm was used to statistically model cluster hierarchy, resulting in a tree of phenotypic classifications. Validations of the diagnostic tree included permutations of U133A and U133B arrays, and use of two probe set algorithms (MAS5.0 and MBEI). This showed that the two nuclear envelope defects (EDMD LMNA, EDMD emerin) were highly related disorders and were also related to fascioscapulohumeral muscular dystrophy (FSHD). FSHD has recently been hypothesized to involve abnormal interactions of chromatin with the nuclear envelope. To identify disease-specific transcripts for EDMD, we applied a leave-one-out (LOO) cross-validation approach using LMNA patient muscle as a test data set, with reverse transcription–polymerase chain reaction (RT–PCR) validations in both LMNA and emerin patient muscle. A high proportion of top-ranked and validated transcripts were components of the same transcriptional regulatory pathway involving Rb1 and MyoD during muscle regeneration (CRI-1, CREBBP, Nap1L1, ECREBBP/p300), where each was specifically upregulated in EDMD. Using a muscle regeneration time series (27 time points) we develop a transcriptional model for downstream consequences of LMNA and emerin mutations. We propose that key interactions between the nuclear envelope and Rb and MyoD fail in EDMD at the point of myoblast exit from the cell cycle, leading to poorly coordinated phosphorylation and acetylation steps. Our data is consistent with mutations of nuclear lamina components leading to destabilization of the transcriptome in differentiated cells.

Keywords: Skeletal muscle; lamin A/C; emerin; Emery-Dreifuss muscular dystrophy

Abbreviations: EDMD = Emery–Dreifuss muscular dystrophy; FSHD = fascioscapulohumeral muscular dystrophy; IDG = individual discriminatory genes; JDG = jointly discriminatory genes; LGMD = limb-girdle muscular dystrophy; LOO = leave-one-out; RT–PCR = reverse transcription–polymerase chain reaction; VISDA = Visual and Statistical Data Analyzer; wFC = weighted Fisher criterion

Received September 25, 2005. Revised December 23, 2005. Accepted January 11, 2006.
Advance Access publication February 14, 2006

Introduction

LMNA and the nuclear lamina

The nuclear envelope is composed of inner- and outer-membrane components, with an intermediate filament-based nuclear skeleton (lamina) bound to the inner membrane. The major components of the lamina are the nuclear lamins. Nuclear lamins are type V intermediate filaments and have globular amino and carboxyl termini, with four alpha-helical central domains separated by non-helical spacers. Nuclear lamins are classified into A and B types; B types (B1 and/or B2 genes) are ubiquitously expressed. The A-type lamins (lamin A and lamin C; LMNA gene) show more restricted expression, often at specific developmental stages or differentiation states.

LMNA protein isoforms have both structural and signalling roles. LMNA serves as scaffolding for binding of many other proteins, and many of these lamin binding proteins have ascribed functions. Three of the better characterized lamin binding proteins (emerin, LAP2, MAN1) contain a 40-amino acid LEM domain that interacts with a conserved chromatin protein, barrier to auto-integration factor (BAF) (Laguri *et al.*, 2001; Lee *et al.*, 2001; Shumaker *et al.*, 2001; Wolff *et al.*, 2001). Importantly, BAF acts as a transcriptional repressor, presumably at chromatin attachment sites at the nuclear envelope (Wang *et al.*, 2002). While LMNA and emerin show widespread expression, murine knockouts are viable, possibly via functional redundancy among family members (Sullivan *et al.*, 1999; Melcon *et al.*, 2006). On the other hand, BAF is required for cell viability (Bengtsson and Wilson, 2004).

Many additional functions have been ascribed to LMNA and its binding proteins, leading to many possible pathophysiological models when LMNA is mutated (*see* below). Functions include binding and localization of histones, actin, retinoblastoma protein, sterol-response-element-binding protein (SREPB) and components of RNA polymerase II transcription complexes (Gruenbaum *et al.*, 2000; Bengtsson and Wilson, 2004; Zastrow *et al.*, 2004). Thus, LMNA is probably involved in nuclear morphology, cytoskeleton organization, nuclear positioning in the cell, chromatin organization, transcriptional regulation and mechanically coupling the nuclear envelope and cytoskeleton (Starr *et al.*, 2001; Lee *et al.*, 2002; Starr *et al.*, 2002; Malone *et al.*, 2003).

LMNA, emerin and genetic disease

A series of disorders have been found to be caused by mutations in LMNA (Mounkes *et al.*, 2003). The most common is an autosomal dominant form of Emery–Dreifuss muscular dystrophy (EDMD) showing early onset of severe contractures (particularly of the posterior cervical, elbows and ankles), cardiac conduction defect (generally requiring a pace maker by age 20 years), cardiomyopathy and slowly progressive weakness and wasting in a humeroperoneal distribution (Bonne *et al.*, 1999; Brodsky *et al.*, 2000). The large majority of patients show a dominant inheritance pattern, generally caused by missense mutations (Di Barletta *et al.*, 2000). Limb-girdle muscular dystrophy type 1B (LGMD1B) is a very rare dominantly inherited dystrophy with few or no contractures, and this was also found to harbour different mutations of the LMNA gene (Muchir *et al.*, 2000).

A series of other disorders have been found to be caused by LMNA mutations. Familial partial lipodystrophy is a disorder where patients appear normal until puberty, then begin to show an altered subcutaneous fat distribution, with loss of limb fat and relative gain of face, neck and upper trunk fat leading to a cushingoid appearance (Kobberling *et al.*, 1975). Cao and Hegele (2000) identified the same R482Q missense mutation in five Canadian familial partial lipodystrophy pedigrees. A form of inherited dilated cardiomyopathy, with and without cardiac conduction defects (CMD1A) is caused by LMNA mutations. A series of 11 families were studied and six mutations identified in the LMNA gene (Fatkin *et al.*, 1999; Brodsky *et al.*, 2000). A de-myelinating peripheral neuropathy, Charcot-Marie-Tooth type 2B1 (CMT2B1), showed homozygous mutation in the LMNA rod domain (R298C) in three consanguineous Algerian families (De Sandre-Giovanolli *et al.*, 2002). Hutchinson–Gilford progeria syndrome is a disorder of rapid ageing (precocious senility), where patients typically die of arteriosclerosis prior to age 10 years. In the majority of patients, the same *de novo* cryptic splice site mutation was found in exon 11 of LMNA (Eriksson *et al.*, 2003). This splice site change leads to the loss of 50 amino acids near the carboxyl-terminus, including the proteolytic site, to produce mature lamin A. Mandibuloacral dysplasia (MAD) patients show severe dental crowding owing to a hypoplastic mandible, thin skin, stiff joints and wide cranial sutures. Many of the reported patients show an altered fat distribution, strikingly similar to that seen in familial partial lipodystrophy (Simha *et al.*, 2002). A series of consanguinous families have all shown the same homozygous R587H missense mutations in LMNA (Novelli *et al.*, 2002; Simha *et al.*, 2003). Finally, restrictive dermopathy, a congenital tight skin contraction syndrome, has been found to be caused

by a splicing mutation of LMNA in two patients (Navarro *et al.*, 2004).

While the LMNA gene shows many different dominantly inherited mutations resulting in a wide variety of clinical phenotypes, all emerin mutations lead to X-linked recessive loss of function of the protein, and the same EDMD phenotype. It is intriguing that emerin is more widely expressed than LMNA (Lee *et al.*, 2001), yet emerin loss of function leads to a consistent tissue-restricted EDMD phenotype. Some EDMD mutations of LMNA (but not other LMNA phenotypes) lead to mislocalization of emerin, where it is no longer retained within the nuclear envelope, but instead is abnormally released into the endoplasmic reticulum (Vaughan *et al.*, 2001; Holt *et al.*, 2003). Thus, loss of emerin from myofibre nuclei may be a unifying biochemical theme between both dominant LMNA EDMD and X-linked recessive emerin EDMD. However, the majority of LMNA mutations do not lead to mislocalization of emerin.

Molecular pathology of EDMD

The cellular and biochemical abnormalities caused by LMNA and emerin mutations are poorly understood. Both LMNA and emerin show widespread expression though most cells and tissues. However, different mutations of LMNA cause tissue- and age-specific pathologies. Mice made null for LMNA show a runted phenotype and early death, but do not show many of the symptoms of human LMNA disorders (Sullivan *et al.*, 1999; Lammerding *et al.*, 2004). Emerin null mice are phenotypically normal (Melcon *et al.*, 2006). Pathophysiological models include instability of the nuclear envelope (Raharjo *et al.*, 2001; Lammerding *et al.*, 2004), perturbations in chromatin organization and attachment, abnormal development of cells and tissues and perturbations of mRNA expression patterns (Hutchison, 2001; Maraldi *et al.*, 2002; Östlund *et al.*, 2003). For example, cultured cells and lymphocytes from progeria patients showed nuclear morphology abnormalities, a relative loss of lamin A from the majority of nuclei (~80%) and incorrect localization of lamin B into the nucleoplasm (De Sandre-Giovanolli *et al.*, 2003). However, the correlations between nuclear mophology, tissue involvement and patient symptoms have been less than gratifying. There have been some efforts to map the different mutations to specific subdomains of the LMNA protein, as an explanation for the different phenotypes. However, the overlap in phenotypes complicates this effort, and there are few clear genotype/phenotype correlations with regard to LMNA protein domains.

Here, we tested the model of altered gene regulation as a downstream consequence of LMNA and emerin mutations in EDMD patients. We used mRNA expression profiling analysis of 13 diagnostic groups, including EDMD patients with either emerin or LMNA mutations. We show that the expression profiles of these two nuclear envelope disorders are very closely related, and then use disease-specific transcriptional changes to build a biochemical model for specific perturbations of mRNA regulation pathways. Our model is that loss of emerin from the nuclear envelope leads to inappropriate timing of Rb and MyoD phosphorylation and acetylation steps during the myoblast/myotube transition. This model is validated and extended through targeted muscle regeneration studies in mouse knockouts of emerin in a second paper (Melcon *et al.*, 2006).

Methods

Patient muscle biopsies

Muscle biopsies were from patients being tested for neuromuscular disease biochemical analyses (IRB 'Candidate gene and protein studies in neuromuscular disease'), or enrolled into specific IRB-approved research studies. All biopsies were flash-frozen in isopentane cooled in liquid nitrogen immediately after excision and stored in air-tight hydrated containers until use. All muscle samples were tested histologically by cryosection and showed excellent preservation of muscle morphology.

All muscle biopsies showed excellent preservation by haematoxylin–eosin histology and were evaluated for histological preservation by a single co-author (E. Hoffman). Characteristics of the patient population are provided in Supplementary Table 1.

Expression profiling methods and quality control

Expression profiling from flash frozen muscle biopsies was done using methods and quality control procedures as we have described previously (Tumor Analysis Group, 2004). Briefly, RNA was isolated using standard Trizol reagent and methods, and quality of RNA was assessed by gel electrophoresis and/or Agilent Bioanalyzer. Ten micrograms of total RNA was then reverse-transcribed into double stranded cDNA using a T7-oligo-dT primer, and cRNA was produced using two biotinylated nucleotide precursors. A 4-fold amplification of starting material (total RNA) to biotinylated cRNA was set as the minimum threshold for continuing with microarray hybridizations.

Fifteen micrograms of cRNA was fragmented and hybridized to Affymetrix U133A microarrays. Hybridization solution was removed, and the same solution was then hybridized to Affymetrix U133B microarrays. Microarrays were washed and scanned as per standard procedure with scaling factors and present calls within acceptable ranges (Tumor Analysis Group, 2004).

All microarrays were entered into a single normalized project using dCHIP/MBEI (Li and Wong, 2001) and chip outliers were detected. Four microarrays were detected by outlier analysis and were excluded from all further studies. Both MAS5.0 and dCHIP difference model probe set algorithms were used, as we have found these to give good signal/noise ratios in human tissue projects (Seo *et al.*, 2004).

Leave-one-out and weighted Fisher criteria

We used weighted Fisher criterion (wFC) to measure the individually discriminatory power of each gene hence to identify the individual discriminatory genes (IDG) and jointly discriminatory genes (JDG). A leave-one-out (LOO) procedure was then employed

to measure the bias and variance of wFC. The details of the proposed robust gene selection method are described below.

First, gene i is selected as an IDG if its discriminatory power across all clusters, measured by wFC

$$J_{\mathrm{wFC}}(i) = \sum_{k=1}^{k_0-1} \sum_{l=k+1}^{k_0} \pi_k \pi_l \omega(\Delta_{kl}) \frac{\left(\mu_{i,k} - \mu_{i,l}\right)^2}{\sigma_{i,k}^2 + \sigma_{i,l}^2} \qquad (1)$$

is above an empirically determined threshold, where K_0 is the number of the clusters, π_k and π_l are the relative mass of classes k and l and $\mu_{i,k}$ and $\mu_{i,l}$ are the mean expression levels of gene i in classes k and l, respectively, with corresponding standard deviations $\sigma_{i,k}$ and $\sigma_{i,l}$. The weighting function, $\omega(\Delta_{kl})$, is designed to give more weight to the proximate cluster pairs. According to Loog *et al.* (2001), $\omega(\Delta_{kl})$ can be mathematically defined as

$$\omega(\Delta_{kl}) = \frac{1}{2\Delta_{kl}^2} erf\left(\frac{\Delta_{kl}}{2\sqrt{2}}\right),$$

where Δ_{kl} is the Mahalanobis distance between the classes k and l. As a special case, if $K_0 = 2$ (i.e. only one cluster pair), the weighting function $\omega(\Delta_{kl})$ will be set as 1.

Second, the computation of $J_{\mathrm{wFC}}(i)$ is augmented by a LOO procedure to improve the robustness of IDG selection. We describe the LOO procedure of IDG selection as follows. Suppose there are N samples in the study. By excluding one sample at a trial, we will obtain one measurement of $J_{\mathrm{wFC}}(i)$ for each gene each time. After N times of the LOO procedure, we obtain N measurements of $J_{\mathrm{wFC}}(i)$. We will then calculate the mean and variance of $J_{\mathrm{wFC}}(i)$ to determine the IDGs. There are two practical ways to use the mean and variance of $J_{\mathrm{wFC}}(i)$ to determine the individually discriminatory power: (i) using only the mean and (ii) using the least value of $J_{\mathrm{wFC}}(i)$ (worst case, considering the variance of the measurements). In this study, we use the mean measurement of $J_{\mathrm{wFC}}(i)$ across all genes to select and rank the IDGs accordingly.

Quantitative RT–PCR validations

Two hundred nanograms of total RNA was reverse-transcribed into single-stranded cDNA, and aliquots (10 ng) were processed for real-time polymerase chain reaction (PCR) using either 'assays demand' or custom-designed primers against regions of the target sequence similar to those detected by Affymetrix probe sets. Real-time PCR assays were done on an ABI 7700 TaqMan unit.

Double-tagged MyoD proteomics

Methods for constructions, transfections and identification of isolated proteins are as described previously (Puri *et al.*, 2002). Briefly, a double-tagged MyoD construct was utilized, with an amino-terminal Myc tag and carboxy-terminal FLAG-tag. HeLa cells were transfected with either a plasmid containing the double-tagged MyoD or a empty vector, using Superfect (Qiagen). Total cell lysate was prepared 48 h after transfection in lysis buffer. Proteins were then run through two immunoaffinity columns, first for Myc, then for FLAG. Equal amounts of eluted protein from double-tagged MyoD-transfected cells, and vector-transfected cells, were run on one-dimensional SDS–polyacrylamide gel electrophoresis (SDS–PAGE) gels. Coomassie blue staining was done, and bands that were presented in the MyoD cells, but not the vector-transfected cells, were excised. Ten stained gel slices were sent to the Harvard Microchemistry Facility. In gel digestion and mass spectometry peptide maps with MS/MS, determination of sequences was done by database searches as described previously (Eng *et al.*, 1994; LeRoy *et al.*, 1998). Single microcapillary runs were done on a reverse-phase HPLC column, directly coupled to the nano-electrospray ionization source of an ion-trap mass spectrometer (Finnigan LCQ DECA quadrupole ion-trap mass spectrometer).

Results

Generation of a 125 biopsy (250 microarrays) muscle biopsy data set from 13 diagnostic groups

A description of the samples utilized for the human muscle data set is provided in Supplementary Table 1. Thirteen diagnostic groups were studied, with 125 muscle biopsies total. None of these U133 data set expression profiles has been published previously. We have previously published expression profiles on a small subset of the patient muscle biopsies using an earlier microarray (U95A) [acute quadriplegic myopathy (AQM), spastic paraplegia and normal muscle (3 out of 18 samples)]. References are given below for these reports. One of the 13 diagnostic groups was composed of normal adult volunteers enrolled in exercise physiology studies ('Normal'; $n = 18$). Ten of the thirteen groups were defined by molecular diagnostic results (gene and/or protein testing) considered diagnostic of the specific disorder, and all showed histological findings, clinical phenotype and family histories consistent with that diagnosis: Duchenne muscular dystrophy, Becker muscular dystrophy, Fukutin-related protein (FKRP), dysferlin deficiency (LGMD2B), calpain 3 deficiency (LGMD2A), fascioscapulohumeral muscular dystrophy (FSHD), hereditary spastic paraplegia type 4 (HSP; or SPG4), emerin deficiency (X-linked EDMD) and LMNA mutations (autosomal dominant EDMD). Two of the thirteen groups were given the assigned diagnosis by clinical examination, using current criteria: amyotrophic lateral sclerosis (ALS) and AQM. All muscle biopsies were flash frozen in isopentane cooled in liquid nitrogen and were stored at −80°C until analysis. All biopsies showed excellent preservation by histological analysis of cryosections by haematoxylin and eosin staining.

Normal muscle biopsies ($n = 18$) were from normal volunteers participating in exercise physiology studies (*see* Chen *et al.*, 2003). Juvenile dermatomyositis (JDM) ($n = 25$) is an autoimmune disease of children, and biopsies were part of molecular studies of this disorder (*see* Tezak *et al.*, 2002). Calpain 3 ($n = 10$) is a muscle-specific calcium-activated protease that shows loss-of-function mutations of the corresponding gene in LGMD2A; mutation data on two subjects have been published (Chou *et al.*, 1999), whereas others had mutations found for this study (Supplementary Table 1). FKRP ($n = 7$) is a glycosidase involved in glycosylation of dystroglycan, and all patients were homozygous for the common missense mutation C826A (Leucine276Isoleucine) seen in LGMD2I. Duchenne muscular dystrophy patients ($n = 10$)

showed complete lack of dystrophin by immunoblot and immunostaining of muscle (*see* Chen *et al.*, 2000; Bakay *et al.*, 2002*b*; Chen *et al.*, 2005). Becker muscular dystrophy patients (n = 5) showed dystrophin protein of abnormally small molecular weight by immunoblot, consistent with in-frame deletions of the dystrophin gene (n = 5). Dysferlin-deficient patients (LGMD2B; n = 10) showed complete absence of dysferlin by immunoblot analyses. FSHD patients had chromosome 4q sub-telomeric deletions characteristic of this disorder, and a subset of these patients have been reported previously (Winokur *et al.*, 2003) (n = 14). AQM patients (also called critical care myopathy) have been recently published by us (DiGiovanni *et al.*, 2004) (n = 5). Upper motor neuron disease (SPG4; HSP) patients were heterozygous for spastin gene mutations and were previously reported by us (Molon *et al.*, 2004) (n = 4). ALS biopsies were from patients enrolled into a specific research study (n = 9).

Emery–Dreifuss patients were in two groups: autosomal dominant due to LMNA gene mutations (n = 4) and X-linked recessive due to emerin mutations (n = 4). The autosomal dominant patients harboured novel mutations of the LMNA gene detected by gene sequencing, and have not been previously reported. Of the X-linked recessive patients, three were from a single pedigree that has been reported previously (Nevo *et al.*, 2001; Family 4; exon 6 1675-1678delTCCG), and these patients were brought to the NCRR Clinical Research Center at Children's National Medical Center for muscle biopsy for the purposes of this current study. The remaining X-linked patients had an intronic duplication of the emerin gene detected by PCR analyses of genomic DNA.

We used one round biotinylated cRNA production from total RNA isolated from each muscle biopsy. We generated 250 high-quality mRNA expression profiles from 125 patient muscle biopsies from 13 diagnostic groups (125 U133A and 125 U133B Affymetrix microarrays). All microarrays were generated by the Research Center for Genetic Medicine according to a standard operating procedure, and all profiles passed numerous quality control metrics for RNA quality and quantity, image intensity and quality and lack of chip outliers by MBEI analyses (Li and Wong, 2001; Tumor Analysis Group, 2004).

Data from this study is available from the PEPR resource (http://pepr.cnmcresearch.org/browse.do?action=list_prj_exp&projectId = 319), and NCBI GEO (http://www.ncbi.nih.gov/geo). Both raw data and processed files have been made public for 121 of 125 patients (242 of 250 U133A and U133B profiles). Four of the JDM sample (JDM group) profiles corresponding to male patients have not been made public owing to pending publication of these in a manuscript focused on JDM sex differences (L. Pachman, Y.W. Chen; personal communication). These four samples are not necessary for the interpretation of the Emery–Dreifuss data presented in this current manuscript. In our PEPR portal, analysis with five probe set algorithms is available for all 242 profiles (MAS5.0, dCHIP PM/MM, dCHIP PM only, ProbeProfiler, RMA). Public access analysis tools for this data set are available through NCBI GEO, PEPR, and also through our SAS web server (http://sas.cnmcresearch.org).

Data analysis of the human biopsy data set

We have previously reported that both MAS5.0 and dCHIP difference model probe set algorithms provide a good signal/noise balance in human muscle expression profiling projects (Seo *et al.*, 2003, 2004). However, it is also known that different probe set algorithms lead to very different data interpretations. We, therefore, normalized all data separately for MAS5.0 and dCHIP (MBEI) probe set algorithms. To filter out low signals and poorly performing probe sets, a 10% MAS5.0 'present call' filter was used on all data.

We then tested the relationships of the 13 diagnostic groups within the 125 muscle biopsy data set by creating a diagnostic classification tree using Visual and Statistical Data Analyzer (VISDA) (Wang *et al.*, 2000, 2003) in both supervised and unsupervised modes. Different from the conventional hierarchical clustering, our analysis used a sub-space Fisher hierarchical clustering (SFHC) mechanism to discover the tree of phenotype among the diagnostic groups in two steps. First, samples were assigned to known diagnostic groups in which the information about how related each individual within each diagnostic category is kept in the within-group scatter matrix $\mathbf{S}_w$. Second, the relationship of the different groups was determined (unsupervised) by comparing the relatedness $\mathbf{S}_{kl}$ between diagnostic categories using sub-space Fisher criterion as the distance measure $D_{k,l} = Trace(\mathbf{S}_w^{-1}\mathbf{S}_{kl})$ that exploits the information about how related each individual within each diagnostic group is as compared with the relatedness between the diagnostic groups (Loog *et al.*, 2001). VISDA, a hierarchical statistical data visualization and cluster exploration algorithm with hierarchical statistical models and visualization space, has been proven to be able to discover hidden cluster structure and statistically estimate cluster parameters. Furthermore, VISDA is designed to use the top-level model (in a statistical hierarchical mixture model) and top-level projection (of the entire data set) to explain structural information of the entire data set, and then use lower-level models and projections (of 'softly' partitioned sub-clusters) to explain local and internal structure between individual clusters that may not be obvious at the upper levels. Owing to the hierarchical nature of VISDA, the relationships (e.g. distance in the space) among classes are also unveiled during clustering, and such relationships can be shown in a tree structured class grouping figure. The more distant classes in the gene expression space (e.g. in different branches) that are generally less similar classes in terms of their gene expression levels are more easily captured by the VISDA at the upper levels, and therefore are separated at the upper levels of the tree.

We first used the U133A data set with MAS5.0 normalization and grouped the 13 muscle biopsy diagnostic classes into a tree structured class grouping with subgroups of related disorders using VISDA, with the number of IDG

(Xuan *et al.*, 2004) shown at each node (Fig. 1A). This analysis showed four disorders to be most distantly related to the other nine groups, with branching of these four groups from the remainder high in the clustering tree (AQM, JDM, ALS, HSP). This is consistent with each having an underlying disease that is much different from the other groups. The tree then showed two major branches. The left branch included most of the 'dystrophic myopathies', inborn single gene disorders causing degeneration/regeneration of muscle fibres (DMD, BMD, calpain 3, FKRP, dysferlin). The right branch included normal muscle, both nuclear lamina disorders (emerin, LMNA) and FSHD. The two nuclear envelope disorders then formed their own group, indicating that the mRNA profiles were very closely related. The co-segregation of FSHD with EDMD as a distinct branch of the dystrophies is interesting, as FSHD has recently been shown to disrupt chromatin attachment sites to the nuclear envelope (Masny *et al.*, 2004).

To validate the derived tree structure, we used classification with adaptive hierarchical sub-space experts (AHSE) (Fig. 1B). At each node, one multi-layer perceptron (MLP) is designed to classify the specific classes or groups; and the inputs of MLPs are JDG selected for these specific tasks (Xuan *et al.*, 2004). Since the most distant and separable classes or groups among other remaining classes are expected to be partitioned at each node, the misclassification rate of its corresponding MLP should be small. The misclassification rates at each node for both the training and test sets are shown. Despite the many confounding variables intrinsic to human biopsies (inter-individual variation, tissue heterogeneity) we found that most training and test misclassification rates were quite low (0–3%) at most nodes. The one exception was Node 6 (N6; Fig. 1B), where the misclassification rate of the test data set was 7%. The good classification performance at most nodes of the tree implies that the tree reflects a valid picture of the relationships among classes.

We then ran a series of validations for the tree structure, using two distinct probe set algorithms (MAS5.0 and MBEI), and also both U133A and U133B microarrays for the same samples. U133A and U133B microarrays have very little overlap in probe set content, and each interrogates about half of known or proposed transcript units in the genome. The IDG/VISDA and JDG training/test analyses (Fig. 1A and B) were then validated with three additional data sets: U133A MBEI/dCHIP probe set normalization, and then the semi-independent U133B data (both MAS5.0 and MBEI/dCHIP normalizations). All analyses resulted in a similar tree structure, with co-segregation of EDMD LMNA and EDMD emerin into the same branch (Fig. 1C and D for U133B MBEI/dCHIP) (U133A MBEI, and U133B MAS5.0 not shown). While the U133A data gave consistently low misclassification rates, the U133B data gave generally higher test misclassification rates (Fig. 1D). This is probably due to the relatively high proportion of probe sets corresponding to expressed sequence tags (ESTs) or other less well-characterized transcript units, and the generally lower number of 'present calls' on this microarray, as we have shown previously for the U95 A, B, C, D, E chip sets in muscle (Bakay *et al.*, 2002*b*).

There are many confounding variables in such an analysis. Sex, age, ethnic background and severity of pathology are all uncontrolled variables. In addition, three of the four Emerin patients utilized were members of the same family, and this could lead to higher grouping of these subjects. Nevertheless, the finding of similar tree structures using two different probe set algorithms, and two different microarrays, suggests that diagnosis (molecular defect) rather than hidden admixture due to confounding variables drives the tree structure shown. These data serve to generate a hypothesis that must be extensively tested by other methods and approaches.

Gene selection optimized for biochemical pathways suggests disease-specific perturbations of a lamin A–emerin–Rb–MyoD–CREBBP/p300 acetylase biochemical pathway

The IDG/VISDA and JDG approach above is optimized so that the lowest number of genes with the most discriminatory power is selected. This approach is too stringent for discerning biochemical pathways, as the goal of identification of biochemical pathways should not include a positive weight for low numbers of pathway members, but rather a positive weight for the highest proportion of pathway members shown to be dysregulated.

To identify potential biochemical pathways specific to the nuclear envelope disorders (EDMD LMNA and EDMD emerin), we used the following approach. First, we selected potential pathway members by selecting genes that can distinguish the LMNA mutation-positive EDMD group ($n = 4$ biopsies) and the other 11 groups as a single combined group based on wFC, a statistical cluster separability measure (Loog *et al.*, 2001). We used only the LMNA patients for this analysis for two reasons: the LMNA patients were unrelated, whereas some of the emerin patients were from the same family, and we wanted to use the emerin group as a validation set using alternative methods. To assess the generalizability of the selected gene subset, stability analysis was performed via a LOO cross-validation (Xuan *et al.*, 2004), in which the gene subset with most discriminative power is selected on the basis of a training set with one sample profile being left out. The profile is then put back into the data set, a different profile then excluded, and the analysis run again. Each time a gene presents in any one subset, the gene is given one 'vote'. This process is repeated 121 times, so that the maximum number of votes is 121 for a gene that is fully informative in discriminating between the two groups. Discriminatory genes were ranked by vote number and deemed as potential LMNA biochemical pathway members.

We then compared each of the top-ranked probe sets against normal skeletal muscle (e.g. 4 LMNA profiles versus

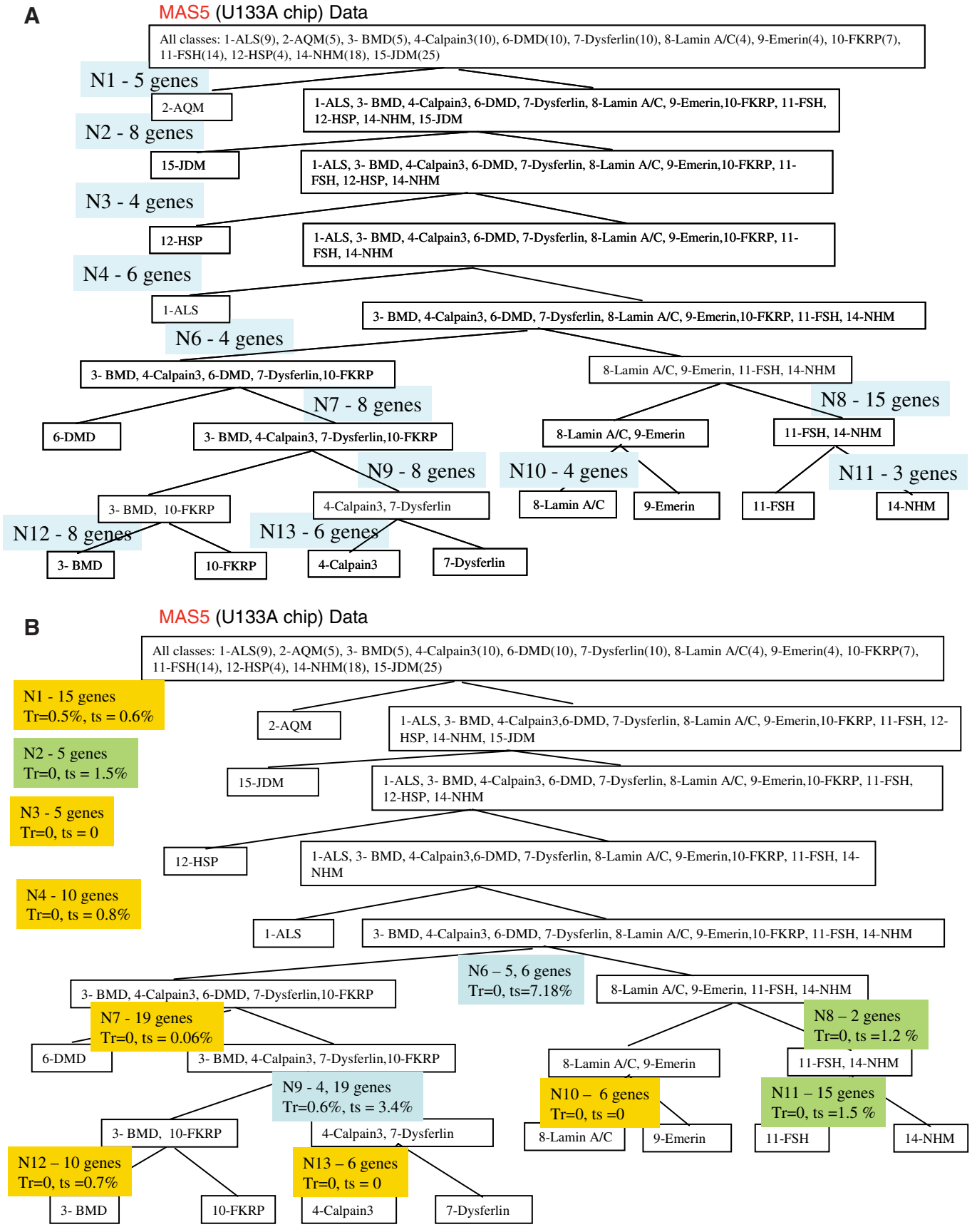
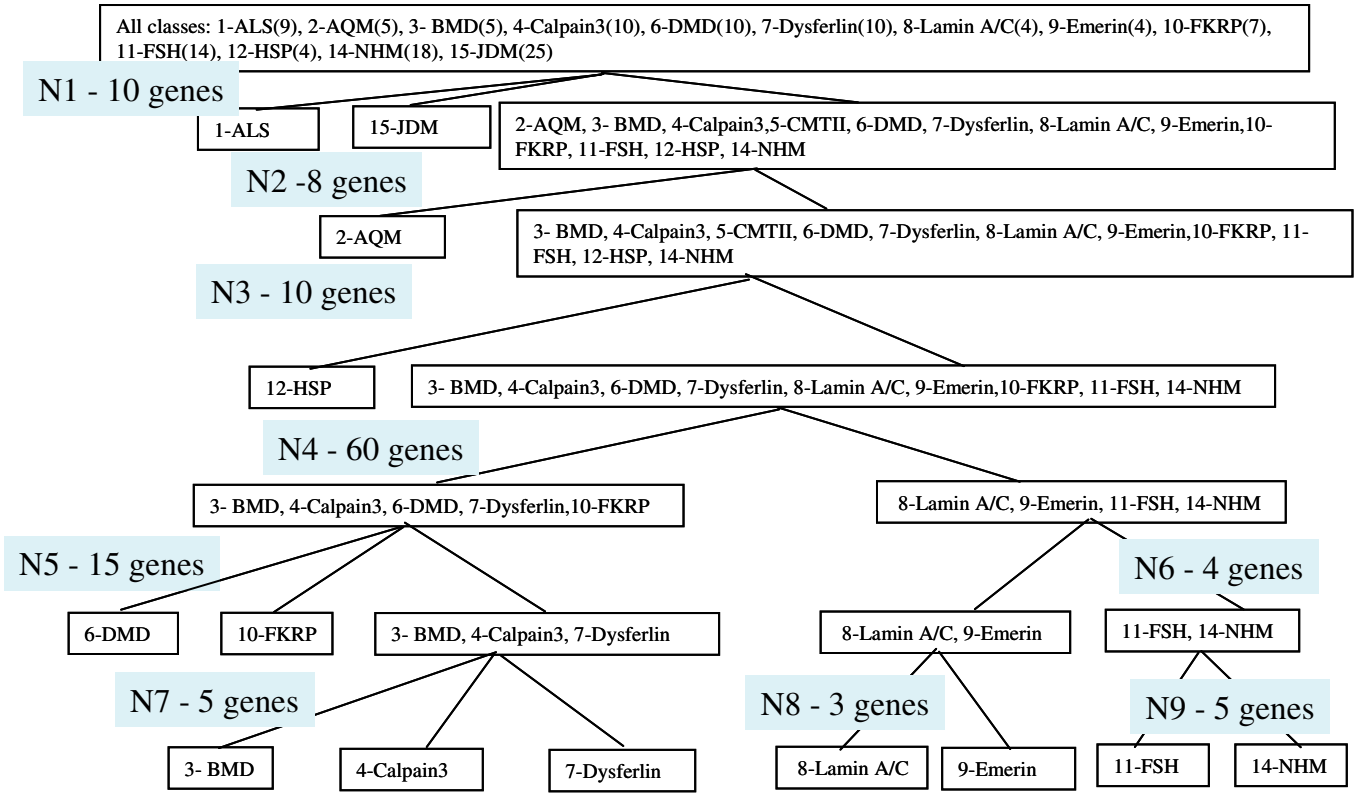
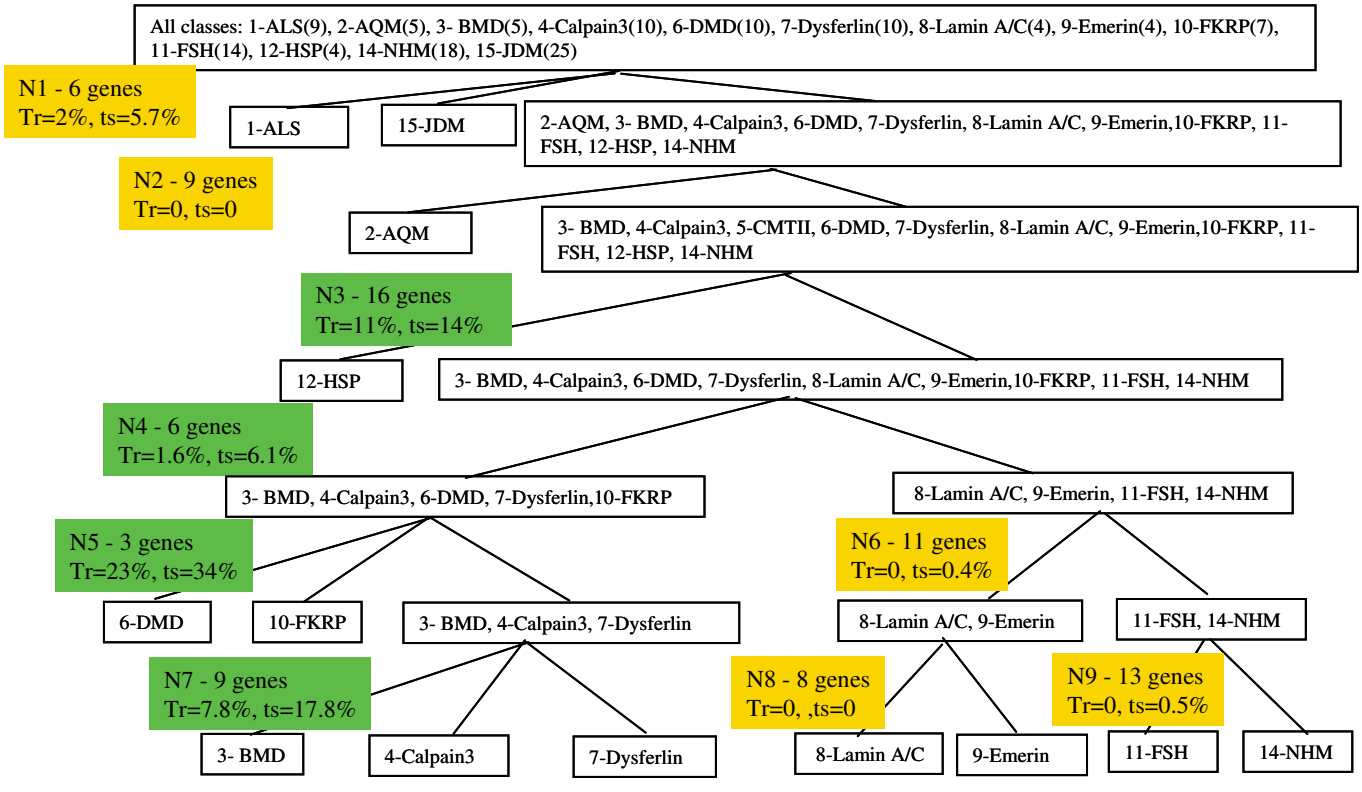


Fig. 1 Continued

C MBEI/dCHIP (U133B chip) Data



D MBEI/dCHIP (U133B chip) Data



18 normal muscle biopsies). The output of this analysis was a *P*-value and relative fold-change (Supplementary Table 2). This was done for both MAS5.0 and dCHIP, again using a 10% present call data filter. This analysis provided 101 probe sets with between 84 and 21 votes (Supplementary Table 2), then no probe sets between 21 and 8 votes.

In studying the assigned functions of the 101 selected probe sets we found that 7 of the top 36 probe sets (20%) could be assigned a known or probable role in a myogenic differentiation pathway at the time point of terminal differentiation of myoblasts into myotubes (Table 1). Each of these seven showed significant upregulation in LMNA EDMD muscle. Specifically CREBBP is an acetylase involved in activation of MyoD, and also acetylation of histones on downstream MyoD transcriptional targets (Li *et al.*, 2000; Puri *et al.*, 2001; Magenta *et al.*, 2003). CRI-1 is an inhibitor of CREBBP, which controls the timing of acetylase activity (MacLellan *et al.*, 2000). NAP1L1 is a protein that binds acetylated histones and serves as a chaperone to open up chromatin on downstream targets of MyoD/CREBBP (Shikama *et al.*, 2002). LAP2 is a component of the inner nuclear lamina, where it cooperates with LMNA to bind hypophosphorylated Rb at the nuclear envelope (Markiewicz *et al.*, 2002). RBL2 (also p130) is a retinoblastoma-like protein involved in terminal differentiation of many cell types.

We validated the increased expression of the Rb–MyoD pathway members, using quantitative reverse transcription–polymerase chain reaction (RT–PCR) in muscle biopsies from patients with both LMNA and emerin mutations. The four tested pathway members (CREBBP, p300, CRI-1, NAP1L1) showed similar upregulation (Table 2 and Fig. 2). The upregulation appeared blunted in all cases in EDMD emerin patient muscle compared with EDMD LMNA patient muscle; however, three of four EDMD emerin subjects were much later in the progressive disease process than the LMNA patients. Thus, this may reflect a consequence of severity of pathology.

The abnormally regulated transcripts specific for the two nuclear envelope disorders (EDMD due to LMNA and emerin) were each significantly increased in expression (CREBBP, CRI-1, EP300, RBL2, LAP2/TMPO and two probe sets for NAP1L1). We hypothesized that the upregulation of these transcripts was in response to a biochemical block in one or more biochemical pathways involving MyoD, Rb and the transition from proliferating myoblasts to post-mitotic myotubes in muscle regeneration (Li *et al.*, 2000; MacLellan *et al.*, 2000; Kitzmann *et al.*, 2001; Puri *et al.*, 2001; Peschiaroli *et al.*, 2002).

Construction of a biochemical model based upon temporal profiling *in vivo*

We have previously described time series data during muscle degeneration/regeneration *in vivo*, where degeneration was induced with cardiotoxin injection, and muscles harvested at 16 time points during regeneration, typically at 12–24 h intervals (Zhao *et al.*, 2002, 2003, 2004, 2006) (*see* http://pepr.cnmcresearch.org for public access to data set and query tools). This data set allows the placement of activation of gene transcription of myogenic genes into a temporal context, thereby achieving cause/effect in transcriptional cascades. To build a model for biochemical perturbations downstream of LMNA and emerin mutations in muscle, we synthesized data from the literature on biochemical pathways involving Rb, MyoD and CREBBP/p300 acetylase and the other EDMD-specific transcripts, and then validated and extended the pathways using a 27 time point *in vivo* muscle regeneration time series data set (Zhao *et al.*, 2002, 2003, 2004). Most previously published studies have focused on an immortalized murine cell line (C2C12) and have not studied these pathways in the context of normal regeneration *in vivo* (Iezzi *et al.*, 2002; Charge and Rudnicki, 2004; Simone *et al.*, 2004). It is becoming increasingly recognized that muscle regeneration *in vivo* requires an intact basal lamina that is absent *in vitro*. Indeed, *in vitro* differentiation of C2C12 cells does not involve most of the complex series of events known to be important *in vivo*, with interplay of infiltrating immune cells, activation of quiescent satellite (stem) cells, proliferation of myoblast to fill up the pre-existing basal lamina, fusion to myotubes and maturation of the myotubes into muscle fibres. Most of the proposed pathway members have not been studied in regeneration *in vivo*. Indeed, muscle regeneration proceeds very poorly with a dysfunctional basal lamina, as is the case with the congenital muscular dystrophies (Pegoraro *et al.*, 1996).

Fig. 1 Diagnostic/phenotypic tree construction from a 125 biopsy, 250 profile data set shows that EDMD due to LMNA mutations and EDMD due to emerin mutations are highly related. (**A**) Shown is the diagnostic tree defined using the U133A data set (125 profiles) of 13 diagnostic groups, normalized using MAS5.0 probe set algorithm. The number of biopsies per subgroup is provided in the top box. The bioinformatic method used was VISDA (Wang *et al.*, 2000, 2003) based upon individual discriminatory probe sets (IDG). The number of IDGs used at each node is shown in the boxes at each node. This data analysis shows the muscular dystrophies to branch into a 'membrane defect' group [left branch of node 6 (N6)] and a nuclear envelope branch [LMNA and emerin mutations; left branch of node 8 (N8)]. This data suggests that the two forms of EDMD share similar perturbations of mRNA profiles. (**B**) The data and tree structure in **A** were then tested for misclassification rates using jointly discriminatory gene approach (JDG) and AHSE approach. The misclassification rates are given for both the training (Tr) and test (ts) sets. Variable, node-specific scaling is shown, and was found to improve performance. (**C**) The U133B data set was normalized by MBEI/dCHIP, and then a new diagnostic tree constructed using the same bioinformatics approaches as in **A** (IDG VISDA). This tree shows a very similar structure to that in **A**, despite using a different microarray, and different probe set algorithm. Again, there is a distinction between the plasma membrane dystrophies and the nuclear envelope dystrophies. (**D**) Shown is the data in **C** tested for misclassification rates. Misclassification rates are found to be higher than in **B**, probably due to poorer performance of the probe sets on the U133B microarray.

Table 1 Gene selection for lamin A/C EDMD ($n = 4$) versus all other ($n = 121$) using wFC, LOO and biochemical pathway building

Committee vote	Fold change ED-versus-NHM	*P*-values	AFFY ID	Symbol	Gene name
79	1.3	0.04798	202160_at	CREBBP	CREB binding protein (Rubinstein–Taybi syndrome)
73	1.7	0.01118	211698_at	CRI1	CREBBP/EP300 inhibitory protein 1
60	1.8	0.00747	202221_s_at	EP300	E1A binding protein p300
58	2.2	0.00002	212331_at	RBL2	retinoblastoma-like 2 (p130)
46	1.5	0.00147	209754_s_at	LAP2/TMPO	lamina-associated polypeptide 2/thymopoietin
45	1.8	0.00017	208754_s_at	NAP1L1	nucleosome assembly protein 1-like 1
37	1.8	0.00036	208753_s_at	NAP1L1	nucleosome assembly protein 1-like 1

Fold change and *P*-values are relative to normal skeletal muscle ($n = 18$).

Table 2 Validation of the Rb–MyoD–CREBBP/p300 pathway perturbations in both lamin A/C and emerin EDMD biopsies using quantitative RT–PCR

Gene	Lamin A/C			Emerin		
	Fold change	*P*-value	*n*	Fold change	*P*-value	*n*
CREBBP	+5.1	0.001	4	+1.8	0.057	4
ECREBBP/P300	+2.9	0.020	3	+2.2	0.020	4
CRI-1	+2.7	0.010	4	+1.7	0.040	4
RBL-1	+4.8	0.040	3	+1.8	0.004	4
NAP1L1	+2.2	0.002	4	+1.6	0.040	4

To build a model of Rb–MyoD biochemical pathways *in vivo*, we relied heavily on a 27 time point *in vivo* muscle regeneration profiling series (Zhao *et al.*, 2003, 2004, 2006). Muscle regeneration was induced in adult mice using a 10-needle injection manifold covering the entire gastrocnemius, and histologically matched muscles expression profiled at the indicated time points. We then used our web Oracle query database to generate the temporal profiles for each candidate gene in the pathways over the 27 time points (Chen *et al.*, 2004) (http://pepr.cnmcresearch.org). Query of each transcript allowed us to sign a relatively precise time point where the transcript was expressed during regeneration.

With regard to the nuclear envelope, lamin B receptor (LBR), lamin B1 and LAP1 showed expression early in regeneration (day 1), with gradual decrease of expression during differentiation to the basal level at Day 16 (Fig. 3). Emerin was equally expressed at all time points (Fig. 3). Lamin B2, LAP2 and LMNA showed later transcriptional induction by the Day 3.5 time point. The 3.0–3.5 day time point is the time frame where we have previously shown strong induction of MyoD and its downstream targets (Zhao *et al.*, 2002, 2003, 2004) (Fig. 4). This data suggests that there is a remodelling of the nuclear envelope at or about the time of myoblast exit from the cell cycle and terminal differentiation.

We then studied the Rb–MyoD–CREBBP/p300 acetylase pathway during muscle regeneration. Rb, CREBBP/p300, NAP1L1 and MEF2A all showed peak expression commensurate with that of MyoD and the change in structure of the nuclear envelope (Fig. 4). HDAC was expressed at about 2 days regeneration. As Rb is known to be regulated by phosphorylation, we used the Kegg databases to search for candidate kinases of Rb and found only one of these, CDK4, to be highly expressed in muscle and differentially regulated during regeneration. We propose that CDK4 is the Rb kinase relevant to muscle regeneration (Figs 4 and 5). Targets of MyoD binding and transcriptional induction, such as the alpha subunit of the acetylcholine receptor (AchRalpha), were induced 12 h later than MyoD, at the Day 3.5 time point.

Assembling literature data on protein–protein interactions *in vitro*, together with query of gene family members in our regeneration time series, allowed us to build a model of relevant events during the specific time frame in which myoblasts exit the cell cycle and commit to terminal differentiation during muscle regeneration *in vivo* (Fig. 5A), and how these might be perturbed in disorders of the nuclear envelope (Fig. 5B). At Day 3.0 of regeneration, MyoD, Nap1L1 and CDK4 are induced. MyoD is bound to pre-existing HDAC, where it is kept hypo-acetylated and inactive. At Day 3.5 of regeneration, Rb1is induced, but it is phosphorylated by CDK4 that is pre-existing (from Day 3), and thus inactive. The CREBBP acetylase and three nuclear lamina components (lamin A/C, LAP2, lamin B2) are all induced at Day 3.5, at which point the machinery to sense appropriate time to exit from the cell cycle is in place in the proliferating myoblasts. When sufficient myoblasts are resident in myotubes, Rb becomes de-phosphorylated, whereupon it recruits HDAC from MyoD, and the Rb/HDAC associates with the nuclear envelope via lamin A/C and LAP2. This allows MyoD to bind the acetylase CREBBP, become acetylated and act on downstream differentiation targets. CREBBP then acetylates histones on MyoD target genes, and Nap1L1 then escorts the acetylated histones from the chromatin to open up the chromatin structure.

In muscle with EDMD due to either lamin A/C or emerin mutations, we propose that a biochemical block then occurs, preventing the appropriate phosphorylation and acetylation steps, owing to failure of Rb to associate correctly with

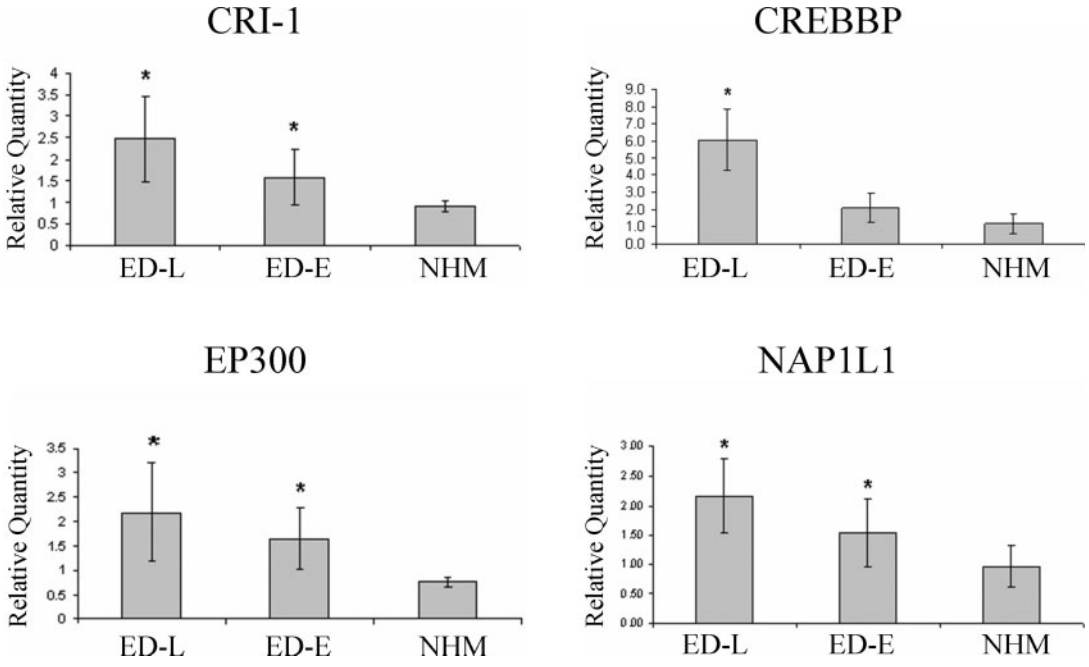


Fig. 2 Validation of upregulation of the Rb–MyoD–CREBBP/p300 pathway in both LMNA and emerin EDMD biopsies by quantitative RT–PCR. Shown are results of quantitative RT–PCR from muscle biopsy RNA from EDMD patients having mutations of either LMNA (autosomal dominant) or emerin (X-linked recessive). All tested pathway members validated by this alternative approach in both disorders. *$P < 0.05$.

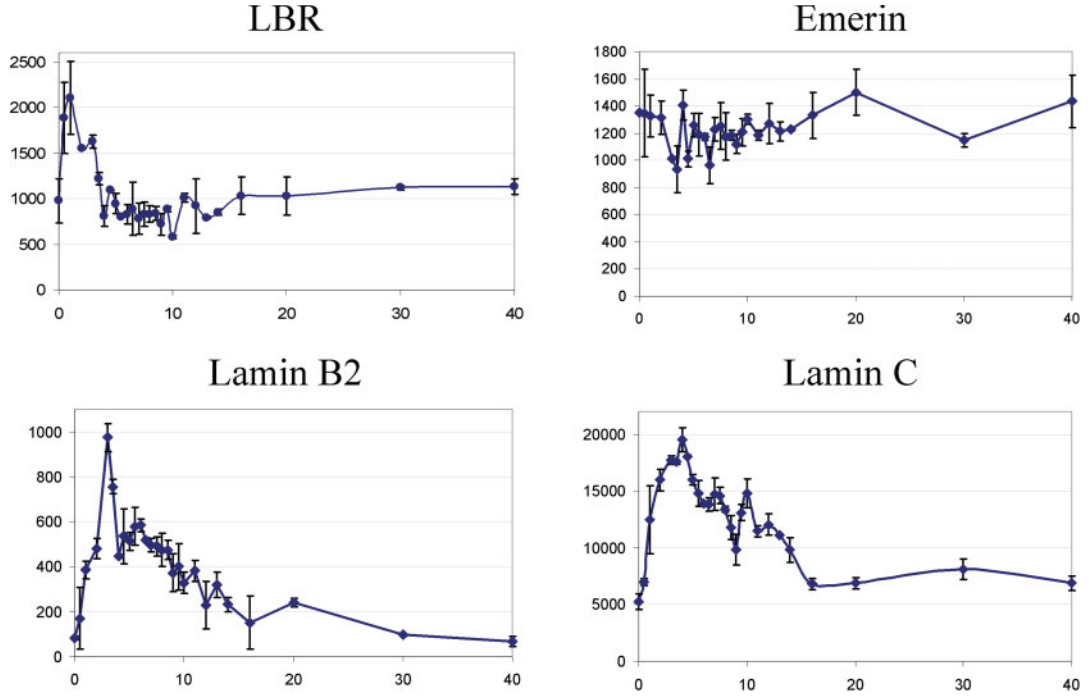


Fig. 3 Coordinated nuclear envelope remodelling during muscle regeneration *in vivo*. Shown are the expression patterns of genes in the nuclear envelope correlated with MyoD induction. Two histologically matched muscles are shown profiled at each time point after staged muscle degeneration/regeneration, with graph shown of the average of the two values. The data set (Zhao *et al*., 2003, 2004) as well as the public access graphics chart tool used for data visualizations (Chen *et al*., 2004) have been described previously. Twenty-seven time points, with 54 expression profiles, are shown. Lamin B receptor (LBR) is induced early during regeneration (day 1). Lamin B2 and Lamin C appear later (day 3) commensurate with terminal differentiation of myoblasts into myotubes. Emerin expression remains constant during regeneraiton.

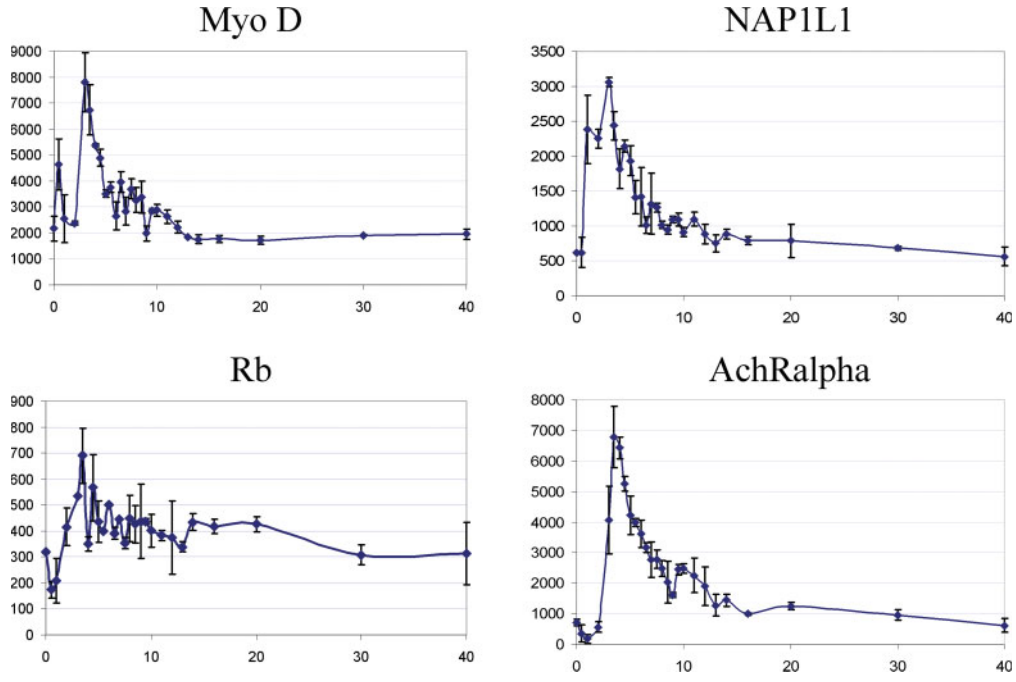


Fig. 4 Transcriptional induction of MyoD/Rb pathway members commensurate with myoblast/myotube transition.

the nuclear envelope (Fig. 5B). The other members of this pathway show compensatory upregulation in an effort to overcome this biochemical block, as observed in the expression profiles of both X-linked emerin deficiency and dominant missense mutations of lamin A/C.

MyoD as a binding partner for LMNA

It is thought that de-phosphorylated Rb directly binds LMNA (Mancini *et al.*, 1994; Ozaki *et al.*, 1994; Markiewicz *et al.*, 2002), and that LMNA is required for Rb stability in LMNA null cells (Johnson *et al.*, 2004). Recent findings have also suggested that Rb is acetylated during the differentiation of myoblasts, and that the acetylase is the same p300 (and P/CAF) acetylase responsible for acetylation of MyoD (Nguyen *et al.*, 2004). We felt that it was possible that MyoD might bind to LMNA directly, and thus LMNA and emerin might assist in the orchestrated acetylation/de-acetylation of Rb and MyoD that is known to be important for myogenic differentiation. To provide evidence for direct binding of MyoD to LMNA, we used stable transfection of double-tagged MyoD construct (FLAG and HA tagged) into HeLa cells, followed by double immunoisolation of MyoD under non-denaturing conditions. Equal amounts of chromatographed protein extracts from double-tagged MyoD and vector transfectants were then loaded on one-dimensional SDS–PAGE gels, and bands specific for MyoD transfected cells were excised from Coomassie blue stained gels. Eight peptides matching the lamin A sequence were identified by peptide maps and MS/MS protein sequencing (Table 3). These findings suggest that LMNA is part of a protein complex containing MyoD.

Discussion

Cross-sectional mRNA profiling to define biochemical pathways

mRNA profiling of human tissues using microarrays is an experimental design with many confounding, uncontrolled variables. Patient tissues are typically difficult or impossible to match for age, sex, ethnicity, stage of disease, tissue pathology or molecular diagnosis, and numbers of biopsy samples are typically limited. There have been many studies of expression profiling of cancers, where diagnostic biomarkers and key biochemical pathways in tumour progression have been identified. To our knowledge, this report presents the largest and most comprehensive data set of expression profiling in human genetic disease, and we have shown that valuable 'pathway-specific' information can be obtained from human cross-sectional studies. There are a number of key points that we believe permitted the sensitivity to detect disease-specific biochemical pathway perturbations in our study. First, it is advantageous to study a relatively large number of well-defined diagnostic groups, as this allows generic 'pathology' transcriptional changes to be cancelled out bioinformatically (e.g. necrosis, inflammation). Intrinsic to this design was the use of mutation-verified groups of patients where the primary genetic defect was known and the availability of diagnostic muscle biopsies that are appropriately processed for mRNA studies (e.g. rapid post-surgical freezing and storage). Second, quality control and standard

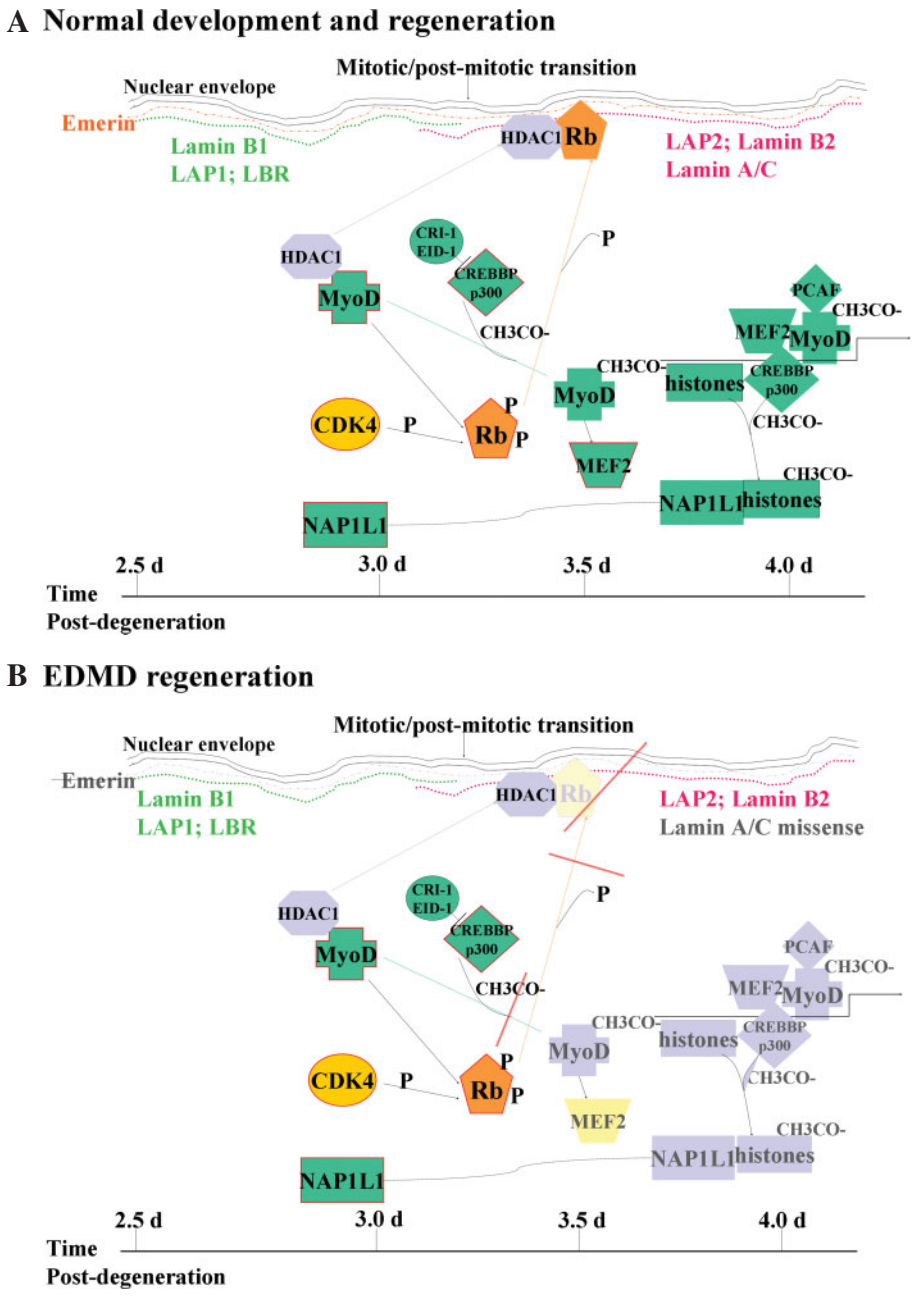


Fig. 5 (**A**) Shows the coordinate regulation of a series of proteins at the Day 3.0–3.5 time point, with remodelling of the nuclear envelope (induction of LMNA, LAP2 and lamin B2) and induction of key myogenic regulatory proteins (Rb, MyoD, CREBBP/p300, NAP1L1, MEF2, CDK4). (**B**) Shows the proposed block in this regulatory pathway, where Rb and MyoD fail to associate correctly with LMNA and emerin (owing to dominant mutations of LMNA and secondary loss of emerin, or primary loss of emerin due to X-linked recessive emerin gene mutations). The correct timing of phosphorylation and acetylation steps is perturbed, leading to a failure of downstream activation of MyoD targets. In this model, the upregulation of the pathway members (see Table 1) appears to reflect a compensatory mechanism to overcome this biochemical block. similar compensatory upregulation of the Rb–MyoD–p300–HDAC pathway is seen in staged regeneration in emerin knockout mice (Melcon et al. 2005). The higher sensitivity of the staged degeneration/regeneration also showed failure of appropriate MyoD target activation and failure of appropriate repression of E2F/Rb target genes.

operating procedures must be followed to reduce technical noise (Tumor Analysis Group, 2004). Third, use of advanced bioinformatics and statistical models, as presented here, is critical to success. Particularly important is the use of different computational methods for 'diagnosis' (VISDA) versus 'pathway building' (AHSE), as described here. Both VISDA and AHSE are founded on more robust statistical principles relative to the traditional hierarchical clustering methods. Fourth, we were fortunate to have a 'disease validation' set, in that we could use lamin A/C patients as a test data set, and emerin patients as a validation data set. In the studies presented here, the use of two different disease

Table 3 Mass spectrometry MS/MS data identifying Lamin A/C (gi|125962) as a potential MyoD interacting protein

Peptide sequence		TIC	Ions	Scans
NSNLVGAAHEELQQSR	+17	1.1E7	21/30	1864
SVGGSGGGSFGDNLVTR	+11	9.2E6	21/32	1899
LKDLEALLNSK	+18	5.3E6	17/20	2179
IDSLSAQLSQLQK	+19	3.6E6	18/24	2102
VAVEEVDEEGK	+17	5.5E6	17/20	1614
LEAALGEAK	+16	3.2E6	12/16	1635
ITESEEVVSR	+14	3.2E6	13/18	1595
QNGDDPLLTYR	+4	4.5E6	9/20	1970

genes, involving the same sub-cellular structure (nuclear envelope), and known to be interacting proteins (emerin and LMNA) provided an internal validation set for identification of potential pathway members. We feel that the approach successfully used here may serve as a model for other larger human disease data sets.

Physical and functional associations between LMNA, Rb and MyoD

We present a working model of the downstream effects of EDMD LMNA and emerin mutations on myogenic transcriptional pathways. This model was derived from a cross-sectional analysis of human muscle biopsies from mutation-positive EDMD patients and cross-referencing disease-specific transcripts to a 27 time point *in vivo* muscle regeneration temporal series.

Our model builds upon emerging data concerning interactions of MyoD, Rb and LMNA. Specifically, Rb has been shown to associate with LMNA and LAP2 (Markiewicz *et al.*, 2002), and Rb has been shown to be targeted for proteosomal degradation in LMNA null cells leading to Rb loss (Johnson *et al.*, 2004). Thus, there is ample evidence for a structural and functional association between LMNA and Rb.

There is also ample evidence for interplay of Rb and MyoD. In proliferating myoblasts, MyoD is associated with HDAC1/2, where it is inhibited from binding to downstream target promoters via hypo-acetylation (Puri *et al.*, 2001). MyoD does induce some expression of pro-differentiation Rb via an independent CREB/p300 binding complex (Magenta *et al.*, 2003), but in the presence of proliferation stimuli, cyclin/CDK complexes hyper-phosphorylate Rb, thereby inactivating it. On the basis of correlation of known Rb kinases with our *in vivo* expression profiles, we hypothesize that CDK4 is the Rb kinase relevant to muscle regeneration (Fig. 5). With the onset of differentiation cues, Rb becomes hypo-phosphorylated by p21 and other cell cycle regulators. Hypo-phosphorylated Rb has been shown in separate studies to bind to LMNA and LAP2α (Mancini *et al.*, 1994; Ozaki *et al.*, 1994; Dechat *et al.*, 2000; Markiewicz *et al.*, 2002), and to compete for HDAC1/2 binding with MyoD (Puri *et al.*, 2001). The freeing of MyoD from HDAC permits acetylation of MyoD by CREBBP/p300/PCAF (Puri *et al.*, 1997; Polesskaya *et al.*, 2000; Polesskaya *et al.*, 2001), and the establishment of functional gene transcriptional machinery with CREBBP/p300, MEF2 and PCAF complex bound to MyoD target promoters (Puri *et al.*, 2001), with CREBBP/p300 then acetylating neighbouring histones (Bergstrom *et al.*, 2002; McManus *et al.*, 2003). In parallel, at the point of myoblast exit from the cell cycle, we show that LMNA is transcriptionally induced, and is thus available to bind hypo-phosphorylated Rb (Markiewicz *et al.*, 2002). At this same time point (Day 3.0 of regeneration *in vivo*) we show that an acetylated histone chaperone, NAP1L1, is induced, facilitating removal of acetylated histones from the MyoD target promoters. Finally, we show proteomic experiments using double-tagged MyoD transfections into HeLa cells, with mass spectrometry identification of double affinity-purified MyoD complexed proteins, where we identify LMNA as one of the top-ranked MyoD binding proteins. Thus, this provides evidence that LMNA binds directly (or indirectly) to MyoD.

Our data suggests that there is a general remodelling of the nuclear envelope at the point of myoblast differentiation *in vivo*. Lamin B1, LAP1 and LBR are highly expressed at the time points of active myoblast proliferation; however, these then decline as lamin B2, LAP2 and LMNA are induced at the point of myoblast exit from the cell cycle and terminal differentiation (Day 3 post-degeneration) (Fig. 3). Emerin is ubiquitously expressed; however, an extrapolation from our data is that normal interactions between emerin and LMNA are required for appropriate MyoD–Rb interactions. Thus, if either emerin is lost (X-linked EDMD) or LMNA harbours specific dominant missense mutations (autosomal dominant EDMD), the carefully orchestrated interaction of Rb, MyoD and CREBBP/p300 acetylases becomes poorly timed or inappropriate, leading to the downstream disruption of transcriptional programming during this key stage in myogenesis.

It is important to note that phosphorylation of Rb appears to be tied to its role in cell cycle regulation, while acetylation of Rb plays a role in myogenic differentiation. Acetylation of Rb during myogenic differentiation appears to be done by the same acetylases (p300 and PCAF) responsible for acetylating MyoD (Chan *et al.*, 2001; Nguyen *et al.*, 2004). Thus, our model of transcriptional pathway perturbations induced by LMNA and emerin mutations may integrate both cell cycle abnormalities and differentiation abnormalities during myogenesis. Melcon *et al.* (2006) tested and extended our model, using staged muscle regeneration in emerin null and LMNA null mice, and validated both the Rb1–MyoD model presented here, while also extending the perturbations to failure of suppression of the Rb–E2F–HP1–SUV39 histone methylation pathway for permanent exit of the cell cycle. The role of Rb in stabilizing heterochromatic regions to permanently silence genes is increasingly well documented (Gonzalo *et al.*, 2005).

The model presented here, and further studied and expanded in the accompanying paper of Melcon *et al.* (2006), can provide a molecular explanation for some of the enigmatic features of EDMD. Patients with EDMD show an unusual early onset of contractures, where muscle growth and/or homeostasis is imbalanced between opposing muscle groups, leading to fixed deformities. Most myofibre growth during development occurs at the myotendinous junction, where myoblasts and sarcomeres are added to the ends of the growing, syncytial muscle fibre. Our model suggests that the timing of myoblast exit from cell cycle and differentiation is perturbed in EDMD patients, and this inappropriate development at the myotendinous junction might be expected to lead to a contracture phenotype. EDMD patients also show an early muscle wasting phenotype, where muscle regeneration appears inhibited. From our model, we would expect that EDMD shows less regenerative capacity than normal muscle. The last enigmatic feature of EDMD is cardiac conduction block. This is more difficult to explain, as little is known about the developmental processes leading to formation of the cardiac conduction system. This system is muscle-derived, and we hypothesize that disruption of the Rb–MyoD–LMNA pathway could lead to inappropriate development or homeostasis of the cardiac conduction system. These models are speculative, based only upon the data here, and must be proven by future targeted studies. For example, the contractures could also result from differential involvement of agonist and antagonist muscle groups. We note that Melcon *et al.* (2006) have validated the regeneration defect; however, the exact mechanisms of development and contractures and cardiac conduction defect will require future studies in the murine knockout and other models.

Our data have implications for other phenotypes associated with different mutations within the LMNA gene. Rb-related pathways have been shown to be critical for the development and homeostasis of many cell types. Rb and Rb-like proteins RBL-1 [p107] and RBL-2 [p130] are important during steps in differentiation and remodelling of muscle, heart and adipocytes (Carnac *et al.*, 2000; Guan *et al.*, 2002; Sindermann *et al.*, 2002). We found RBL-2 (p130) to be an EDMD-specific transcript, and disruption of the E2F–Rb pathways are likely to disrupt development or homeostasis of adipose, cardiac and also muscle tissues; other phenotypes of LMNA mutations are lipodystrophy, cardiac failure and cardiac conduction defects. Our data suggests that the interactions between Rb and LMNA may be quite specific, involving other additional binding proteins, with different missense mutations perturbing different pathways downstream of Rb. Consistent with this, LMNA has been recently shown to bind SREBP, a transcription factor important for adipogenesis (Lloyd *et al.*, 2002).

Finally, the results presented here, and in the follow-up functional studies of Melcon *et al.* (2006) address an important issue regarding mouse models and human disease. While the emerin-deficient mouse model described by Melcon *et al.* shows no overt clinical abnormalities, we were able to uncover abnormalities of the Rb–MyoD pathway model described here, as well as extend this model into abnormalities of the related E2F pathways through induction of degeneration/regeneration in muscle. It is often the case that mouse models of human disease possessing the same primary defect do not show clinical symptoms similar to the human disease. In other words, a genetic model is often not a clinical model. It is often hypothesized that the lack of murine phenotype is due to biochemical redundancy; for example, compensatory mechanisms are more effective in the mouse. The paper of Melcon *et al.* (2006) demonstrates this effect, where the block in LMNA–Rb–MyoD pathways due to emerin deficiency results in a dramatic compensatory upregulation of the components of the pathway, and that this compensation is indeed partly effective. The data presented here suggest that human EDMD patient muscle is also attempting to overcome the Rb–MyoD block, as we found similar upregulation of many of the members of this pathway. We hypothesize that the ability of EDMD muscle to overcome this block may deteriorate with advancing age of the patient. The shorter lifespan and relative resistance to apoptosis compared with humans may help prevent the development of pathology in the emerin-deficient mouse.

Acknowledgements

This work was supported by grants from the NIH (NINDS 3RO1 NS29525-09; NHGRI R21/R33 HG002946-01; NICHD R21 HD044891-01; NICHD 1P30HD40677-01; R01 AR48289), Department of Defense (W81XWH-04-01-0081); and donations from the Crystal Ball (Norfolk VA; Muscular Dystrophy Association), the Juvenile ALS Foundation (Erin Godla Fund; www.juvenileALS.org) and Federation to Eradicate Duchenne (FED, Washington DC; www.duchennemd.org). The authors thank Josephine Chen and Sonia Tandon for assistance with databases. This study would not have been possible without the support of the NCRR CRC at Children's National Medical Center. Microarray data is available through both NCBI GEO (http://www.ncbi.nih.gov/geo/) and CNMC PEPR (http://pepr.cnmcresearch.org).

References

Bakay M, Zhao P, Chen J, Hoffman EP. A web-accessible complete transcriptome of normal human and DMD muscle. Neuromuscul Disord 2002b; 12: S125–41.

Bengtsson L, Wilson KL. Multiple and surprising new functions for emerin, a nuclear membrane protein [review]. Curr Opin Cell Biol 2004; 16: 73–9.

Bergstrom DA, Penn BH, Strand A, Perry RL, Rudnicki MA, Tapscott SJ. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. Mol Cell 2002; 9: 587–600.

Bonne G, Di Barletta MR, Varnous S, Becane HM, Hammouda EH, Merlini L, et al. Mutations in the gene encoding LMNA cause autosomal dominant Emery-Dreifuss muscular dystrophy. Nat Genet 1999; 21: 285–8.

Brodsky GL, Muntoni F, Miocic S, Sinagra G, Sewry C, Mestroni L. LMNA gene mutation associated with dilated cardiomyopathy with variable skeletal muscle involvement. Circulation 2000; 101: 473–6.

Cao H, Hegele RA. Nuclear LMNA R482Q mutation in Canadian kindreds with Dunnigan-type familial partial lipodystrophy. Hum Mol Genet 2000; 9:109–12.

Carnac G, Fajas L, L'honore A, Sardet C, Lamb NJ, Fernandez, A. The retinoblastoma-like protein p130 is involved in the determination of reserve cells in differentiating myoblasts. Curr Biol 2000; 10: 543–6.

Chan HM, Krstic-Demonacos M, Smith L, Demonacos C, La Thangue NB. Acetylation control of the retinoblastoma tumour-suppressor protein. Nat Cell Biol 2001; 3: 667–74.

Charge SB, Rudnicki MA. Cellular and molecular regulation of muscle regeneration [review]. Physiol Rev 2004; 84: 209–38.

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; 151: 1321–36.

Chen, YW, Hubal MJ, Hoffman EP, Thompson PD, Clarkson PM. Molecular responses of human muscle to eccentric exercise. J Appl Physiol 2003; 95: 2485–94.

Chen J, Zhao P, Massaro D, Clerch LB, Almon RR, DuBois DC, et al. PEPR: an on-line query tool for microarray time series data with graphical interface. Nucleic Acids Res 2004; 32: D578–81.

Chen YW, Nagaraju K, Bakay M, McIntyre O, Rawat R, Shi R, et al. Early onset of inflammation and later involvement of TGFβ in Duchenne muscular dystrophy. Neurology 2005; 65: 826–34.

Chou FL, Angelini C, Daentl D, Garcia C, Greco C, Hausmanowa-Petrusewicz I, et al. Calpain III mutation analysis of a heterogeneous limb-girdle muscular dystrophy population. Neurology 1999; 52: 1015–20.

Dechat T, Korbei B, Vaughan OA, Vlcek S, Hutchison CJ, Foisner R. Lamina-associated polypeptide 2 alpha binds intranuclear A-type lamins. J Cell Sci 2000; 113: 3473–84.

De Sandre-Giovannoli A, Chaouch M, Kozlov S, Vallat JM, Tazir M, Kassouri N, et al. Homozygous defects in LMNA, encoding LMNA nuclear-envelope proteins, cause autosomal recessive axonal neuropathy in human (Charcot-Marie-Tooth disorder type 2) and mouse. Am J Hum Genet 2002; 70: 726–36.

De Sandre-Giovannoli A, Bernard R, Cau P, Navarro C, Amiel J, Boccaccio I, et al. Lamin A truncation in Hutchinson-Gilford progeria. Science 2003; 300: 2055–6.

Di Barletta MR, Ricci E, Galluzzi G, Tonali P, Mora M, Morandi L, et al. Different mutations in the LMNA gene cause autosomal dominant and autosomal recessive Emery-Dreifuss muscular dystrophy. Am J Hum Genet 2000; 66: 1407–12.

DiGiovanni S, Molon A, Broccolini A, Melcon G, Mirabella M, Hoffman EP, et al. Myogenic atrophy in acute quadriplegic myopathy is specifically associated with activation of pro-apoptotic TGF beta-MAPK cascade. Ann Neurol 2004; 55: 195–206.

Eng JK, McCormick AL, Yates, JRIII. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 1994; 5: 976–89.

Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, et al. Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. Nature 2003; 423: 293–8.

Fatkin D, MacRae C, Sasaki T, Wolff MR, Porcu M, Frenneaux M, et al. Missense mutations in the rod domain of the LMNA gene as causes of dilated cardiomyopathy and conduction-system disease. N Eng J Med 1999; 341: 1715–24.

Gonzalo S, García-Cao M, Fraga MF, Schotta G, Peters A, Cotter SE, et al. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. Nat Cell Biol 2005; 7: 420–8.

Gruenbaum Y, Wilson K, Harel A, Goldberg M, Cohen M. Nuclear lamins—structural proteins with fundamental functions [review]. J Struct Biol 2000; 129: 313–23.

Guan Y, Taylor-Jones JM, Peterson CA, McGehee RE Jr. p130/p107 expression distinguishes adipogenic potential in primary myoblasts based on age. Biochem Biophys Res Commun 2002; 296: 1340–5.

Holt I, Ostlund C, Stewart CL, Man N, Worman HJ, Morris GE. Effect of pathogenic mis-sense mutations in lamin A on its interaction with emerin in vivo. J Cell Sci 2003; 116: 3027–35.

Hutchison CJ, Alvarez-Reyes M, Vaughan OA. Lamins in disease: why do ubiquitously expressed nuclear envelope proteins give rise to tissue-specific disease phenotypes? J Cell Sci 2001; 114: 9–19.

Iezzi S, Cossu G, Nervi C, Sartorelli V, Puri PL. Stage-specific modulation of skeletal myogenesis by inhibitors of nuclear deacetylases. Proc Natl Acad Sci USA 2002; 99: 7757–62.

Johnson BR, Nitta RT, Frock RL, Mounkes L, Barbie DA, Stewart CL, et al. A-type lamins regulate retinoblastoma protein function by promoting subnuclear localization and preventing proteasomal degradation. Proc Natl Acad Sci USA 2004; 101: 9677–82.

Kitzmann M, Fernandez A. Crosstalk between cell cycle regulators and the myogenic factor MyoD in skeletal myoblasts. Cell Mol Life Sci 2001; 58: 571–9.

Kobberling J, Willms B, Kattermann R, Creutzfeldt W. Lipodystrophy of the extremities. A dominantly inherited syndrome associated with lipoatrophic diabetes. Humangenetik 1975; 29: 111–20.

Laguri C, Gilquin B, Wolff N, Romi-Lebrun R, Courchay K, Callebaut I, et al. Structural characterization of the LEM motif common to three human inner nuclear membrane proteins. Structure (Camb) 2001; 9: 503–11.

Lammerding J, Schulze PC, Takahashi T, Kozlov S, Sullivan T, Kamm RD, et al. LMNA deficiency causes defective nuclear mechanics and mechano-transduction. J Clin Invest 2004; 113: 370–8.

Lee KK, Haraguchi T, Lee RS, Koujin T, Hiraoka Y, Wilson KL. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. J Cell Sci 2001; 114: 4567–73.

Lee KK, Starr D, Cohen M, Liu J, Han M, Wilson KL, Gruenbaum Y. Lamin-dependent localization of UNC-84, a protein required for nuclear migration in C. elegans. Mol Biol Cell 2002; 13: 892-901.

LeRoy G, Orphanides G, Lane WS, Reinberg D. Requirements of RSF and FACT for transcription of chromatin templates in vitro. Science 1998; 282: 1900–4.

Li FQ, Coonrod A, Horwitz M. Selection of a dominant negative retinoblastoma protein (RB) inhibiting satellite myoblast differentiation implies an indirect interaction between MyoD and RB. Mol Cell Biol 2000; 20: 5129–39.

Li C, Hung Wong W. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2001; 2: 1–11.

Lloyd DJ, Trembath RC, Shackleton S. A novel interaction between lamin A and SREBP1: implications for partial lipodystrophy and other laminopathies. Hum Mol Genet 2002; 11: 769–77.

Loog M, Duin RPW, Haeb-Umbach R. Multiclass linear dimension reduction by weighted pairwise Fisher criteria. IEEE Trans Pattern Anal Machine Intell 2001; 23: 762–6.

MacLellan WR, Xiao G, Abdellatif M, Schneider MD. A novel Rb- and p300-binding protein inhibits transactivation by MyoD. Mol Cell Biol 2000; 20: 8903–15.

Magenta A, Cenciarelli C, De Santa F, Fuschi P, Martelli F, Caruso M, et al. MyoD stimulates RB promoter activity via the CREB/CREBBP/p300 nuclear transduction pathway. Mol Cell Biol 2003; 23: 2893–906.

Malone CJ, Misner L, Le Bot N, Tsai MC, Campbell JM, Ahringer J, et al. The C. elegans Hook protein, ZYG-12, mediates the essential attachment between the centrosome and nucleus. Cell 2003; 115: 825-36.

Mancini MA, Shan B, Nickerson JA, Penman S, Lee WH. The retinoblastoma gene product is a cell cycle-dependent, nuclear matrix-associated protein. Proc Natl Acad Sci USA 1994; 91: 418–22.

Maraldi NM, Squarzoni S, Sabatelli P, Lattanzi G, Ognibene A, Manzoli FA. Emery-Dreifuss muscular dystrophy, nuclear cell signaling and chromatin remodeling [review]. Adv Enzyme Regul 2002; 42: 1–18.

Markiewicz E, Dechat T, Foisner R, Quinlan RA, Hutchison CJ. LMNA binding protein LAP2alpha is required for nuclear anchorage of retinoblastoma protein. Mol Biol Cell 2002; 13: 4401–13.

Masny PS, Bengtsson U, Chung SA, Martin JH, Van Engelen B, Van Der Maarel SM, et al. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? Hum Mol Genet 2004; 13: 1857–71.

McManus KJ, Hendzel MJ. Quantitative analysis of CBP- and P300-induced histone acetylations in vivo using native chromatin. Mol Cell Biol 2003; 23: 7611–27.

Melcon G, Kozlov S, Cutler DA, Sullivan T, Hernandez L, Zhao P, et al. Loss of emerin at the nuclear envelope disrupts the Rb1/E2F and MyoD pathways during muscle regeneration. Hum Mol Genet. 2006; 15: 637-51.

Molon A, Di Giovanni S, Chen YW, Clarkson PM, Angelini C, Pegoraro E, et al. Large-scale disruption of microtubule pathways in morphologically normal human spastin muscle. Neurology 2004; 62: 1097–104.

Mounkes LC, Kozlov S, Hernandez L, Sullivan T, Stewart CL. A progeroid syndrome in mice is caused by defects in A-type lamins. Nature 2003; 423: 298–301.

Muchir A, Bonne G, van der Kooi AJ, van Meegen M, Baas F, Bolhuis PA, et al. Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). Hum Mol Genet 2000; 9: 1453–9.

Navarro CL, De Sandre-Giovannoli A, Bernard R, Boccaccio I, Boyer A, Genevieve D, et al. Lamin A and ZMPSTE24 (FACE-1) defects cause nuclear disorganization and identify restrictive dermopathy as a lethal neonatal laminopathy. Hum Mol Genet 2004; 13: 2493–503.

Nevo Y, Ahituv S, Yaron Y, Kedmi M, Shomrat R, Legum C, et al. Novel mutations in the emerin gene in Israeli families. Hum Mutat 2001; 17: 522.

Nguyen DX, Baglia LA, Huang SM, Baker CM, McCance DJ. Acetylation regulates the differentiation-specific functions of the retinoblastoma protein. EMBO J 2004; 23: 1609–18.

Novelli G, Muchir A, Sangiuolo F, Helbling-Leclerc A, D'Apice MR, Massart C, et al. Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding LMNA. Am J Hum Genet 2002; 71: 426–31.

Östlund C, Worman HJ. Nuclear envelope proteins and neuromuscular diseases [review]. Muscle Nerve 2003; 27: 393–406.

Ozaki T, Saijo M, Murakami K, Enomoto H, Taya Y, Sakiyama S. Complex formation between lamin A and the retinoblastoma gene product: identification of the domain on lamin A required for its interaction. Oncogene 1994; 9: 2649–53.

Pegoraro E, Mancias P, Swerdlow SH, Raikow RB, Garcia C, Marks H, et al. Congenital muscular dystrophy (CMD) with primary laminin α2 deficiency presenting as inflammatory myopathy. Ann Neurol 1996; 40: 782–91.

Peschiaroli A, Figliola R, Coltella L, Strom A, Valentini A, D'Agnano I, et al. MyoD induces apoptosis in the absence of RB function through a p21(WAF1)-dependent re-localization of cyclin/cdk complexes to the nucleus. Oncogene 2002; 21: 8114–27.

Polesskaya A, Duquet A, Naguibneva I, Weise C, Vervisch A, Bengal E, et al. CREB-binding protein/p300 activates MyoD by acetylation. J Biol Chem 2000; 275: 34359–64.

Polesskaya A, Naguibneva I, Duquet A, Bengal E, Robin P, Harel-Bellan A. Interaction between acetylated MyoD and the bromodomain of CBP and/or p300. Mol Cell Biol 2001; 21: 5312–20.

Puri PL, Sartorelli V, Yang XJ, Hamamori Y, Ogryzko VV, Howard BH, et al. Differential roles of p300 and PCAF acetyltransferases in muscle differentiation. Mol Cell 1997; 1: 35–45.

Puri PL, Iezzi S, Stiegler P, Chen TT, Schiltz RL, Muscat GEO, et al. Class I histone deacetylases sequentially interact with MyoD and pRb during skeletal myogenesis. Mol Cell 2001; 8: 885–97.

Puri PL, Bhakta K, Wood LD, Costanzo A, Zhu J, Wang JY. A myogenic differentiation checkpoint activated by genotoxic stress. Nat Genet 2002; 32: 585–93.

Raharjo WH, Enarson P, Sullivan T, Stewart CL, Burke B. Nuclear envelope defects associated with LMNA mutations cause dilated cardiomyopathy and Emery-Dreifuss muscular dystrophy. J Cell Sci 2001; 114: 4447–57.

Seo J, Bakay M, Zhao P, Chen YW, Clarkson P, Shneiderman B, et al. Interactive color mosaic and dendrogram displays for signal/noise optimization in microarray data analysis. IEEE ICME 2003; III: 461–2.

Seo J, Bakay M, Chen YW, Hilmer S, Shneiderman B, Hoffman EP. Interactively optimizing signal-tonoise ratios in expression profiling: project-specific algorithm selection and detection p-value weighting in Affymetrix microarrays. Bioinformatics 2004; 20: 2534–44.

Shikama N, Chan HM, Krstic-Demonacos M, Smith L, Lee CW, Cairns W, et al. Functional interaction between nucleosome assembly proteins and p300/CREB-binding protein family coactivators. Mol Cell Biol 2002; 20: 8933–43.

Shumaker DK, Lee KK, Tanhehco YC, Craigie R, Wilson KL. LAP2 binds to BAF-DNA complexes: requirement for the LEM domain and modulation by variable regions. EMBO J 2001; 20: 1754–64

Simha V, Garg A. Body fat distribution and metabolic derangements in patients with familial partial lipodystrophy associated with mandibuloacral dysplasia. J Clin Endocrinol Metab 2002; 87: 776–85.

Simha V, Agarwal AK, Oral EA, Fryns J-P, Garg A. Genetic and phenotypic heterogeneity in patients with mandibuloacral dysplasia-associated lipodystrophy. J Clin Endocrinol Metab 2003; 88: 2821–4.

Simone C, Forcales SV, Hill DA, Imbalzano AN, Latella L, Puri PL. p38 pathway targets SWI-SNF chromatin-remodeling complex to muscle-specific loci. Nat Genet 2004; 36: 738–43.

Sindermann JR, Smith J, Kobbert C, Plenz G, Skaletz-Rorowski A, Solomon JL, et al. Direct evidence for the importance of p130 in injury response and arterial remodeling following carotid artery ligation. Cardiovasc Res 2002; 54: 676–83.

Starr DA, Hermann GJ, Malone CJ, Fixsen W, Priess JR, Horvitz HR, et al. unc-83 encodes a novel component of the nuclear envelope and is essential for proper nuclear migration. Development 2001; 128: 5039–50.

Starr DA, Han M. Role of ANC-1 in tethering nuclei to the actin cytoskeleton. Science 2002; 11: 406–9.

Sullivan T, Escalante-Alcalde D, Bhatt H, Anver M, Bhat N, Nagashima K, et al. Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. J Cell Biol 1999; 147: 913–20.

Tezak Z, Hoffman EP, Lutz J, Fedczyna T, Stephan D, Bremer EG, et al. Gene expression profiling in DQA1*0501+ children with untreated dermatomyositis: a novel model of pathogenesis. J Virol 2002; 168: 4154–63.

Tumor Analysis Best Practices Working Group. Expression profiling—best practices for Affymetrix array data generation and interpretation in clinical trials. Nat Rev Genet 2004; 5: 229–37.

Vaughan O, Alvarez-Reyes M, Bridge J, Broers J, Ramaekers F, Wehnert M, et al. Both emerin and lamin C depend on lamin A for localization at the nuclear envelope. J Cell Sci 2001; 114: 2577–90.

Wang Y, Luo L, Freedman MT, Kung SY. Probabilistic principal component subspaces: a hierarchical finite mixture model for data visualization. IEEE Trans Neural Networks 2000; 11: 625–36.

Wang X, Xu S, Rivolta C, Li LY, Peng GH, Swain PK, et al. Barrier to autointegration factor interacts with the cone-rod homeobox and represses its transactivation function. J Biol Chem 2002; 277: 43288–300.

Wang Z, Wang Y, Lu J, Kung SY, Zhang J, Lee R, et al. Discriminatory mining of gene expression microarray data. J VLSI Signal Process Sys 2003; 35: 255–72.

Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, et al. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. Hum Mol Genet 2003; 12: 2895–907.

Wolff N, Gilquin B, Courchay K, Callebaut I, Worman HJ, Zinn-Justin S. Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery-Dreifuss muscular dystrophy. FEBS Lett 2001; 501: 171–6.

Xuan J, Dong Y, Khan J, Clarke R, Hoffman E, Wang Y. 'Gene selection by weighted Fisher criterion for multiclass prediction'. Proc Int Conf Bioinformatics and Applications (ICBA2004), Fort Lauderdale, Florida, Dec. 2004.

Xuan J, Dong Y, Khan J, Hoffman E, Clarke R, Wang Y. Robust gene selection by weighted Fisher criterion for multiclass prediction in gene expression profiling. Int Conf Pattern Recognition, Vol. 2, Cambridge, UK; 2004. p. 291–4.

Zastrow MS, Vlcek S, Wilson KL. Proteins that bind A-type lamins: integrating isolated clues [review]. J Cell Sci 2004; 117: 979–87.

Zhao P, Iezzi S, Sartorelli V, Dressman D, Hoffman EP. Slug is downstream of myoD: identification of novel pathway members via temporal expression profiling. J Biol Chem 2002; 277: 20091–101.

Zhao P, Seo J, Wang Z, Wang Y, Shneiderman B, Hoffman EP. In vivo filtering of in vitro MyoD target data: an approach for identification of biologically relevant novel downstream targets of transcription factors. C R Biol 2003; 326: 1049–65.

Zhao P, Hoffman EP. Myogenesis pathways in muscle regeneration. Dev Dyn 2004; 229: 380–92.

Zhao P, Caretti G, Mitchell S, McKeehan WL, Boskey AL, Pachman LM, et al. Fgfr4 is required for effective muscle regeneration in vivo: delineation of a MyoD-Tead2-Fgfr4 transcriptional pathway. J Biol Chem 2006; 281: 429–38.