Global regulation of alternative splicing during myogenic differentiation

Christopher S. Bland[1,2], Eric T. Wang[3], Anthony Vu[1], Marjorie P. David[4], John C. Castle[5], Jason M. Johnson[6], Christopher B. Burge[3] and Thomas A. Cooper[1,2,7,*]

[1]Department of Pathology and Immunology, [2]Interdepartmental Program in Cell and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, [4]University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA, [5]Institute for Translational Oncology and Immunology, Mainz, Germany 155131, [6]Merck Research Laboratories, Boston, MA 02115 and [7]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

ABSTRACT

Recent genome-wide analyses have elucidated the extent of alternative splicing (AS) in mammals, often focusing on comparisons of splice isoforms between differentiated tissues. However, regulated splicing changes are likely to be important in biological transitions such as cellular differentiation, or response to environmental stimuli. To assess the extent and significance of AS in myogenesis, we used splicing-sensitive microarray analysis of differentiating C2C12 myoblasts. We identified 95 AS events that undergo robust splicing transitions during C2C12 differentiation. More than half of the splicing transitions are conserved during differentiation of avian myoblasts, suggesting the products and timing of transitions are functionally significant. The majority of splicing transitions during C2C12 differentiation fall into four temporal patterns and were dependent on the myogenic program, suggesting that they are integral components of myogenic differentiation. Computational analyses revealed enrichment of many sequence motifs within the upstream and downstream intronic regions near the alternatively spliced regions corresponding to binding sites of splicing regulators. Western analyses demonstrated that several splicing regulators undergo dynamic changes in nuclear abundance during differentiation. These findings show that within a developmental context, AS is a highly regulated and conserved process, suggesting a major role for AS regulation in myogenic differentiation.

INTRODUCTION

Current estimates are that ~95% of multi-exon genes in humans are subject to alternative splicing (AS), greatly expanding the transcriptome (1). AS also serves a crucial regulatory role by altering the function, localization and expression level of gene products, often in response to the activities of key signaling pathways (2–5). Misregulation of AS is implicated in the pathogenic mechanisms of several diseases (6–9). Splicing regulatory proteins are subject to multiple levels of regulation during development (10–12) and AS regulation has been shown to occur during a number of developmental processes including heart development (13), neurogenesis (14–16) and T-cell differentiation (17). Despite increased recognition of the prevalence of AS and its relevance to development, tissue identity and disease, little is known about the mechanisms that regulate natural splicing transitions. In addition, the broad biological relevance of the extensive transcript diversity generated by AS continues to be debated (18–21).

Recent efforts to examine splicing on a global scale using high-throughput techniques such as splicing sensitive microarrays (22,23) and deep sequencing (1,24,25) have focused primarily on comparing splicing patterns in adult tissues or examining events affected by depletion of *trans* acting factors. Other studies have used purely computational approaches to ascertain the global impact of AS (20,26), often relying on EST databases, which are heavily biased towards transcripts derived from brain and cancer tissues (27,28). By restricting global AS analyses to adult tissues, temporally regulated aspects of AS biology are overlooked. Analysis of global AS transitions during key biological transitions such as development provides an experimental system in which to identify the regulatory mechanisms and biological relevance of AS.

---

*To whom correspondence should be addressed. Tel: 713 798 3141; Fax: 713 798 5838; Email: tcooper@bcm.edu

AS is enriched in skeletal muscle (22), as are several splicing factors, such as the FOX and Muscleblind-like (MBNL) families (29,30), suggesting that myogenesis is accompanied by high levels of AS regulation. The C2C12 mouse myoblast cell line is a subclone of a cell line derived from adult muscle satellite cells (31,32). The cells are committed to the myogenic pathway and are highly proliferative when maintained in high serum/low confluence conditions. Exposing confluent C2C12 cells to low serum conditions induces differentiation. Cultures up-regulate the myogenic transcription factor myogenin within 24 h, exit the cell cycle within 36 h and myoblasts fuse within 72 h to form multinucleated myotubes that exhibit morphological and biochemical similarities to immature skeletal muscle tissue (Figure 1A) (33,34).

Our goal was to utilize myogenic differentiation as a model system to study developmentally-associated AS regulation. We specifically set out to define networks of AS transitions that occur during myogenic differentiation and to identify *cis*-acting elements and *trans*-acting factors that control these regulatory networks. Using splicing sensitive microarrays we identified ~100 alternatively spliced regions that undergo robust splicing transitions during C2C12 differentiation. Most of the strongly regulated splicing transitions tested are conserved during differentiation of mammalian and avian myoblasts and are dependent on the myogenic program. Computational analysis identified multiple sequence motifs that are significantly enriched and conserved within the intronic regions surrounding the regulated alternative regions. Many of these motifs were significantly associated with specific temporal subsets of splicing transitions. The identified motifs included binding sites for known splicing regulators such as FOX, CELF, hnRNP L, TIA and PTB. A large number of novel sequence motifs were also enriched near regulated regions. Furthermore, many of the splicing regulators with enriched and conserved binding motifs exhibit changes in nuclear abundance during C2C12 differentiation, either due to apparent nuclear/cytoplasmic transitions (MBNL2 and MBNL3) or changes in total steady-state levels (MBNL1, CUGBP1 and CUGBP2, PTB, FOX1 and hnRNP C), while other splicing regulators exhibit relatively constant nuclear steady-state levels (FOX2). These results reveal regulatory networks controlling natural splicing transitions that are likely to be directly relevant to physiological changes associated with muscle differentiation.

MATERIALS AND METHODS

Cell culture

C2C12 myoblasts (ATCC) were maintained in high-glucose DMEM, supplemented with 10% FBS, 1% penicillin/streptomycin and 1% L-glutamine (all Invitrogen). C2C12 cells were differentiated at ~95% confluence by adding differentiation media containing high-glucose DMEM, supplemented with 2% horse serum (Invitrogen), 1% penicillin/streptomycin and 1% L-glutamine. Low-passage (<10 passages) cells were used for all experiments. QM7 cells (ATCC) (35) were maintained in Earle's M199 basal medium, supplemented with 10% fetal calf serum, 10% Tryptose phosphate broth (both Difco), 1% penicillin/streptomycin and 1% L-glutamine. QM7 cells were differentiated at ~95% confluence in Earle's M199 basal medium, 0.01% fetal calf serum, 1% penicillin/streptomycin and 1% L-glutamine. QM7 cells were allowed to differentiate for 96 h, at which point large numbers of myotubes were visible. All cell lines were maintained in a humidified, 37°C incubator with 5% CO2.

BDM treatment

A 5M stock solution of 2,3, butanedione monoxime (BDM; Sigma) dissolved in DMSO was added to C2C12 differentiation media to a final concentration of 15 µM, (or an equal volume of DMSO; vehicle control). The treated media was then added to ~95% confluent C2C12 cells (Hour 0 of differentiation). Fresh treated differentiation media was added to cultures every two days. C2C12 cells were allowed to differentiate for up to 120 h. RNA and whole-cell protein lysate were collected as described below at −24, 0, 24, 72 and 120 relative to differentiation induction.

Splicing sensitive microarray, RNA preparation and hybridization

Splicing arrays were designed and RNA isolation and hybridization performed as previously described (13). Briefly, the array design included probes to 10 111 genes expressed in muscle or heart. Optimized 60-nt probes were designed to monitor each exon, and 36-nt probes were centered across each exon–exon junction (36). Exons <60 nt were monitored by using shorter probes attached to T-stilts. In total, 248 316 oligonucleotide probes, including 110 367 exon and 93 382 junction probes were synthesized on a six-array set. Total RNA for splicing microarray analysis was prepared from C2C12 cultures at −24 and 120 h relative to differentiation induction using RNeasy kits (Qiagen), and the RNA quality was assessed spectrophotometrically (Agilent) at the microarray core facility at Baylor College of Medicine. Cy dye labeling, hybridization and array analysis was performed as described earlier (13).

RT–PCR validation and quantification of splicing transitions

For RT–PCR assays, total RNA was extracted using TRIzol (Invitrogen). RT–PCR primers were designed complementary to constitutive exonic regions flanking the predicted alternative regions. One-Step RT–PCR reactions were performed using specific primers on 0.5 µg of total RNA samples isolated from C2C12 or QM7 cells at various time points throughout differentiation. After gene-specific reverse transcription at 42°C for 35 min, the resulting cDNA was subjected to 18–25 cycles of 95°C for 45 s, 57°C for 45 s and 72°C for 60 s, then separated using PAGE, followed by ethidium bromide staining, imaging and molecular weight-corrected quantification using the Kodak Gel logic 2200 and Molecular Imaging Software. Alternative region percent inclusion was quantified by
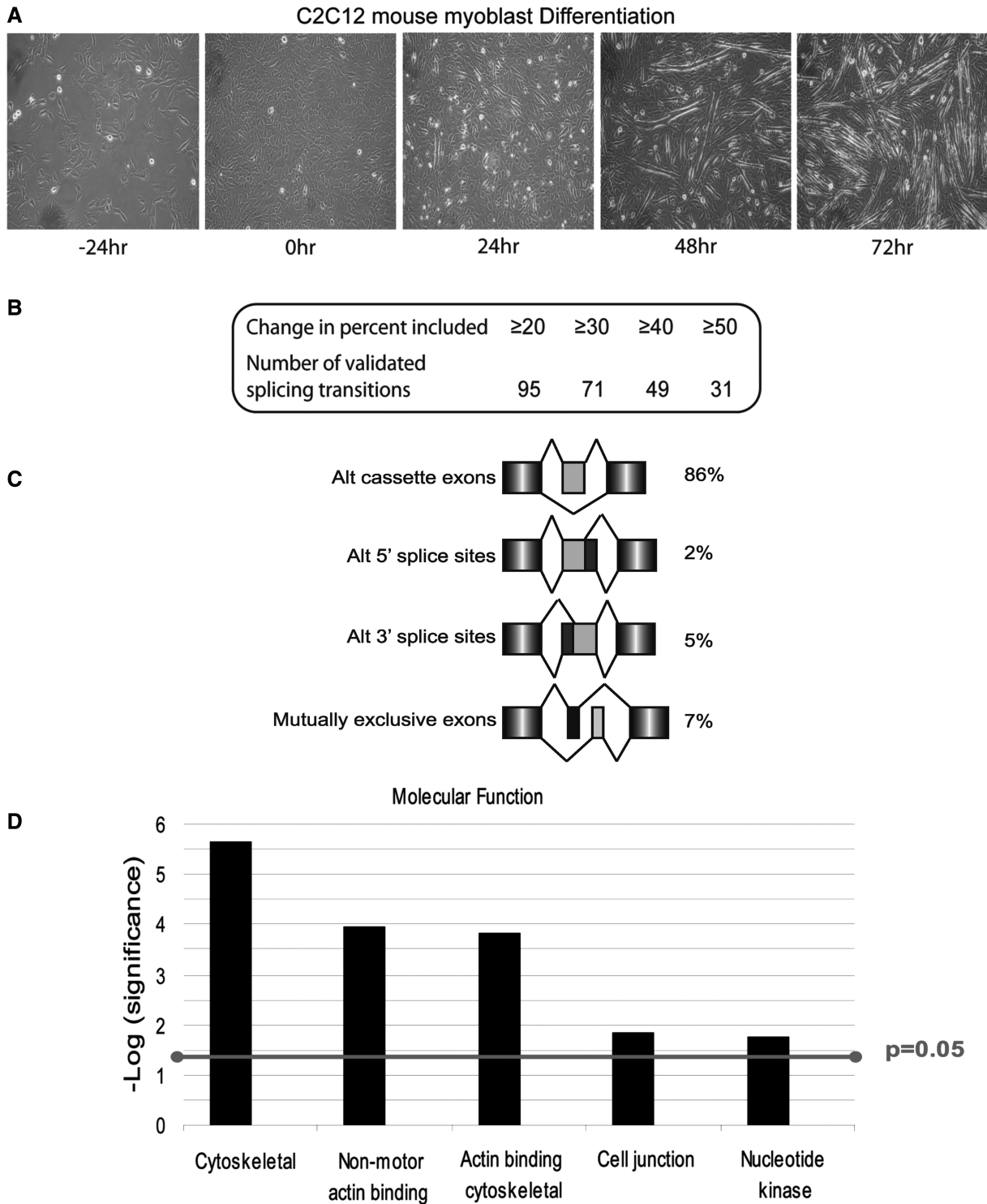
Figure 1. Characterization of validated splicing transitions associated with C2C12 myoblast differentiation. (**A**) Phase-contrast micrographs showing a time course of C2C12 differentiation. (**B**) The number of splicing events (out of 117 total validated events) that undergo splicing transitions of ≥20 percentage points. (**C**) Summary of the different types of validated splicing transitions included within the data set of 117 splicing transitions. (**D**) GO analysis for significantly ($P \leq 0.05$) enriched molecular functions in validated splicing transitions of ≥20 percentage points.

adjusting band intensity for length of the PCR product, dividing the intensity the band including the regulated alternative region by the total product [inclusion band/(inclusion + exclusion band) ×100]. All bands of interest were gel-isolated, amplified and confirmed by sequencing.

Gene ontology analysis

Gene ontology (GO) analysis was performed using PANTHER according to the instructions provided at www.pantherdb.org/. Validated splicing transitions from C2C12 or QM7 data sets were analyzed against a reference

list (10 111 genes on the microarray) for significant enrichment of molecular function terms, calculated using binomial testing with Bonferroni correction for multiple testing as described earlier (37) Only processes with a significant ($P \leq 0.05$) enrichment were discussed in this article.

Computational analysis of motif enrichment and conservation

Significantly enriched pentameric motifs within flanking introns of AS regions undergoing splicing transitions $\geq 10\%$ across all time points were detected according to a GC-content controlled first-order Markov model. The intronic regions analyzed consist of the first and last 250 nt of the upstream and downstream introns, relative to the regulated alternative region, excluding the first 9 and last 30 nt of the intron which contain the conserved 5′ and 3′ splice sites, respectively. Sequences were binned according to their G+C content into 10 groups. Expected pentamer frequency was calculated for each pentamer by using the first-order Markov model in introns of each GC group respectively. Pentamer enrichment was then evaluated by comparing the occurrence frequency of each pentamer to the overall expected frequency calculated by summing up the expected counts of all G+C groups. A P-value was calculated by using the binomial distribution. All results were subjected to false discovery rate filter (Benjamini–Hochberg, FDR < 0.05). Significantly conserved pentameric motifs among splicing transitions changing $\geq 10\%$ across all time points were identified within the above intronic regions using alignments to seven other mammalian genomes that have at least $5\times$ sequence coverage in the UCSC 28-way multi-genome alignment. A conservation rate (CR) was calculated as the fraction of aligned and conserved motif occurrences. The significance of the CR for each pentamer was evaluated by comparing to 10 other pentamers with similar expected CR, according to a first-order Markov model. P-values were obtained using binomial distribution, and subjected to a false discovery rate filter (Benjamini–Hochberg, FDR < 0.05).

Motif regression analysis

Regression analysis was performed on motifs determined to be significantly enriched or conserved, discovered using introns flanking events which change $\geq 10\%$ across all time points, events which increase $\geq 10\%$ across all time points, and events which decrease $\leq 10\%$ across all times points. The percent of alternative region inclusion at each time point (i) of a C2C12 differentiation time course ($-24, 0, 12, 24, 72$ and 120 h relative to differentiation induction) was compared to every other time point (j), to obtain different vectors of splicing changes (y_{ij}), where the length of the vector was proportional to the number of splicing events within the validated data set. For each of the above mentioned intronic regions (r) flanking each regulated splicing event, the density of motifs (x_r) was regressed against each vector of splicing changes ($y_{ij} = A x_r + b$). Significant correlations ($P < 0.05$) were noted and plotted in a heat map format (Supplementary Figure S1) in which colors denote $-\log 10$

(P-value) for P-values <0.05, with darker colors represent increasing significance. Red shades denote a positive coefficient (A), where presence of the motif is associated with an increase in alternative region inclusion ($\Psi_{Time2} - \Psi_{Time1} > 0$), and blue shades denote a negative coefficient (A), where presence of the motif is associated with an increase in alternative region skipping ($\Psi_{Time2} - \Psi_{Time1} < 0$).

Clustering analysis

Affinity propagation analysis was performed to identify similar clusters of splicing transitions, (38) using a matrix of all pair wise comparisons between all splicing events. To obtain the similarity matrix, the Pearson correlation coefficients was computed between the raw inclusion levels of regulated alternative regions for each pair of splicing events across time points $-24, 0, 12, 24, 48, 72$ and 120 of C2C12 differentiation. Various affinity propagation preference levels were tested; preference values of 0.25–0.5 yielded the same number of clusters, and these clusters are shown in Figure 6 and Supplementary Table S5. Clustering graphs illustrate splicing profiles of a cluster of splicing events across time points; the y-axis represents inclusion levels at various time points, normalized so that means and SDs are uniform. Each cluster contains an 'exemplar' (shown in red) which best 'exemplifies' the cluster, according to the affinity propagation algorithm. Splicing profiles of the other events in each cluster are plotted in various shades of gray, where darker gray represents a larger maximum difference in inclusion level across all time points. Motif enrichment and conservation analysis was performed for the events within each cluster in a manner identical as that described earlier.

Nuclear, cytoplasmic and whole-cell protein extraction

Nuclear and cytoplasmic protein fractions were prepared from C2C12 cells by washing cells three times in ice cold PBS, scraping in fractionation buffer [10 mM HEPES (pH 7.5), 10 mM MgCl2, 5 mM KCl, 0.1 mM EDTA, 0.1% Triton X-100, 0.2 mM PMSF, 1 mM DTT, $1\times$ Complete Protease Inhibitor Mixture (Roche)] and passing cells through a 27 gauge needle 10 times. The resulting lysate was then centrifuged at 600g for 15 min at 4°C to pellet nuclei. The cytoplasmic fraction was acetone precipitated overnight and resuspended in lysis buffer [10 mM HEPES (pH 7.5), 0.32 M sucrose, 1% SDS, 5 mM MG132 and 5 mM EDTA with protease inhibitors]. The nuclear pellet was washed two times in fractionation buffer, and then resuspended in lysis buffer. Both nuclear and cytoplasmic fractions were sonicated and total protein was quantified using the BCA assay (Pierce). All samples were stored at -80°C.

Western blot analysis

Thirty micrograms of nuclear, cytoplasmic or total protein was fractionated on a 10% Tris–glycine SDS–PAGE gels and transferred onto PVDF membranes (Immobilon). The protein loading and transfer quality were confirmed by Ponceau S red staining (Sigma). Membranes were blocked in TBST with 5% non-fat milk, then incubated

with the primary antibody overnight at 4°C, washed in TBST, incubated with the appropriate secondary HRP-conjugated antibody at room temperature for 2 h, washed in TBST and visualized using an HRP chemiluminescence system (Pierce). The following primary antibodies were used: monoclonal anti-CUGBP1 antibody (1:1000; 3B1; Upstate), monoclonal anti-CUGBP2 [1:500; 1H2 (39) conjugated with HRP], polyclonal anti-MBNL1 [1:1000 (40)], monoclonal anti-MBNL2 (1:1000; and monoclonal anti-MBNL3 (1:1000; both generously provided by Glenn E. Morris) (41), monoclonal anti-FOX-1 [1:500; generated in-house (13)] polyclonal anti-FOX-2 [1:120 000; generated in-house (13)], monoclonal anti-PTB (1:1000; R164; generously provided by Doug Black), monoclonal anti-Myogenin (1:500; F5D Abcam), monoclonal anti-hnRNP L (1:50 000; 4D11; Abcam), monoclonal anti-hnRNP C (1:10 000; Sigma), polyclonal Anti-p27 (1:500; sc-518; Santa Cruz) monoclonal anti-TBP (1:1000; ab818; Abcam), and monoclonal anti-GAPDH (1:10 000; Biogenesis). Secondary antibodies used were HRP-conjugated goat anti-rabbit (1:50 000; Calbiochem) and HRP-conjugated sheep anti-mouse (1:5000; Jackson).

RESULTS

Skeletal muscle differentiation is associated with a large number of AS transitions

AS transitions during C2C12 differentiation were identified using a previously described custom splicing sensitive microarray (13) as well as from the literature focusing on exons found to be preferentially included in adult mouse skeletal muscle tissue (23,42). Overall, 117 splicing transitions were identified, 66 of which were detected using splicing-sensitive microarrays, and the remaining events were found by performing EST comparisons and through the literature. All events were validated by RT–PCR using total RNA collected from C2C12 cells 24 h before differentiation induction (−24 h) and 5 days post differentiation induction (120 h). The identities of the predominant (≥10% of total) RT–PCR products were confirmed by sequencing. To define the kinetics of splicing transitions, validated events were further analyzed by RT–PCR in a series of time course experiments, using total RNA collected at −24, 0, 12, 24, 72 and 120 h relative to differentiation induction (Supplementary Table S1). AS is expressed as the percent of mRNA that includes the variable region and splicing transitions are expressed as the change of percent of mRNAs containing the variable region between time points; thus an event that changes from 20 to 50% inclusion is a 30% point change. At least two biological replicates were performed for all time points of all events analyzed in the time course. In addition, all events were tested in at least two additional biological replicates at the −24 and 120 time points. The magnitude of most splicing transitions was highly consistent between biological replicates.

Of the 117 validated splicing events, 95 exhibited strong transitions of ≥20 percentage points with 49 of these exhibiting a change of 40 percentage points or more (Figure 1B and Supplementary Table S1). Sixteen events changed between 10 and 20 percentage points and six events displayed changes <10 percentage points. Of the 95 strong events, the average alternative region length was 96 nt, with a SD of 49 nt. A total of 69 of the 95 strong events (73%) showed increased inclusion during differentiation, and the remaining 26 (27%) showed increased skipping. Among all 117 validated transitions, there were 100 cassette exons (85%), including 16 multiple cassette exons, 8 mutually exclusive exons (7%) and 9 alternative 3′ or 5′ splice sites (7%) (Figure 1C). Ninety-six of these 117 events (82%) are predicted to retain the reading frame. In addition, 38 of the 117 events (32%) underwent an average fold change ≥2.0 in gene expression.

To assess the biological roles of the transcripts undergoing AS transitions during C2C12 myoblast differentiation, we performed comparative GO analysis (37,43). The 95 strong events were compared against a reference set of ~10 000 transcripts expressed in mouse striated muscle. Genes expressing transcripts undergoing AS transitions were significantly enriched for GO molecular function terms relating to cytoskeletal, actin binding, cell junction and nucleotide kinase molecular functions (Figure 1D), as well as a significant ($P = 0.021$) over-representation in integrin signaling pathway components, consistent with other muscle specific splicing data sets (23,42). Overall, this data set represents a large cohort of validated, robustly regulated splicing transitions that occur in several temporally coordinated patterns (see below) during myogenic differentiation.

AS transitions associated with muscle differentiation are conserved in mammalian and avian cultures

To asses the biological relevance of the splicing transitions during muscle differentiation, we examined their conservation between the murine C2C12 line and the quail myoblast cell line QM7 (35). From a set of 77 strongly regulated alternative exons from the validated data set, we identified 48 orthologous exons in the published chicken genome (build 2.1). A total of 38 (79%) of these were detected by RT–PCR (Figure 2A and Supplementary Table S2) of which 22 (57%) were regulated similarly to mouse, three (8%) were oppositely regulated and 13 (34%) were alternatively spliced but showed little or no regulation during QM7 differentiation (Figure 2B). We used GO analysis to define the molecular functions of the 25 genes containing exons regulated in both quail and mouse and found that 7 (30%) were within transcripts encoding RNA-binding proteins, a statistically significant enrichment ($P = 1.39 \times 10^{-4}$), while 6 (24%) were within transcripts encoding cytoskeletal components. The high level of conservation of regulated AS transitions in avian and mammalian muscle differentiation strongly suggest that AS regulation plays a biologically significant role and that RNA-binding proteins are conserved targets of splicing regulation.
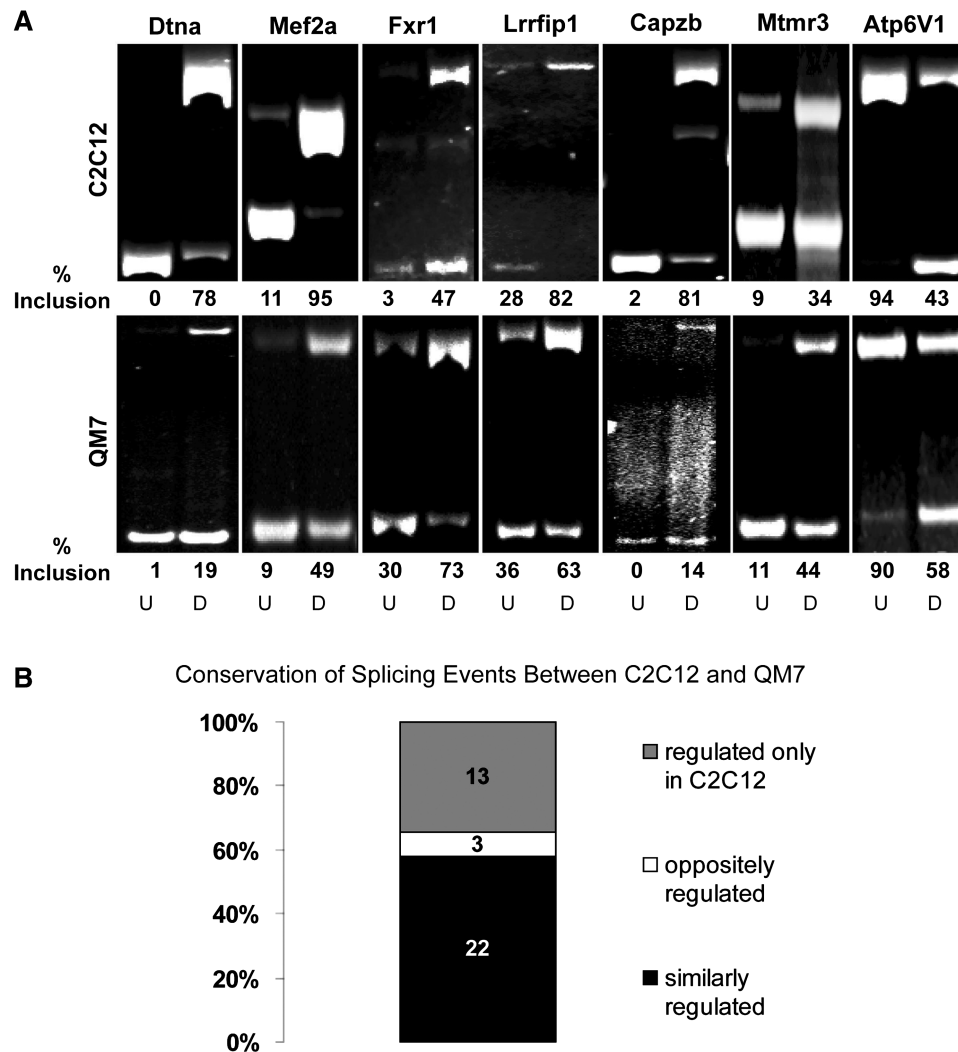
Figure 2. Most AS events common to mouse and quail myoblasts undergo conserved transitions during differentiation. (**A**) Representative RT–PCR results for seven splicing transitions conserved between C2C12 (mouse) and QM7 (quail) myoblast differentiation. Undifferentiated (U) and differentiated (D) cultures are indicated. The percentage alternative exon inclusion (% inclusion) is indicated in undifferentiated and differentiated C2C12 and QM7 cells. (**B**) Summary of splicing transitions conservation between C2C12 and QM7 cells. See Supplementary Table S2 for a full list of all events tested between C2C12 and QM7 cells.

The majority of regulated splicing transitions in C2C12 are dependent on the myogenic differentiation program

Myogenic differentiation in culture is induced by a combination of reduced growth factors and high cell density and is preceded by cell-cycle exit (33,44,45). To determine whether AS regulation is an integrated component of the myogenic program or simply a result of growth factor withdrawal and cell-cycle exit, we used 2,3, butanedione monoxime (BDM) to induce cell-cycle exit and simultaneously block myogenic differentiation (46). As expected, C2C12 cultures treated with 15 μM BDM failed to undergo morphological changes associated with differentiation, including cell alignment and formation of myotubes (Figure 3A). In addition, BDM-treated cultures showed no increase in cell number over time (data not shown), exhibited upregulated expression of the cell-cycle inhibitor p27, and failed to upregulate expression of the myogenic transcription factor myogenin

(Figure 3B). We examined 21 splicing events that undergo strong (≥20 percentage points) differentiation-induced splicing transitions in cells treated at time 0 with BDM compared to cells treated with the DMSO vehicle alone. The majority of these events (13) showed little or no change (<25% of that seen in control cells) (representative events are shown in Figure 3Ci and 3Cii). Of particular interest are the several splicing transitions that occur predominantly between −24 and 0 h of differentiation. These events are initiated much earlier than the appearance of transcriptional markers of myogenesis or overt morphological changes indicative of differentiation. Most of these events revert toward the level of inclusion observed at −24 h when BDM is added at time 0, suggesting dependence on the myogenic differentiation program despite the transition initiating before addition of differentiation media (Figure 3Cii). Four events showed a moderate degree of regulation in the presence of BDM (>25% but

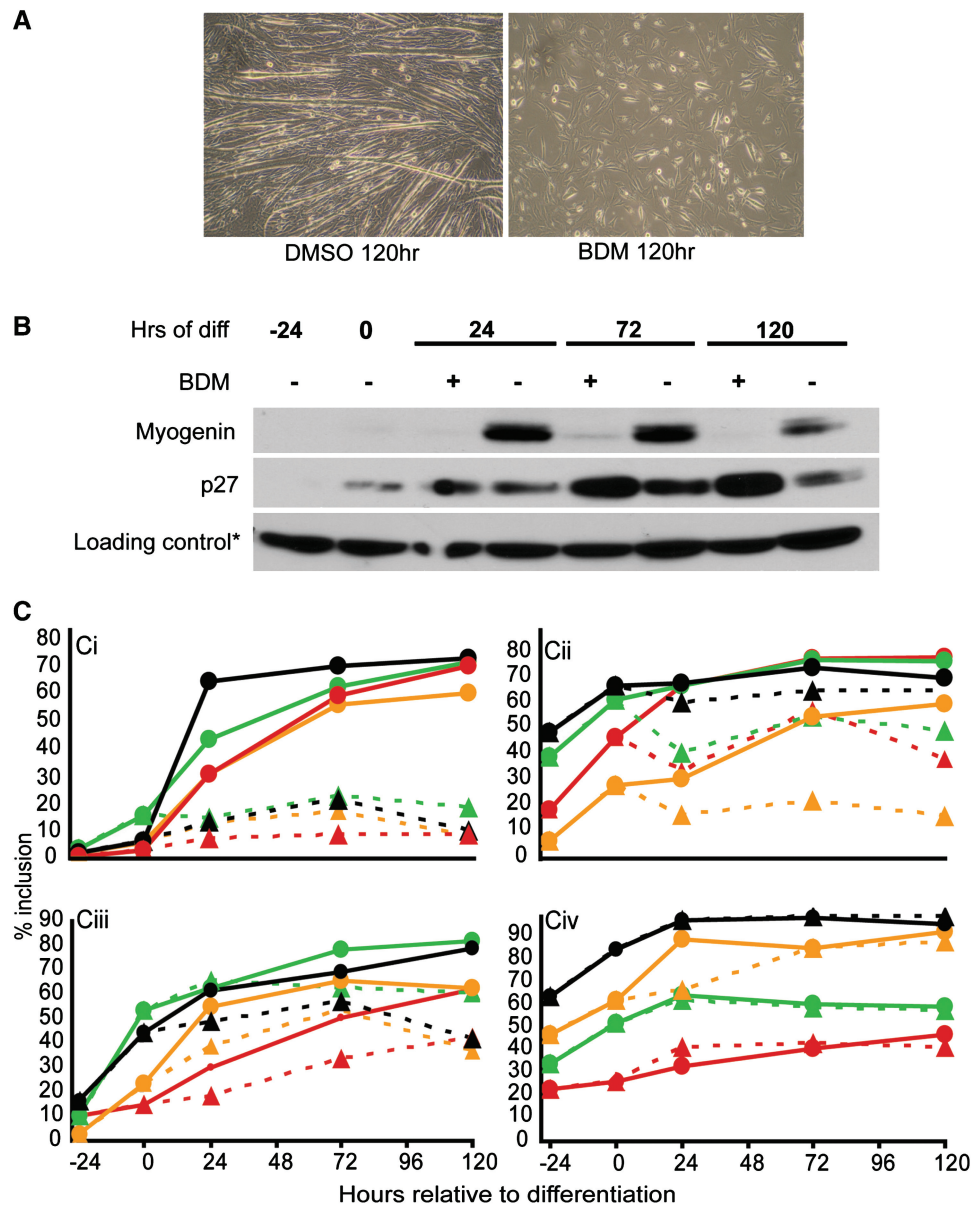**Figure 3.** Most splicing transitions are dependent upon myogenic differentiation. (**A**) Phase-contrast micrograph of vehicle (DMSO) and BDM (15 µM) treated C2C12 cells 120 h following induction of differentiation. (**B**) Western blots of whole cell lysate from DMSO or BDM treated C2C12 cells. Myogenin staining was used to measure myogenic differentiation progression and p27 staining was used to assess cell-cycle exit. 'Asterisk' indicates ∼60 kDa non-specific band detected by p27 antibody. (**C**) Four groups each depicting four representative examples of splicing transitions distinguished based on myogenic differentiation dependence. BDM (dashed lines) or DMSO vehicle (solid lines) was added to media at Hour 0. Group i exhibited a complete block of splicing transitions following addition of BDM. Group ii exhibited partial reversion of splicing transitions to undifferentiated (−24 h) state upon treatment with BDM. Group iii continued to undergo splicing transitions after BDM treatment, but at a reduced level compared to DMSO. Group iv events were not affected by addition of BDM. Splicing events depicted (in the order: black, green, red, orange) are as follows: Group i: Dtna_78,93; Atp2b1_87; Lrrfip2_93,102; Anxa7_66. Group ii: Dguok_100; 4632411B12R Rik_78; Mpp6_42; Azi2_47. Group iii: Capzb_113; Art3_30; Akap13_54; Bin1_45. Group iv: R3hdm_42; 5830434P21Rik_82; 5230400G24Rik_60; Pkp4_129.

<75% of that seen in control) (Figure 3Ciii) and four events showed comparable levels of regulation in both BDM treated and control conditions (Figure 3Civ). Similar results were obtained using C2C12 suspension cultures which, similar to BDM treatment, uncouples cell-cycle exit from differentiation (47) (data not shown).

We conclude that the majority of AS transitions identified are dependent upon and integrated into the myogenic differentiation program. Interestingly, this analysis also identified splicing transitions that are induced by cell-cycle withdrawal independent of the myogenic program, suggesting separate networks of regulated AS transitions controlled by distinct signaling mechanisms during myogenic differentiation.

Cis-regulatory elements associated with myogenic splicing transitions

To identify *cis*-acting regulatory elements associated with the validated splicing transitions in our data set, we first

examined the introns flanking all alternatively spliced regions that undergo transitions of ≥ 10 percentage points for enrichment of pentanucleotide motifs relative to a first-order Markov background model (13). We analyzed the first and last 250 nt of the upstream and downstream introns relative to the regulated alternative region. To avoid the inclusion of sequence motifs associated with constitutive splicing, the first 9 and last 30 nt of the intron which contain the conserved 5' and 3' splice sites, respectively were excluded from the analysis. We identified several significantly enriched motifs ($P < 0.05$ after Bonferroni correction, *italics* indicates corrected $P < 0.0001$ in Figure 4), including recognition sequences for FOX, CELF, PTB and hnRNP L splicing factors of which CELF and hnRNP L sites were most enriched ($P < 1.13 \times 10^{-13}$ after Bonferroni correction) (Figure 4 upper panel and Supplementary Table S3). Additionally, several motifs not associated with known RNA-binding proteins were identified.

A second analysis was performed to identify motifs within the flanking introns of regulated alternative regions that were conserved between mouse and seven other mammalian species. This analysis identified several of the same motifs including FOX, PTB and the NA(A>C)TAAY (STAR) motif (Figure 4 lower panel and Supplementary Table S4) previously found to be associated with skeletal muscle specific AS in adult tissue (23,42). FOX motifs within the first 250-nt downstream of the alternative region were both significantly enriched ($P = 1.25 \times 10^{-6}$) and conserved ($P \leq 5.25 \times 10^{-9}$). Based on the significance values, frequency of binding sites, and combined enrichment and conservation, we conclude that FOX, CELF, PTB, STAR and hnRNP L splicing regulators are particularly strong candidates for proteins that regulate different or overlapping subsets of splicing transitions during myogenic differentiation.

To identify enriched motifs that correlate with specific temporal splicing patterns, we performed regression analysis using the time course data regressed against enriched motifs within different intronic regions. Different enriched motifs were identified within specific intronic regions that strongly correlated with temporal splicing patterns (Figure 5 and Supplementary Figure S1). FOX motifs downstream from regulated alternative regions strongly correlated with continuous increased inclusion of variably spliced regions during differentiation, while upstream FOX motifs were correlated with an early decrease. For PTB, motifs located upstream of the variably spliced region correlated with continuous inclusion and downstream motifs correlated with late exclusion. Upstream hnRNP-L motifs correlated with an early increase in variably spliced region inclusion, while upstream STAR motifs correlated with an early decrease in inclusion. Overall, these data suggest different position- and time-dependent effects for several known splicing regulators and identified novel motifs associated with specific location and time-dependent effects.

Coordinated waves of splicing transitions occur throughout myogenic differentiation

To determine whether splicing transitions show different coordinated temporal patterns during myogenic differentiation, we examined the inclusion levels of 117 validated events at seven to nine time points from 24 h before, until 120 h after addition of differentiation media (Supplementary Table S1). Clusters of splicing events that share similar temporal patterns of regulation were identified by subjecting the time course data set to affinity propagation analysis, in which the behaviors of splicing events are compared to identify distinct clusters that share a common pattern of regulation with a set of 'exemplar' events (38). Four major clusters of ≥ 10 events
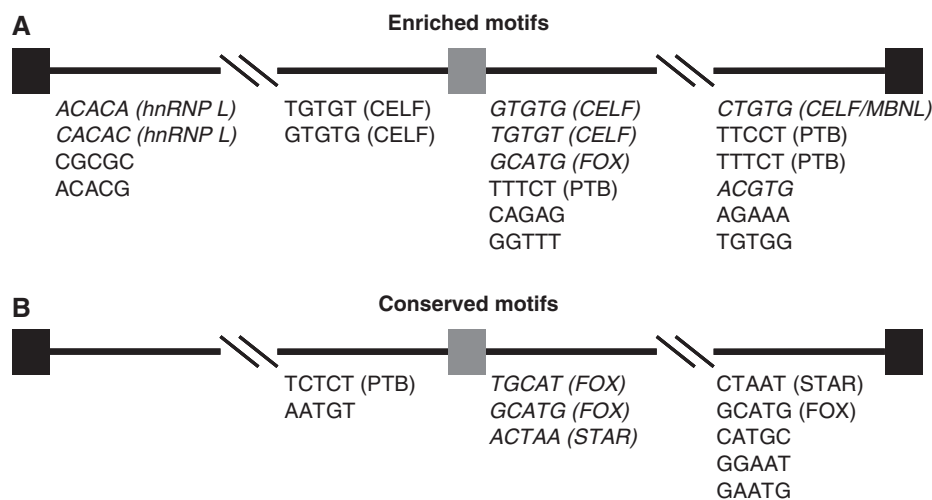

Figure 4. Motif enrichment and conservation in intronic regions flanking regulated variable regions. Pentameric motifs that are significantly enriched (upper panel) within the flanking intronic sequences of exons exhibiting AS changes during C2C12 differentiation or conserved (lower panel) among seven other mammalian species. Motifs were identified using a first-order Markov background model (corrected $P < 0.05$, motifs in italics indicate corrected $P < 0.0001$). Intronic regions analyzed included the first and last 250 nt of the upstream and downstream introns, excluding the first 9 nt and last 30 nt of the introns. Motifs matching the recognition sequences of known splicing regulators are indicted. Only motifs with significance $P < 0.05$ (standard text) or $P < 0.0001$ (italic text) are shown in the figure. For a complete list of all enriched and conserved motifs, see Supplementary Tables S3 and S4, respectively.
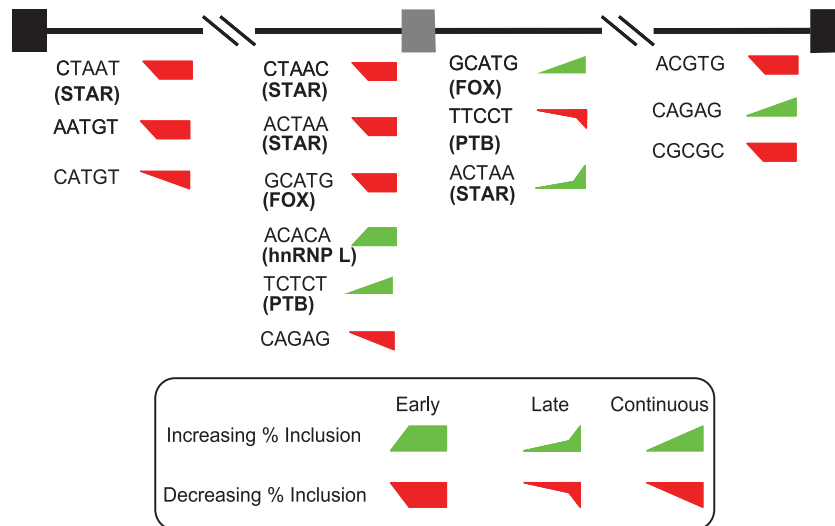
Figure 5. Regression analysis identified enriched motifs associated with temporal splicing transitions. Regression analysis was used to identify correlations between enriched motifs and specific temporal patterns of splicing transitions. The enrichment of each motif within the various intronic regions was regressed against the magnitude of each splicing transition (% of total change) at −24, 0, 12, 24, 48, 72 and 120 h relative to differentiation induction. Green denotes positive correlations (motif associated with increased inclusion of alternative region during differentiation); red denotes inverse correlations (decreased inclusion of alternative region during differentiation). Early (before 24 h), late (after 24 h) and continuous refers to period in the differentiation time course to which a given motif is most strongly correlated. For a full display of all regression analysis heatmap data see Supplementary Figure S1.

(comprising 102 out of the 111 with ≥10 point change events that were analyzed by affinity propagation) were defined based on the timing of maximal transition and are named based on the exemplar event that best represents the behavior of the cluster (Figure 6). Among the four major clusters, there were 82 events that showed increasing alternative region inclusion and 20 events that showed increased skipping (Figure 6A). The ALS2CR2e8 cluster (10 events) exhibits an extremely early increase in alternative region inclusion between −24 and 12 h relative to induction of differentiation, the DTNA cluster (30 events) exhibits an early (12–24 h) increase in inclusion, while the KIF1B cluster (42 events) underwent a late increase in alternative region inclusion between 24 and 72 h. The PHKA1 cluster (20 events) exhibited early/continuous increased skipping. For a full list of events and their assigned cluster see Supplementary Table S5. In addition, two minor clusters consisting of ≤8 events each (9 total) are shown in Supplementary Figure S2.

To identify *cis*-regulatory elements associated with specific splicing clusters, the four largest clusters were analyzed for motif enrichment and conservation (Figure 6B and Supplementary Tables S6 and S7). This analysis provides an independent assessment of motifs associated with temporal patterns that can be compared to the regression analysis presented in Figure 5. TIA and hnRNP-L motifs were significantly enriched within the upstream intronic regions of events in the KIF1B cluster, while CELF, FOX and hnRNP-L motifs were significantly enriched in the DTNA (early increase) cluster. Interestingly, hnRNP-L motifs were identified in the late as well as the early cluster suggesting the potential involvement in different temporal events as components of

combinatorial regulation (see Discussion). Conservation analysis of different clusters identified significantly conserved FOX motifs within the proximal downstream intronic regions of both the KIF1B (late increase) and DTNA (early increase) clusters. These data, in addition to the regression analysis described earlier, strongly suggest a role for FOX, CELF and hnRNP L proteins in the regulation of AS transitions during myogenic differentiation that is specific with regard to binding site location and timing of activity.

Several splicing factors undergo different temporal transitions during myogenic differentiation

We performed western blot analysis on nuclear and cytoplasmic fractions during a time course of C2C12 differentiation to assess the steady-state nuclear abundance of splicing regulators during muscle differentiation (Figure 7). Of particular interest were those regulators for which binding sites were enriched near the variably spliced region. The protein expression levels of CELF, FOX, hnRNP L and MBNL family members were examined during C2C12 differentiation. All western blots shown in Figure 7 contain 30 μg of nuclear and cytoplasmic extracts prepared from the same differentiation time course. Except CUGBP1 which shows small differences, the results are representative of what was observed in at least two independent cultures (data not shown). TATA-Binding Protein (TBP) and GAPDH were used to determine the purity of nuclear and cytoplasmic fractions, respectively. Myogenin was used as an early marker of differentiation. It is difficult to find proteins that show constant steady-state levels during C2C12 differentiation, therefore, Ponceau S red staining was used to confirm even loading.
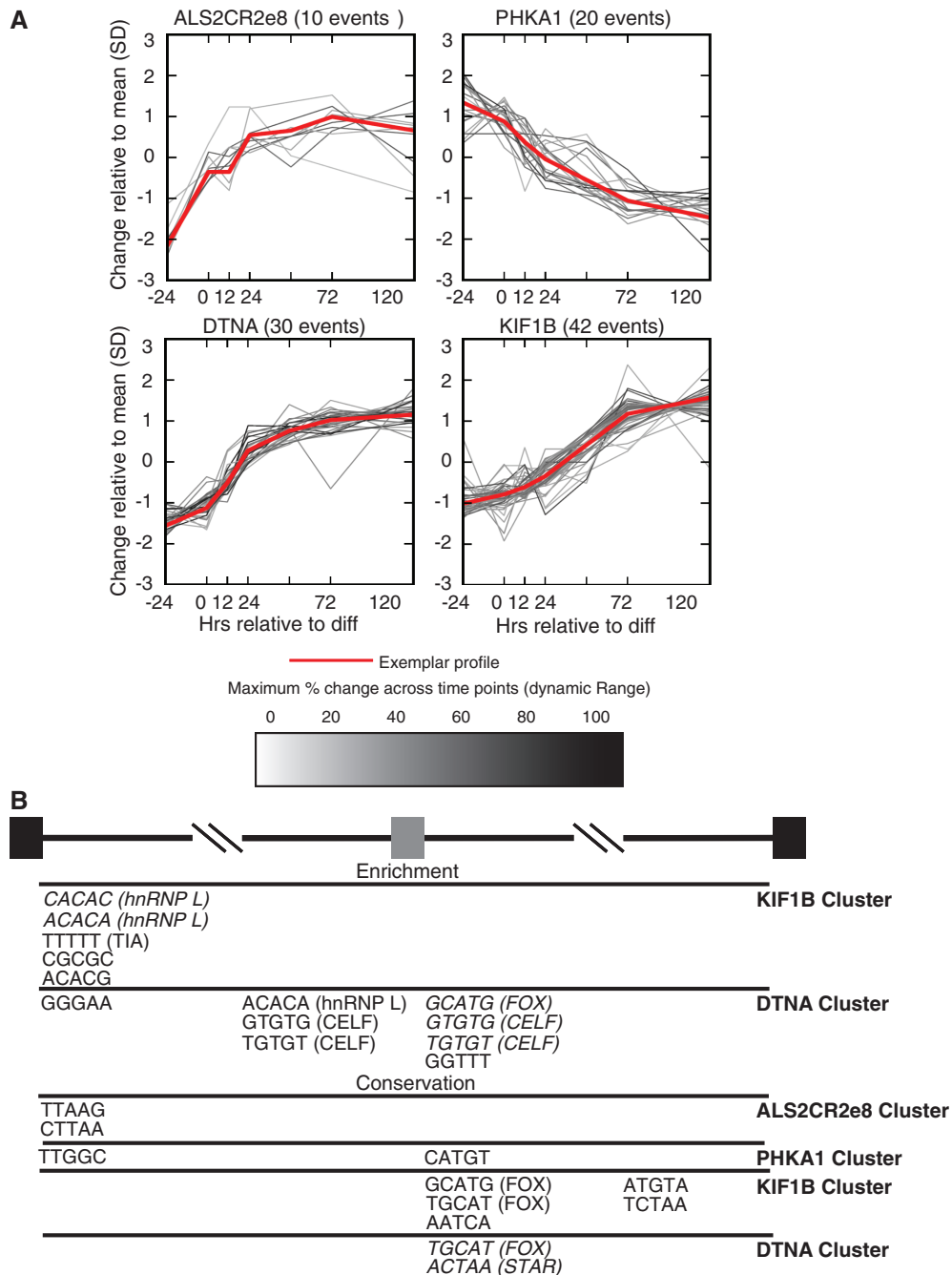
Figure 6. Clusters of splicing transitions with similar temporal patterns exhibit enrichment of specific motifs. Splicing transitions that share similar temporal patterns were grouped together in clusters using affinity propagation analysis. Each cluster contains a single 'exemplar' event (red), which typifies the temporal behavior of that cluster. The remaining events in each cluster are plotted in various shades of gray, where darker gray represents a larger maximum difference in inclusion level across all time points. (**A**) Four major clusters ($n \geq 10$) representing 102 out of 111 events analyzed by affinity propagation analysis. A total of six clusters were identified in all; the two remaining contain ≤ 8 events and are shown in Supplementary Figure S2. (**B**) Motif enrichment and conservation analysis identical to that described in Figure 4A was performed on the four clusters shown in Figure 5A. Significantly enriched (corrected $P < 0.05$, italics indicate corrected $P < 0.0001$) motifs are shown. Motifs matching the recognition sequences of known splicing regulators are indicted. For a list of all cluster-specific enriched motifs, see Supplementary Table S5. For a full list of all cluster specific enrichment and conservation date see Supplementary Tables S6 and S7, respectively.

We found that several splicing regulators undergo dramatic changes in nuclear steady-state levels that correlate with different temporal clusters of splicing transitions. Factors showing early transitions (within the first 24 h) include all three MBNL family members and hnRNP C. MBNL1 shows a particularly dramatic increase of a

doublet at ~42 kDa between 0 and 24 h as well as a weaker ~32 kDa band that decreases during differentiation. Both MBNL2 and MBNL3 undergo changes in nuclear abundance as well as nuclear and cytoplasmic distribution early in differentiation, with nuclear protein first detectable at 12 h. Five proteins (hnRNP L, CUGBP1,
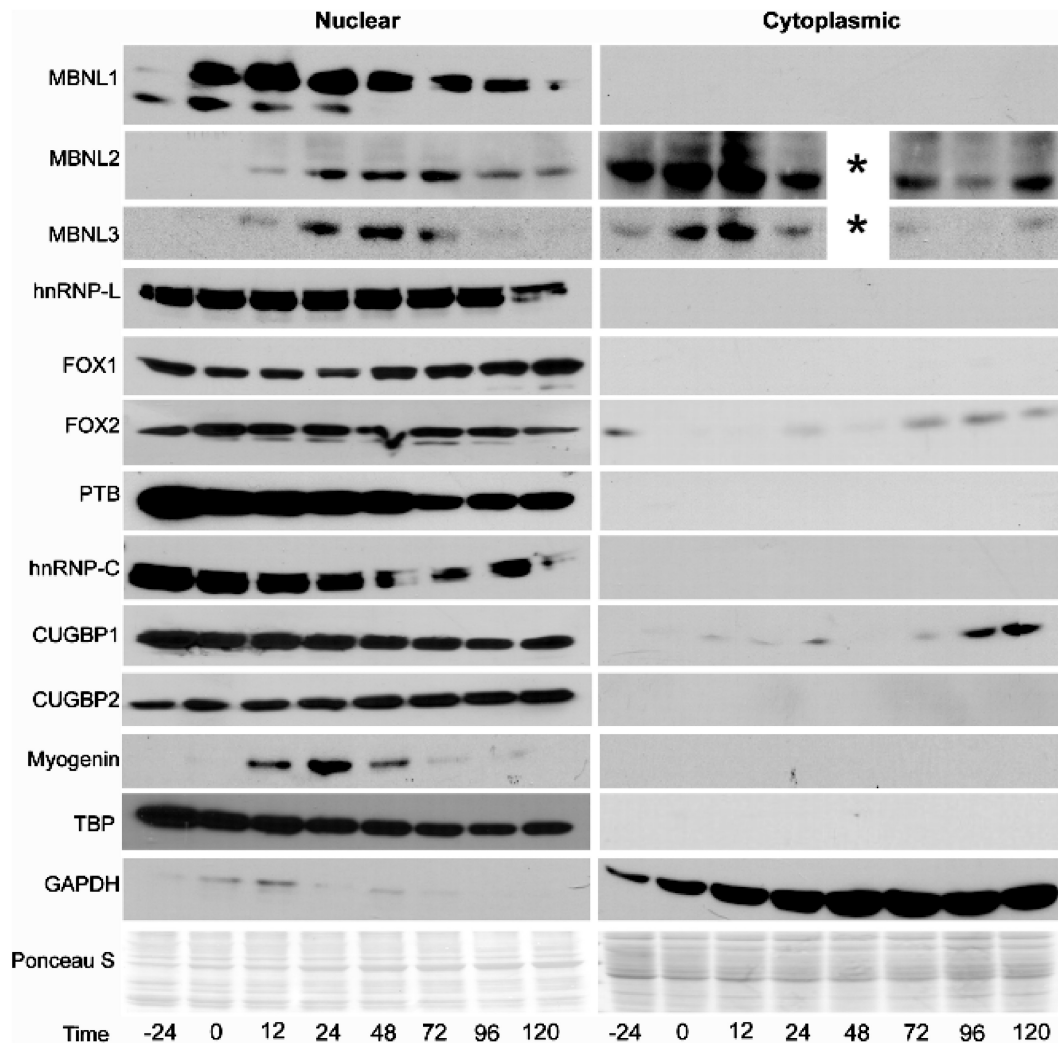
Figure 7. Western blot analysis shows dynamic regulation of candidate splicing factors during C2C12 differentiation. Steady-state nuclear and cytoplasmic protein levels of MBNL, CELF, FOX, PTB, hnRNP C and hnRNP L splicing regulators during C2C12 differentiation. TBP and GAPDH serve as nuclear and cytoplasmic markers, respectively. Ponceau S staining serves as a loading control.

CUGBP2, FOX1 and PTB) showed modulation of nuclear abundance >24 h following induction of differentiation. Nuclear levels of CUGBP1 and CUGBP2 proteins shown a respective slight decrease and increase during C2C12 differentiation. FOX1 levels increased modestly, while FOX2 protein levels remained constant throughout differentiation and PTB decreases during differentiation.

Overall, changes in nuclear levels of some proteins correlated well with the preceding analyses associating known binding motifs with specific temporal patterns. Motifs associated with hnRNP L and PTB correlated predominantly with late changes, which is consistent with the late changes in protein expression. It is important to note that regulated splicing is likely to involve combinatorial control such that expression levels of individual proteins might be involved in but not solely determinative of splicing patterns (see Discussion). We conclude that several of the splicing regulators identified through enriched/conserved binding motifs and that undergo

changes in nuclear abundance are likely to play important roles in regulating myogenic splicing transitions.

DISCUSSION

Although there is ample evidence for extensive regulation of AS in metazoans, there are conflicting views as to the extent of its biological significance (18–21). By amassing a large validated data set of AS transitions that occur in response to myogenic differentiation, we characterized global splicing regulation directly and in a biologically relevant context. We found nearly 100 splicing transitions that were robustly regulated during differentiation, most AS events detected in both mouse and quail myoblasts underwent conserved transitions during differentiation, were dependent on the myogenic program, and can be grouped into four temporal patterns, demonstrating coordinated regulation. In addition, AS transitions initiate prior to the appearance of the earliest transcriptional markers of differentiation, indicative of links to

myogenic differentiation program within proliferating myoblasts. These findings strongly support a biologically significant role for AS as an integral component of myogenic differentiation. This conclusion is further supported by results from mRNA-Seq analysis published while this manuscript was in revision, demonstrating extensive splicing changes during C2C12 differentiation (48).

We found that more than half of the splicing transitions detectable in the QM7 quail cell line were conserved with mouse. This high level of conservation is particularly striking given the 300 million years of evolution that separates mammals and birds and the fact that not only is AS conserved, but also a regulated transition of splicing patterns during a shared developmental process. Genome-wide analysis using EST and full-length cDNA databases have detected low levels of AS conservation across species (21,49–51). However, here we find that a high proportion of splicing transitions in myogenic differentiation are conserved, consistent with previous results (13). Previously, we identified a conserved network of splicing transitions that occur during postnatal murine heart development (13). It is interesting to note that 49 out of 78 (62.8%) of the splicing transitions that occur during postnatal heart development are regulated during C2C12 differentiation. Additionally, in both systems splicing transitions occurred in tightly grouped temporal waves, suggesting that the splicing regulatory programs in myoblast differentiation and heart development at least partially overlap. GO analysis of the regulated transcripts in C2C12 showed a strong enrichment in molecular function terms relating to muscle development and function, even when compared against a muscle specific background (Figure 1D), consistent with other muscle splicing data sets (23,42). Additionally, several of the regulated AS transitions we observed during C2C12 differentiation have already been experimentally shown to be important for normal muscle function (52–56), further supporting the view that AS regulation is affecting processes that are biologically significant to muscle. A substantial subset of alternative exons that have been previously shown to be enriched in human (42) and mouse (23) muscle were found to undergo transitions in exon inclusion during C2C12 differentiation. Specifically, of the 56 muscle-enriched human exons described by Das *et al.*, 21 (37.5%) were orthologous to exons we observed to undergo splicing transitions during C2C12 differentiation. Consistent with our results, this study identified enrichment of Fox and CELF-binding sites within 200-nt downstream of muscle enriched exons. Sugnet *et al.* reported 28 exons enriched in adult mouse skeletal muscle. Of these 28 exons, 12 (42.9%) were found to undergo regulated splicing transitions during C2C12 differentiation. However, it should be noted that both of these studies examined adult tissues, thus it is likely that many of the events they identified were excluded from our study because they do not undergo regulated changes in the level of alternative exon inclusion during C2C12 differentiation (despite being enriched in adult muscle). It is also important to note that we did not attempt to exhaustively identify all exons that undergo changes in inclusion

levels during C2C12 differentiation, but instead our goal was to identify a large set of robust splicing transitions which we could then use for computational and molecular analysis, and it is likely that many splicing transitions that occur during C2C12 differentiation were not detected in our study.

To determine whether the splicing transitions are specific to the myogenic differentiation pathway or are a general response to cell-cycle withdrawal, we included BDM in the differentiation medium which allows induction of cell-cycle arrest but prevents progression of the differentiation program (46). The results indicated that most of the splicing transitions tested are dependent on induction of the myogenic program and cannot be induced through serum deprivation, cellular confluence and/or cell-cycle exit alone. Time course experiments showed that ≥10 splicing transitions occur largely before the addition of differentiation media (Hour 0), suggesting that these events are regulated by signals associated with increasing cell density or cell-cycle exit. It is particularly interesting that for the majority of the early events tested, the transition initiated in proliferative cells was reversed when myogenic differentiation was blocked by using BDM treatment (Figure 3Cii). These results indicate the presence of a signaling mechanism that links early differentiation-dependent splicing transitions with components of the myogenic program that initiate as proliferating myoblasts become confluent.

Two independent computational analyses to identify motifs associated with specific temporal patterns showed that specific sequence motifs are associated with distinct temporal splicing transition patterns. Both analyses identified hnRNP-L motifs (ACACA) within the proximal upstream intronic region of alternative regions as strongly associated with an early increase in alternative region inclusion (Figures 5 and 6B). FOX-binding motifs were also identified in both analyses as strongly associated with increased inclusion when located within the proximal downstream intronic region of alternative regions. Such concordance of two independent computational analyses is strongly supports a position-dependent role for these two proteins in AS regulation.

Computational analysis demonstrated that motifs associated with known regulators of AS, including the FOX, CELF, PTB, MBNL and hnRNP-L families of splicing regulators, are enriched within the intronic regions surrounding the variably spliced regions that undergo regulation during differentiation. Many of these enriched binding motifs are also conserved with regard to sequence and relative location among eight mammalian species. Regression analysis predicted that FOX-binding upstream from alternative regions promotes skipping, and downstream promotes inclusion, which is in agreement with experimental evidence from other groups (57–59). While FOX-binding sites are highly enriched near differentiation-induced splicing events, steady-state nuclear FOX2 levels remain relatively constant, while FOX1 levels modestly increase very early in C2C12 differentiation but remain constant thereafter (Figure 5A). Regression analysis showed that FOX sites are most enriched in early AS changes, suggesting that the early

increase in FOX1 steady-state levels could account for early AS transitions observed. However, the FOX2 transcript contains two exons which undergo regulated splicing changes of >20 percentage points during C2C12 differentiation, which could alter its activity independent of its steady-state levels, leading to another mechanism for the regulation of FOX activity. This possibility is supported by reports that show AS in the FOX1 and FOX2 transcripts result in alterations in splicing regulatory activity (11,60).

MBNL family members show dynamic regulation in steady-state nuclear protein levels early in myogenic differentiation. Considering the large proportion of validated splicing transitions that show dramatic changes within the first 24 h of differentiation, these data indicate that MBNL family members are likely contributors to developmentally regulated myogenic AS. While MBNL-binding motifs were detected in our computational analysis, the high stringency of the analysis combined with the relatively high variability of MBNL-binding motifs likely led to their under-representation. Nevertheless, MBNL-binding motifs were still found to be significantly enriched (corrected $P = 9.55 \times 10^{-6}$) in the downstream distal intronic regions of regulated alternative region. Furthermore, the well characterized role that MBNL family members play in the pathogenesis of myotonic dystrophy suggest they are important for normal muscle development (29,61). Overall, these data demonstrate that AS during myogenic differentiation is highly conserved, extensively regulated and suggests that splicing regulation is influenced by multiple regulatory factors associated with distinct temporal clusters of splicing transitions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Pan,Q., Shai,O., Lee,L.J., Frey,B.J. and Blencowe,B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**, 1413–1415.
2. Goncalves,V., Matos,P. and Jordan,P. (2008) The beta-catenin/TCF4 pathway modifies alternative splicing through modulation of SRp20 expression. *RNA*, **14**, 2538–2549.
3. Cheng,C., Yaffe,M.B. and Sharp,P.A. (2006) A positive feedback loop couples Ras activation and CD44 alternative splicing. *Genes Dev.*, **20**, 1715–1720.
4. Shin,C. and Manley,J.L. (2004) Cell signalling and the control of pre-mRNA splicing. *Nat. Rev. Mol. Cell Biol.*, **5**, 727–738.
5. Ip,J.Y., Tong,A., Pan,Q., Topp,J.D., Blencowe,B.J. and Lynch,K.W. (2007) Global analysis of alternative splicing during T-cell activation. *RNA*, **13**, 563–572.
6. Cooper,T.A., Wan,L. and Dreyfuss,G. (2009) RNA and disease. *Cell*, **136**, 777–793.
7. Grosso,A.R., Martins,S. and Carmo-Fonseca,M. (2008) The emerging role of splicing factors in cancer. *EMBO Rep.*, **9**, 1087–1093.
8. Kuyumcu-Martinez,N.M. and Cooper,T.A. (2006) Misregulation of alternative splicing causes pathogenesis in myotonic dystrophy. *Prog. Mol. Subcell. Biol.*, **44**, 133–159.
9. Wolfe,M.S. (2009) Tau mutations in neurodegenerative diseases. *J. Biol. Chem.*, **6**, 6021–6025.
10. Boutz,P.L., Stoilov,P., Li,Q., Lin,C.H., Chawla,G., Ostrow,K., Shiue,L., Ares,M. Jr and Black,D.L. (2007) A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.*, **21**, 1636–1652.
11. Lee,J.A., Tang,Z.Z. and Black,D.L. (2009) An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev.*, **23**, 2284–2293.
12. Makeyev,E.V., Zhang,J., Carrasco,M.A. and Maniatis,T. (2007) The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol. Cell*, **27**, 435–448.
13. Kalsotra,A., Xiao,X., Ward,A.J., Castle,J.C., Johnson,J.M., Burge,C.B. and Cooper,T.A. (2008) A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc. Natl Acad. Sci. USA*, **105**, 20333–20338.
14. Chawla,G., Lin,C.H., Han,A., Shiue,L., Ares,M. Jr and Black,D.L. (2009) Sam68 regulates a set of alternatively spliced exons during neurogenesis. *Mol. Cell Biol.*, **29**, 201–213.
15. Li,Q., Lee,J.A. and Black,D.L. (2007) Neuronal regulation of alternative pre-mRNA splicing. *Nat. Rev. Neurosci.*, **8**, 819–831.
16. Licatalosi,D.D., Mele,A., Fak,J.J., Ule,J., Kayikci,M., Chi,S.W., Clark,T.A., Schweitzer,A.C., Blume,J.E., Wang,X. *et al.* (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, **456**, 464–469.
17. Lynch,K.W. (2004) Consequences of regulated pre-mRNA splicing in the immune system. *Nat. Rev. Immunol.*, **4**, 931–940.
18. Grabowski,P.J. and Black,D.L. (2001) Alternative RNA splicing in the nervous system. *Prog. Neurobiol.*, **65**, 289–308.
19. Kriventseva,E.V., Koch,I., Apweiler,R., Vingron,M., Bork,P., Gelfand,M.S. and Sunyaev,S. (2003) Increase of functional diversity by alternative splicing. *Trends Genet.*, **19**, 124–128.
20. Melamud,E. and Moult,J. (2009) Stochastic noise in splicing machinery. *Nucleic Acids Res.*, **37**, 4873–4886.
21. Sorek,R., Shamir,R. and Ast,G. (2004) How prevalent is functional alternative splicing in the human genome? *Trends Genet.*, **20**, 68–71.
22. Castle,J.C., Zhang,C., Shah,J.K., Kulkarni,A.V., Kalsotra,A., Cooper,T.A. and Johnson,J.M. (2008) Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. *Nat. Genet.*, **40**, 1416–1425.
23. Sugnet,C.W., Srinivasan,K., Clark,T.A., O'Brien,G., Cline,M.S., Wang,H., Williams,A., Kulp,D., Blume,J.E., Haussler,D. *et al.* (2006) Unusual intron conservation near tissue-regulated exons found by splicing microarrays. *PLoS Comput. Biol.*, **2**, e4.
24. Sultan,M., Schulz,M.H., Richard,H., Magen,A., Klingenhoff,A., Scherf,M., Seifert,M., Borodina,T., Soldatov,A., Parkhomchuk,D. *et al.* (2008) A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*, **321**, 956–960.

25. Wang,E.T., Sandberg,R., Luo,S., Khrebtukova,I., Zhang,L., Mayr,C., Kingsmore,S.F., Schroth,G.P. and Burge,C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470–476.

26. Bonizzoni,P., Mauri,G., Pesole,G., Picardi,E., Pirola,Y. and Rizzi,R. (2009) Detecting alternative gene structures from spliced ESTs: a computational approach. *J. Comput. Biol.*, **16**, 43–66.

27. Gupta,S., Zink,D., Korn,B., Vingron,M. and Haas,S.A. (2004) Strengths and weaknesses of EST-based prediction of tissue-specific alternative splicing. *BMC Genomics*, **5**, 72.

28. Kan,Z., Rouchka,E.C., Gish,W.R. and States,D.J. (2001) Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res.*, **11**, 889–900.

29. Pascual,M., Vicente,M., Monferrer,L. and Artero,R. (2006) The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation*, **74**, 65–80.

30. Kuroyanagi,H., Ohno,G., Mitani,S. and Hagiwara,M. (2007) The Fox-1 family and SUP-12 coordinately regulate tissue-specific alternative splicing in vivo. *Mol. Cell Biol.*, **27**, 8612–8621.

31. Yaffe,D. and Saxel,O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature*, **270**, 725–727.

32. Blau,H.M., Pavlath,G.K., Hardeman,E.C., Chiu,C.P., Silberstein,L., Webster,S.G., Miller,S.C. and Webster,C. (1985) Plasticity of the differentiated state. *Science*, **230**, 758–766.

33. Silberstein,L., Webster,S.G., Travis,M. and Blau,H.M. (1986) Developmental progression of myosin gene expression in cultured muscle cells. *Cell*, **46**, 1075–1081.

34. Andres,V. and 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.

35. Antin,P.B. and Ordahl,C.P. (1991) Isolation and characterization of an avian myogenic cell line. *Dev. Biol.*, **143**, 111–121.

36. Castle,J., Garrett-Engele,P., Armour,C.D., Duenwald,S.J., Loerch,P.M., Meyer,M.R., Schadt,E.E., Stoughton,R., Parrish,M.L., Shoemaker,D.D. *et al.* (2003) Optimization of oligonucleotide arrays and RNA amplification protocols for analysis of transcript structure and alternative splicing. *Genome Biol.*, **4**, R66.

37. Thomas,P.D., Campbell,M.J., Kejariwal,A., Mi,H., Karlak,B., Daverman,R., Diemer,K., Muruganujan,A. and Narechania,A. (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.*, **13**, 2129–2141.

38. Frey,B.J. and Dueck,D. (2007) Clustering by passing messages between data points. *Science*, **315**, 972–976.

39. Ladd,A.N., Stenberg,M.G., Swanson,M.S. and Cooper,T.A. (2005) Dynamic balance between activation and repression regulates pre-mRNA alternative splicing during heart development. *Dev. Dyn.*, **233**, 783–793.

40. Wang,G.S., Kearney,D.L., De Biasi,M., Taffet,G. and Cooper,T.A. (2007) Elevation of RNA-binding protein CUGBP1 is an early event in an inducible heart-specific mouse model of myotonic dystrophy. *J Clin. Invest.*, **117**, 2802–2811.

41. Holt,I., Jacquemin,V., Fardaei,M., Sewry,C.A., Butler-Browne,G.S., Furling,D., Brook,J.D. and Morris,G.E. (2009) Muscleblind-like proteins: similarities and differences in normal and myotonic dystrophy muscle. *Am. J. Pathol.*, **174**, 216–227.

42. Das,D., Clark,T.A., Schweitzer,A., Yamamoto,M., Marr,H., Arribere,J., Minovitsky,S., Poliakov,A., Dubchak,I., Blume,J.E. *et al.* (2007) A correlation with exon expression approach to identify cis-regulatory elements for tissue-specific alternative splicing. *Nucleic Acids Res.*, **35**, 4845–4857.

43. Thomas,P.D., Kejariwal,A., Guo,N., Mi,H., Campbell,M.J., Muruganujan,A. and Lazareva-Ulitsky,B. (2006) Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.*, **34**, W645–W650.

44. Guo,K., Wang,J., Andres,V., Smith,R.C. and Walsh,K. (1995) MyoD-induced expression of p21 inhibits cyclin-dependent kinase activity upon myocyte terminal differentiation. *Mol. Cell Biol.*, **15**, 3823–3829.

45. Blau,H.M. and Webster,C. (1981) Isolation and characterization of human muscle cells. *Proc. Natl Acad. Sci. USA*, **78**, 5623–5627.

46. Dhawan,J. and Helfman,D.M. (2004) Modulation of acto-myosin contractility in skeletal muscle myoblasts uncouples growth arrest from differentiation. *J. Cell. Sci.*, **117**, 3735–3748.

47. Milasincic,D.J., Dhawan,J. and Farmer,S.R. (1996) Anchorage-dependent control of muscle-specific gene expression in C2C12 mouse myoblasts. *In Vitro Cell Dev. Biol. Anim.*, **32**, 90–99.

48. Trapnell,C., Williams,B.A., Pertea,G., Mortazavi,A., Kwan,G., van Baren,M.J., Salzberg,S.L., Wold,B.J. and Pachter,L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, **5**, 511–515.

49. Modrek,B. and Lee,C.J. (2003) Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat. Genet.*, **34**, 177–180.

50. Takeda,J., Suzuki,Y., Sakate,R., Sato,Y., Seki,M., Irie,T., Takeuchi,N., Ueda,T., Nakao,M., Sugano,S. *et al.* (2008) Low conservation and species-specific evolution of alternative splicing in humans and mice: comparative genomics analysis using well-annotated full-length cDNAs. *Nucleic Acids Res.*, **36**, 6386–6395.

51. Yeo,G.W., Van Nostrand,E., Holste,D., Poggio,T. and Burge,C.B. (2005) Identification and analysis of alternative splicing events conserved in human and mouse. *Proc. Natl Acad. Sci. USA*, **102**, 2850–2855.

52. Wechsler-Reya,R.J., Elliott,K.J. and Prendergast,G.C. (1998) A role for the putative tumor suppressor Bin1 in muscle cell differentiation. *Mol. Cell Biol.*, **18**, 566–575.

53. van der Flier,A., Kuikman,I., Kramer,D., Geerts,D., Kreft,M., Takafuta,T., Shapiro,S.S. and Sonnenberg,A. (2002) Different splice variants of filamin-B affect myogenesis, subcellular distribution, and determine binding to integrin [beta] subunits. *J. Cell. Biol.*, **156**, 361–376.

54. Burkin,D.J., Gu,M., Hodges,B.L., Campanelli,J.T. and Kaufman,S.J. (1998) A functional role for specific spliced variants of the alpha7beta1 integrin in acetylcholine receptor clustering. *J Cell Biol*, **143**, 1067–1075.

55. Belkin,A.M. and Retta,S.F. (1998) beta1D integrin inhibits cell cycle progression in normal myoblasts and fibroblasts. *J. Biol. Chem.*, **273**, 15234–15240.

56. Collo,G., Starr,L. and Quaranta,V. (1993) A new isoform of the laminin receptor integrin alpha 7 beta 1 is developmentally regulated in skeletal muscle. *J. Biol. Chem.*, **268**, 19019–19024.

57. Ponthier,J.L., Schluepen,C., Chen,W., Lersch,R.A., Gee,S.L., Hou,V.C., Lo,A.J., Short,S.A., Chasis,J.A., Winkelmann,J.C. *et al.* (2006) Fox-2 splicing factor binds to a conserved intron motif to promote inclusion of protein 4.1R alternative exon 16. *J. Biol. Chem.*, **281**, 12468–12474.

58. Yeo,G.W., Coufal,N.G., Liang,T.Y., Peng,G.E., Fu,X.D. and Gage,F.H. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.*, **16**, 130–137.

59. Zhang,C., Zhang,Z., Castle,J., Sun,S., Johnson,J., Krainer,A.R. and Zhang,M.Q. (2008) Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev.*, **22**, 2550–2563.

60. Nakahata,S. and Kawamoto,S. (2005) Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. *Nucleic Acids Res.*, **33**, 2078–2089.

61. Lin,X., Miller,J.W., Mankodi,A., Kanadia,R.N., Yuan,Y., Moxley,R.T., Swanson,M.S. and Thornton,C.A. (2006) Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.*, **15**, 2087–2097.