Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2

Chaolin Zhang,1,2,5 Zuo Zhang,1,5 John Castle,3 Shuying Sun,1,4 Jason Johnson,3 Adrian R. Krainer,1,6 and Michael Q. Zhang1,7

1Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA; 2Department of Biomedical Engineering, The State University of New York at Stony Brook, Stony Brook, New York 11794, USA; 3Rosetta Inpharmatics LLC, a wholly owned subsidiary of Merck and Co., Seattle, Washington 98109, USA; 4Department of Molecular and Cellular Biology, The State University of New York at Stony Brook, Stony Brook, New York 11794, USA

The precise regulation of many alternative splicing (AS) events by specific splicing factors is essential to determine tissue types and developmental stages. However, the molecular basis of tissue-specific AS regulation and the properties of splicing regulatory networks (SRNs) are poorly understood. Here we comprehensively predict the targets of the brain- and muscle-specific splicing factor Fox-1 (A2BP1) and its paralog Fox-2 (RBM9) and systematically define the corresponding SRNs genome-wide. Fox-1/2 are conserved from worm to human, and specifically recognize the RNA element UGCAUG. We integrate Fox-1/2-binding specificity with phylogenetic conservation, splicing microarray data, and additional computational and experimental characterization. We predict thousands of Fox-1/2 targets with conserved binding sites, at a false discovery rate (FDR) of ∼24%, including many validated experimentally, suggesting a surprisingly extensive SRN. The preferred position of the binding sites differs according to AS pattern, and determines either activation or repression of exon recognition by Fox-1/2. Many predicted targets are important for neuromuscular functions, and have been implicated in several genetic diseases. We also identified instances of binding site creation or loss in different vertebrate lineages and human populations, which likely reflect fine-tuning of gene expression regulation during evolution.

Keywords: Tissue-specific alternative splicing, splicing regulatory network, Fox-1/A2BP1, Fox-2/RBM9, UGCAUG; comparative genomics

Supplemental material is available at http://www.genesdev.org.

Received June 6, 2007; revised version accepted July 28, 2008.

The sequencing of complete genomes revealed that complex metazoans, including mammals, have only slightly more genes than unicellular yeast [International Human Genome Sequencing Consortium 2001]. Organismal complexity must have resulted largely from mechanisms for diversifying the expression products, and the temporal and spatial patterns, from a limited set of genes. It is crucial to understand how gene expression is orchestrated to determine developmental stages, specify cell types, and respond to external stimuli [Maniatis and Reed 2002]. Alternative splicing (AS), the process for removing introns from pre-mRNA transcripts and joining exons in different combinations, is an essential step of post-transcriptional regulation [Cartegni et al. 2002; Black 2003]. In mammals, more than two-thirds of genes are alternatively spliced [Johnson et al. 2003]. The choice of exons and splice sites is largely determined by many RNA-binding proteins, or splicing factors, which interact with cis-regulatory elements to activate or repress particular splicing events.

Many splicing factors have restricted and dynamic expression patterns, and play important roles in tissue-specific or developmentally regulated splicing of particular transcripts. However, the mechanisms and impact of these AS events remain poorly understood. A well-studied example is Sxl, Tra, Tra-2, and several other splicing factors in Drosophila, which regulate a cascade of AS events to control sex determination [Lopez 1998]. In mammals, tissue-specific splicing factors include Nova-1/2, PTB/nPTB, Fox-1/2, Muscleblind-like (MBNL) and CELF family proteins, Hu proteins, TIA1/TIAR, and probably many more yet to be characterized [for reviews, see Li et al. 2007; David and Manley 2008]. The identification of the RNA targets of these factors is critical for...
understanding the splicing regulatory networks (SRNs), but remains challenging, in most cases, only a handful of targets have been determined experimentally. Recently, the development of high-throughput technologies that monitor mRNA isoform abundances and protein–RNA interactions, including splicing microarrays [Johnson et al. 2003; Ule et al. 2005b; Sugnet et al. 2006; Boutz et al. 2007b], RNP immunoprecipitation-microarray (RIP-Chip) [Keene et al. 2006], and ultraviolet cross-linking and immunoprecipitation (CLIP) assays [Ule et al. 2005a], provided new opportunities to identify in vivo RNA targets and characterize SRNs genome-wide [for reviews, see Blencowe 2006; Moore and Silver 2008]. This has been demonstrated already in studies of Nova-1/2 targets, which revealed important functions of coregulated targets in the neuronal synapse and in axon guidance, as well as mechanisms by which Nova-1/2 activate or repress exon inclusion, depending on the locations of their binding sites [Ule et al. 2006]. Other studies described subsets of exons with splicing patterns characteristic of particular tissues, or associated with responses to certain stimuli; these exons sometimes show enrichment of RNA sequence motifs and Gene Ontology (GO) terms, suggesting coordinated splicing regulation that could be important for cellular functions [Sugnet et al. 2006; Das et al. 2007; Fagnani et al. 2007; Ip et al. 2007; McKee et al. 2007]. However, more detailed studies are required to identify individual targets for particular splicing factors.

Computational target prediction for specific splicing factors is difficult, largely due to the small size and degeneracy of splicing-factor-binding motifs. An exception to this degeneracy is the hexanucleotide UGCAUG, which is an important intronic element for the splicing of several exons [Huh and Hynes 1994; Kawamoto 1996; Lim and Sharp 1998]. A computational study further suggested that this element is enriched in the introns downstream from a set of neuron-specific exons [Brudno et al. 2001]. Importantly, the zebrafish and mammalian homologs of Caenorhabditis elegans fox-1 were identified as recognizing the [UGCAUG element [Jin et al. 2003; Ponthier et al. 2006]. In C. elegans, the fox-1 gene is critical in the sex determination pathway for X-chromosome dosage compensation [Meyer 2000]. In mammals, Fox-1 [also known as A2BP1] encodes a RNA-binding protein initially identified as an interacting partner of ataxin-2, and has at least one paralog, Fox-2 [also known as RBM9 or Fkh] [Nakahata and Kawamoto 2005]. Human Fox-1 and Fox-2 share very high sequence homology—100% identity in the RRM domain—and bind to the same RNA element [Supplemental Fig. S1; Kiehl et al. 2001; Jin et al. 2003; Auweter et al. 2006; Ponthier et al. 2006]. Both proteins are exclusively or preferentially expressed in brain, heart, and skeletal muscle. In addition, mutation or abnormal expression of Fox-1 has been found in patients with several genetic diseases, including epilepsy, mental retardation [Bhalia et al. 2004], autism [The Autism Genome Project Consortium 2007; Martin et al. 2007; Sebat et al. 2007], and heart disease [Kaynak et al. 2003]. Fox-2 was also implicated in hormone signaling, as a corepressor of tamoxifen activation of the estrogen receptor [Norris et al. 2002]. Therefore, Fox-1/2 are likely essential regulators for tissue-specific splicing, and systematic analysis of their targets may provide important insights into the mechanisms of tissue-specific splicing regulation, the characteristics of SRNs, and their physiological roles.

In this study, we define and characterize the SRNs of Fox-1/2. We combined computational predictions from comparative genomics analysis, with experimental validation and characterization. Strikingly, our analysis revealed thousands of potential Fox-1/2 targets with binding sites highly conserved across vertebrate species. Fox-1/2 can either activate or repress splicing, depending on the locations of their binding sites, and also contribute to more complex splicing patterns. Many of the predicted targets play important roles in neuromuscular functions and disorders. We also discuss the evolution of Fox-1/2-binding sites across different vertebrate lineages and among different human populations, and the potential phenotypic implications.

Results

Comparative genomics analysis defines extensive Fox-1/2 SRNs with high specificity

Several previous studies found the enrichment of the UGCAUG element in conserved intronic sequences flanking brain-specific exons or conserved alternative exons [Brudno et al. 2001; Minovitsky et al. 2005; Sugnet et al. 2006; Voelker and Berglund 2007; Yeo et al. 2007], but it was unclear if the sequence specificity and conservation are sufficient to predict Fox-1/2 targets. To characterize global features of SRNs, we sought to develop an effective method for genome-wide Fox-1/2 target prediction. We searched all human internal exons, and 200 nucleotides [nt] of upstream and downstream intronic flanking [UIF and DIF] regions for conserved UGCAUG elements, taking advantage of 28 sequenced vertebrate genomes [Miller et al. 2007]. The inclusion of many species for comparison, in contrast to previous studies [Minovitsky et al. 2005; Voelker and Berglund 2007; Yeo et al. 2007], was justified by their power in reducing the false-positive rate in pairwise species analysis [Supplemental Fig. S2, see also Supplemental Material]. To account for the different levels of divergence among the vertebrate genomes, we adapted a branch length score (BLS) method [Stark et al. 2007] to measure the conservation level of each UGCAUG element, as summarized in Supplemental Figure S3. Several examples of elements with different levels of conservation, along with their BLS scores, are shown below (see Fig. 6C, below). The false discovery rate (FDR) associated with each BLS score was then determined using shuffled random motifs as controls.

To determine appropriate BLS thresholds, we initially studied cassette exons, the most frequent form of AS in mammals [Thanaraj et al. 2004]. The conserved fraction of Fox-1/2 sites in UIF and DIF sequences is higher than that of random sites for the whole range of BLS thresholds [Fig. 1A, left and right panels], which confirms and
We then applied these thresholds to all human internal exons, and predicted 1706 target exons (1457 after overlapping exon variants were merged) from 1103 genes, with at least one conserved Fox-1/2-binding site [FDR = 24%] (Supplemental Table S1). This includes 192 exons with two or more sites in the same or different regions [FDR = 0.03]. Compared with 10 known Fox-1/2 targets determined by previous studies, our predictions successfully included six; the other four were missed because their binding sites are too far away from the exons or do not reach the conservation level we required, or the exon was not included in our database [Supplemental Table S2].

The predicted Fox-1/2 target exons have a number of characteristics similar to those of known regulated tissue-specific exons. Overall, 757 exons are alternatively spliced [Fig. 1B, left panel], and this proportion is significantly higher than the overall fraction of AS exons [Fig. 1B, right panel] in the human genome (44.4% vs. 25.7%, \( P = 5 \times 10^{-63} \)). Among predicted targets with AS are 544 cassette exons [including some with multiple types of AS], which is a significant over-representation, compared with the expected proportion estimated from all AS exons (48.6% vs. 29.6%, \( P = 1.4 \times 10^{-27} \)). A more stringent comparison to exons with conserved random motifs gave qualitatively similar results [data not shown]. Importantly, the AS exons predicted as Fox-1/2 targets, and cassette exons in particular, have more conserved AS patterns in mouse/rat [50.7% vs. 10%–20%] [Blencow 2006], smaller exon sizes (75 nt vs. 110 nt median), and a higher tendency to preserve the reading frame (67.5% vs. 42%), compared with all cassette exons, consistent with previous observations from tissue-specific exons [Blencow 2006; Sugnet et al. 2006; Fangnani et al. 2007]. We expect that the proportion of AS exons we observed is an underestimate, because many AS exons with low EST coverage may be misclassified as constitutive exons. Among the exons currently without evidence of AS, 83 (8.7%) have mouse and/or rat orthologous exons associated with AS events, and 176 (18.5%) are predicted as alternative conserved exons [ACEs] [Yeo et al. 2005]. In summary, comparative analysis of multiple genomes provides an effective way to predict functional Fox-1/2 targets.

Different types of AS events correlate with distinct patterns of Fox-1/2 motif distribution

Since all typical types of alternative exons and splice sites are present in our predicted Fox-1/2 targets, we examined the distribution of Fox-1/2-binding sites separately for each type. We found that the positional preference of Fox-1/2-binding sites differs among different types of AS events. For cassette exons, conserved Fox-1/2-binding sites are 1.75-fold more enriched in the DIF vs. the UIF region [Fig. 1C], in contrast to the approximately equal enrichment for conserved random sites [\( P = 1.8 \times 10^{-8} \)]. No preference for the UIF and DIF regions was observed in the case of mutually exclusive or constitutive exons (Fig. 1D,G). Interestingly, for exons

Figure 1. Comparative analysis accurately predicts Fox-1/2 targets. [A, left axis] The number of conserved Fox-1/2-binding sites [blue] and random-motif sites [gray] in UIF [left], exonic [middle], and DIF [right] sequences of cassette exons, using varying thresholds of BLSs. Error bars represent standard error of the mean. [Right axis] The corresponding FDR of prediction is shown in red. The thresholds used in the paper (0.22 for UIF and DIF sites and 0.8 for exonic sites) are indicated by arrowheads. [B] Proportions of different types of splicing patterns for predicted Fox-1/2 targets [left] and all internal exons [right]. [C–G] Distribution of conserved Fox-1/2-binding sites in different regions for cassette exons [C], mutually-exclusive exons [D], alternative 5’ splice sites [E], and 3’ splice sites [F], and constitutive exons [G]. In each panel, the splicing pattern is shown schematically above the histogram. The distribution of conserved Fox-1/2 sites is color-coded as in B. The distribution of conserved random-motif sites is shown in gray for comparison.

extends previous observations [Brudno et al. 2001; Minovitsky et al. 2005; Voelker and Berglund 2007; Yeo et al. 2007]. Importantly, the comparative analysis of vertebrate species makes it possible to achieve very low FDRs. Excess Fox-1/2 site conservation in the exons was observed only when we applied very stringent BLS thresholds, which correspond to conservation beyond mammals [Fig. 1A, middle panel; Supplemental Fig. S2B]. As a trade-off between sensitivity and specificity of target prediction, we required a BLS threshold of 0.22 for UIF and DIF sites, corresponding to an FDR of 0.24 and 0.15, respectively; for exonic sites, a BLS threshold of 0.8 was required to achieve comparable specificity [FDR = 0.24] [Fig. 1A, indicated by arrowheads].
with alternative 5' and 3' splice sites, conserved Fox-1/2 sites tend to be more enriched in the intron involved in alternative splice site selection [Fig. 1E,F]. This is particularly true for alternative 5' splice sites, for which the DIF region has 3.9-fold more putative binding sites than the UIF region ($P = 2.7 \times 10^{-5}$) [Fig. 1E]. This preference is only partly explained by the generally higher level of sequence conservation in intronic sequences regulating alternative splice site selection. As another line of evidence, we examined the distribution of Fox-1/2-binding sites for all exons with alternative 5' or 3' splice sites, without requiring cross-species conservation. Again, a different positional preference was observed, with 1.2-fold more DIF sites for alternative 5' splice sites and slightly more UIF sites for alternative 3' splice sites ($P = 0.004$). The preference for Fox-1/2 sites to be located near alternative splice sites suggests that Fox-1/2 may play an important role in the differential selection of alternative splice sites in a tissue-specific manner.

**Splicing patterns of Fox-1/2 targets across tissues suggest position-dependent and combinatorial regulation**

We next asked how Fox-1/2 enhance or repress target splicing depending on the location of the presumptive binding sites. To address this question in an unbiased manner, we examined the splicing patterns of predicted Fox-1/2 targets in a panel of 47 tissues and cell lines [Supplemental Table S3], as measured using prototype splicing microarrays. This microarray platform includes both exon and exon-junction probes, which interrogate constitutively and alternatively spliced regions detected from EST/cDNA data, and can therefore monitor the expression of both genes and individual mRNA isoforms genome-wide [J.C. Castle, C. Zhang, J.K. Shah, A.V. Kulkarni, T.A. Cooper, and J.M. Johnson, in prep.].

The splicing microarray was designed independently of this study for a different purpose. Therefore, among the 544 cassette exons we predicted as Fox-1/2 targets, only 234 exons are represented on the array (Fig. 2A; Supplemental Table S4). For each of these exons, we used a splicing index to measure the reciprocal change of exon inclusion level in each particular condition, relative to a reference pool [Ule et al. 2005b]. Hierarchical clustering of tissues using the splicing data successfully grouped brain, muscle, and heart tissues in one cluster, and other tissues in the other cluster, consistent with the restrictive or preferential expression of Fox-1/2 in the first group (Fig. 2A,B).

A majority (62%) of the cassette exons showed higher inclusion in brain, heart, and skeletal muscles, compared with other tissues ($P = 0.0004$). Correlating with a 1.6-fold enrichment of predicted Fox-1/2 sites in the DIF

![Figure 2](https://genesdev.cshlp.org/content/2553/Fig2.png)

**Figure 2.** Splicing profiling of predicted Fox-1/2 targets shows position-dependent and complex modes of Fox-1/2-mediated splicing regulation. (A) Hierarchical clustering of splicing indices of 234 cassette exons predicted as Fox-1/2 targets in 47 human tissues and cell lines (for color scale, see bottom of B). The tissue cluster on the right includes brain, heart, and skeletal muscle tissues is labeled. For each exon, the number of conserved Fox-1/2-binding sites in UIF, exonic, and DIF sequences is gray-scale coded on the right, in the same order as in the splicing heatmap (grayscale on the right). Four clusters of exons, with different combinations of splicing levels in brain and heart/muscle, are labeled by dashed boxes. (B) The expression pattern of Fox-1/2 in the same order of tissues as in the splicing heatmap [color scale at the bottom]. (C–G) Average splicing profile (left) and distribution of conserved Fox-1/2 sites [right] for all 234 predicted targets (C) and exons belonging to the four clusters [D–G]. (H, left) The average splicing profile of all cassette exons on the splicing microarrays as controls. (Right) The distribution of random-motif sites for all cassette exons was used to test the enrichment/depletion of Fox-1/2 sites in different regions for exon subsets in C–G. The $P$-values from $\chi^2$ tests are also indicated in [C–G].
region vs. the UIF region, DIF sites were generally associated with inclusion of the upstream exon (Fig. 2C), confirming and extending several previous studies [Jin et al. 2003; Underwood et al. 2005; Sugnet et al. 2006; Das et al. 2007].

We further examined four exon clusters with different combinations of splicing patterns in brain and heart/skeletal muscle tissues: (1) exons specifically included in both brain and heart/muscle [B[+]M[+]]; (2) exons specifically included only in brain [B[+]M[−]]; (3) exons specifically skipped in brain and muscle/heart [B[−]M[−]]; and (4) exons specifically skipped in brain [B[−]M[+]] (Fig. 2A). We reasoned that exons mainly regulated by Fox-1/2 should have consistent splicing patterns in brain and heart/muscle, because Fox-1 and Fox-2 have high expression in these tissues. Therefore, focusing on the B[+]M[+] and B[−]M[−] clusters should minimize complications due to other factors, and help to infer the rules for Fox-1/2-mediated splicing, as influenced by the locations of the binding sites. Indeed, these two clusters have very distinct spatial distributions of putative Fox-1/2-binding sites (Fig. 2D,F). The B[+]M[+] cluster (Fig. 2D) shows a very strong tendency for the sites being located in the DIF region, rather than the UIF region [sixfold enrichment, \( P = 0.007 \), compared with random motif sites] (Fig. 2H). This bias suggests that downstream Fox-1/2-binding sites are potent splicing enhancers in general. In contrast, for the B[−]M[−] cluster, we observed an opposite pattern of Fox-1/2-binding site distribution, with a ninefold enrichment in the UIF region (Fig. 2F) (\( P = 0.0007 \), compared with random motif sites). This in turn strongly suggests that upstream binding sites repress exon inclusion. Therefore, our analysis provides strong evidence that the different effects of Fox-1/2 in splicing regulation generally depend on the locations of the binding sites.

The other two clusters are also intriguing and indicative of combinatorial regulation. In addition to Fox-1/2, other splicing factors may play an important role in tissue-specific splicing of these exons. For example, for the exons in the B[−]M[−] cluster (Fig. 2G), Fox-1/2-binding sites are significantly enriched in the DIF region (\( P = 4 \times 10^{-9} \)), to an extent similar to that observed in the B[+]M[+] cluster. However, these exons show very low inclusion in brain. This pattern could be explained by brain-specific repressors that counteract the enhancing effects of Fox-1/2, or by muscle/heart-specific coactivators that promote exon inclusion together with Fox-1/2.

To confirm the splicing patterns observed from the splicing microarrays, we performed semiquantitative RT–PCR assays of several endogenous transcripts. In all six cases we tested, cassette exons were conserved with downstream intronic sites (FMNL3, PTBP2, and UAP1) showed brain- and/or muscle/heart-specific exon inclusion, whereas those with only conserved upstream intronic sites (PB1, two exons from MBNL1) showed brain- and/or muscle/heart-specific skipping (Fig. 3). Of particular interest is the UAP1 exon, which was included in muscle and heart, but predominantly skipped in brain, despite the presence of multiple downstream binding sites. Different exon-inclusion levels in brain and muscle/heart were also observed for MBNL1 exons. The splicing patterns observed by RT–PCR are generally consistent with those observed with the splicing microarrays, although a quantitative assessment will require additional experiments. More importantly, the results support the idea that Fox-1/2 alone are not always sufficient to determine the tissue-specific splicing pattern of target genes.

**Overexpression and knockdown of Fox-1/2 alter the splicing of predicted Fox-1/2 targets depending on the presence of the UGCAUG element**

To validate the predicted targets and the position-dependent effect of Fox-1/2-binding sites more directly, we next tested the splicing of endogenous pre-mRNAs in the presence or absence of Fox-1/2 proteins. We first examined Fox-1/2 expression in several cell lines by Western blotting. In all the cell lines we tested, including a few neuronal cell lines, we found variable levels of Fox-2 protein, but no Fox-1 [data not shown]. Because RNAi is very effective in HeLa cells, which express a low but readily detectable level of Fox-2, we used this cell line to test for AS of our predicted targets. For comparison, we generated HeLa cell derivatives expressing different combinations of Fox-1/2—i.e., neither Fox-1 nor Fox-2, Fox-1 only, Fox-2 only, and both Fox-1 and Fox-2—by stable transduction of shRNAs targeting Fox-1 or Fox-2, and/or Fox-1/2 cDNAs in retroviral vectors. Initial examination using several predicted targets showed that expression of either Fox-1 or Fox-2 robustly changed splicing of target exons, compared with the cells without Fox-1/2, in a way similar to that in cells expressing both
knockdown of Fox-2. (Lane 2) untreated HeLa cells. (Lane 3) HeLa cells with shRNA knockdown of Fox-2 and stable transduction of Fox-1 cDNA. (B,C) Radiative RT–PCR analysis of predicted Fox-1/2 targets with downstream intronic binding sites [B], or with only upstream intronic binding sites [C]. All examples are cassette exons. For each exon, the gene symbol is shown below, together with the number of conserved Fox-1/2-binding sites in UIF, exonic, and DIF sequences. In each panel, the quantified exon-inclusion level and the number of conserved Fox-1/2-binding sites in UIF, exonic, and DIF UGCAUG elements are indicated above and below the gene symbol, respectively. The size of each PCR product is also labeled. For some of the genes, indicated by an asterisk, the splicing pattern in tissues was also measured by RT–PCR, as shown in Figure 3.

From our list of predicted targets, we chose to test those with binding sites with a wide range of conservation scores, but favored genes involved in gene expression regulation, such as transcription and RNA processing, and with links to genetic diseases. Among 33 tested cassette exons with conserved downstream intronic sites and detectable expression in HeLa cells, 18 (55%) clearly gave a higher level of exon inclusion in the presence of Fox-1 or Fox-2 [Fig. 4B, Supplemental Fig. S5; Supplemental Table S5], whereas the rest did not show a discernible change; none of them gave a reduction in exon inclusion. Among the 22 tested cassette exons with only conserved upstream intronic sites, 13 (59%) showed a clear change of inclusion level when Fox-1 or Fox-2 was expressed [Fig. 4C; Supplemental Fig. S6; Supplemental Table S5]. With the exception of PLOD2, all of these exons [12 of 13] were repressed by Fox-1/2. These results suggest a satisfactory validation rate of 55%–60%, given the fact that we started from pure computational predictions. Consistent with the splicing microarray data, these analyses strongly indicate the enhancer character of downstream intronic sites, and silencer character of upstream intronic sites, that are predictive of Fox-1/2-regulated splicing.

We next tested if Fox-1/2-regulated splicing depends on the presence of UGCAUG element[s]. To this end, we generated two minigene constructs: one from the FMNL3 gene [Fig. 5A] and the other from the PB1 gene [Fig. 5C]. The FMNL3 minigene comprises the two constitutive exons flanking the cassette exon and both introns [Fig. 5A]. Inclusion or skipping of the cassette exon results in the use of different stop codons. In the wild-type minigene, there are four conserved Fox-1/2 sites downstream from the cassette exon. The exon inclusion level greatly increased when Fox-1 was overexpressed, recapitulating the splicing pattern of the endogenous gene [Fig. 5B, lanes 1,2]. In contrast, Fox-1-mediated exon inclusion became much weaker or completely disappeared when two of the sites [Fig. 5B, lanes 3,4 with mutations in sites 1 and 2, lanes 6,7 with mutations in sites 3 and 4] or all four sites [Fig. 5B, lanes 7,8] were mutated.

The PB1 minigene is a chimeric construct consisting of the PB1 cassette exon with partial flanking introns (~250 nt from the upstream and downstream introns, respectively) inserted into intron 1 of a human β-globin gene splicing reporter [Fig. 5C]. There are three conserved Fox-1/2 sites upstream of the cassette exon. As shown in Figure 5D, overexpression of Fox-1 strongly inhibited inclusion of the cassette exon [lanes 1,2]. When we mutated one or more of the UGCAUG sites, the inhibitory effect of Fox-1 was reduced or eliminated [Fig. 5D, lanes 3–10]. Therefore, Fox-1/2-mediated AS of both FMNL3 and PB1 genes depends on the presence of UGCAUG elements.

Predicted Fox-1/2 targets are enriched in genes with neuromuscular functions

The large number of predicted targets raises the important question of how the Fox-1/2 SRNs are organized to perform cellular functions. To achieve a better understanding of the SRNs, we examined GO terms enriched in the predicted Fox-1/2 target genes, in comparison with a control gene set derived from exons with a similar conservation level [see Materials and Methods]. Many Fox-1/2 target genes have neuromuscular functions, which is reflected in top GO terms related to cytoskeleton organization, ion channels, protein phosphorylation, muscle contraction, etc. [Table 1]. Some of these GO terms were identified previously from smaller sets of brain/muscle-
specific exons, or exons with UGCAUG elements (Das et al. 2007; Fagnani et al. 2007; Yeo et al. 2007), likewise suggesting similar functions of many predicted targets.

Consistently, disruptions of our predicted Fox-1/2 target genes were previously implicated in neurological, neurodegenerative, and sensory disorders, as well as heart

Table 1. Representative GO functions of predicted Fox-1/2 target genes

<table>
<thead>
<tr>
<th>Biological process</th>
<th>Count</th>
<th>P-value</th>
<th>Fold change</th>
<th>FDR (Benjamini)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytoskeleton organization and biogenesis</td>
<td>68</td>
<td>1.1E-11</td>
<td>2.4</td>
<td>3.7E-08</td>
</tr>
<tr>
<td>Actin filament-based process</td>
<td>37</td>
<td>6.5E-09</td>
<td>2.9</td>
<td>1.1E-05</td>
</tr>
<tr>
<td>Potassium ion transport</td>
<td>31</td>
<td>1.7E-07</td>
<td>2.9</td>
<td>1.1E-04</td>
</tr>
<tr>
<td>Metal ion transport</td>
<td>56</td>
<td>2.0E-07</td>
<td>2.1</td>
<td>1.1E-04</td>
</tr>
<tr>
<td>Ion transport</td>
<td>91</td>
<td>3.0E-07</td>
<td>1.7</td>
<td>1.4E-04</td>
</tr>
<tr>
<td>Cation transport</td>
<td>65</td>
<td>2.2E-06</td>
<td>1.8</td>
<td>9.0E-04</td>
</tr>
<tr>
<td>System development</td>
<td>61</td>
<td>2.6E-06</td>
<td>1.9</td>
<td>9.6E-04</td>
</tr>
<tr>
<td>Nervous system development</td>
<td>59</td>
<td>8.3E-06</td>
<td>1.8</td>
<td>2.1E-03</td>
</tr>
<tr>
<td>Muscle contraction</td>
<td>28</td>
<td>1.8E-05</td>
<td>2.5</td>
<td>4.3E-03</td>
</tr>
<tr>
<td>Protein amino acid phosphorylation</td>
<td>70</td>
<td>4.2E-05</td>
<td>1.6</td>
<td>6.9E-03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cellular component</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytoskeleton</td>
<td>123</td>
<td>2.5E-15</td>
<td>2.1</td>
<td>1.5E-12</td>
</tr>
<tr>
<td>Actin cytoskeleton</td>
<td>51</td>
<td>1.2E-13</td>
<td>3.2</td>
<td>3.5E-11</td>
</tr>
<tr>
<td>Non-membrane-bound organelle</td>
<td>148</td>
<td>7.1E-09</td>
<td>1.6</td>
<td>1.1E-06</td>
</tr>
<tr>
<td>Myofibril</td>
<td>15</td>
<td>3.2E-08</td>
<td>5.9</td>
<td>3.9E-06</td>
</tr>
<tr>
<td>Synapse</td>
<td>30</td>
<td>9.2E-08</td>
<td>3.0</td>
<td>9.4E-06</td>
</tr>
<tr>
<td>Myosin</td>
<td>18</td>
<td>8.0E-07</td>
<td>4.0</td>
<td>7.0E-05</td>
</tr>
<tr>
<td>Sarcomere</td>
<td>13</td>
<td>1.0E-06</td>
<td>5.5</td>
<td>6.7E-05</td>
</tr>
<tr>
<td>Microtubule-associated complex</td>
<td>24</td>
<td>1.7E-06</td>
<td>3.1</td>
<td>9.7E-05</td>
</tr>
<tr>
<td>Striated muscle thick filament</td>
<td>9</td>
<td>9.1E-06</td>
<td>7.1</td>
<td>4.3E-04</td>
</tr>
<tr>
<td>A band</td>
<td>9</td>
<td>9.1E-06</td>
<td>7.1</td>
<td>4.3E-04</td>
</tr>
<tr>
<td>Post-synaptic membrane</td>
<td>18</td>
<td>8.8E-05</td>
<td>2.9</td>
<td>3.6E-03</td>
</tr>
<tr>
<td>Golgi-associated vesicle membrane</td>
<td>9</td>
<td>5.3E-04</td>
<td>4.5</td>
<td>1.9E-02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Molecular function</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytoskeletal protein binding</td>
<td>79</td>
<td>4.4E-19</td>
<td>3.0</td>
<td>1.1E-15</td>
</tr>
<tr>
<td>Actin binding</td>
<td>55</td>
<td>2.0E-13</td>
<td>3.0</td>
<td>1.6E-10</td>
</tr>
<tr>
<td>Motor activity</td>
<td>38</td>
<td>1.4E-12</td>
<td>3.7</td>
<td>8.8E-10</td>
</tr>
<tr>
<td>Calmodulin binding</td>
<td>37</td>
<td>2.4E-12</td>
<td>3.8</td>
<td>1.2E-09</td>
</tr>
<tr>
<td>Ion channel activity</td>
<td>53</td>
<td>3.0E-07</td>
<td>2.1</td>
<td>1.0E-04</td>
</tr>
<tr>
<td>Enzyme binding</td>
<td>35</td>
<td>2.3E-06</td>
<td>2.4</td>
<td>5.6E-04</td>
</tr>
</tbody>
</table>
disease and muscular dystrophy, according to the OMIM database [Supplemental Table S5; McKusick 1998]. Therefore, our systematic results support and extend several scattered observations in the literature (Kaynak et al. 2003; Bhalla et al. 2004; The Autism Genome Project Consortium 2007; Martin et al. 2007; Sebat et al. 2007). Interestingly, we found that predicted Fox-1/2 targets are more likely than expected by chance to be disease genes: 157 of 1103 predicted target genes (14.2%) are annotated in the OMIM database as disease genes, compared with a proportion of 7.8% for all genes ($P = 8.3 \times 10^{-14}$), or 10.7% for control genes with a comparable conservation level ($P = 0.0001$) [Supplemental Table S7]. This correlation is indicative of the potential pathological impact when conserved tissue-specific SRNs are dysregulated. We also note that the predicted target genes also have many more introns than the control genes (15 vs. eight, Wilcoxon rank sum test). Disease genes tend to be more intron-rich than average, and are thus more susceptible to splicing dysregulation (López-Bigas et al. 2005), which could contribute to the association between Fox-1/2 targets and OMIM phenotypes we observe.

Creation and loss of Fox-1/2-binding sites may contribute to fine-tuning of gene expression

The large number of species included in our comparative analysis makes it possible to study not only the conservation, but also the evolutionary creation and loss of Fox-1/2-binding sites in specific lineages. Here we mainly analyzed the intronic sites, due to the difficulty in decoupling the mixed selective pressures in coding exons. About 17% of the predicted intronic binding sites are conserved in at least one of the five fish species we analyzed, including those in UAP1, Muscleblind-like genes, PBX1, NLGN3, and others. In contrast, ~19% of the sites are conserved only in mammals. We note that these estimates are biased, due to the artificial enrichment of more conserved sites in our prediction. Nevertheless, they point to evolutionary changes in Fox-1/2 splicing regulation, which may contribute to phenotypic differences across different species, or among different individuals in human populations, as illustrated below.

The first example is a 34-nt exon from PTB (exon 11) and nPTB (exon 10), which results in frame-shifting and nonsense-mediated mRNA decay [NMD] when skipped, and is critical for expression of full-length functional products from both genes [Fig. 6A,C; Supplemental Fig. S7B; Boutz et al. 2007b]. Together with regulation by microRNAs, controlled splicing of this exon through autoregulation switches expression from PTB to nPTB in developing neurons, which in turn results in reprogramming of the splicing patterns of their target RNAs (Boutz et al. 2007a,b; Makeyev et al. 2007). In addition to these negative regulators, we found that Fox-1/2 strongly activate inclusion of the nPTB exon, presumably by interacting with the two DIF sites, D-I and D-II, in contrast, the effect of Fox-1/2 on the paralogous PTB exon is more subtle [Fig. 6A]. Comparison of the two paralogs revealed
that PTB has a T-to-C substitution at the first position of the proximal D-I site, creating a CGCAUG element with weaker affinity for Fox-1/2 [Fig. 6 A,C; Supplemental Fig. S7B]. The loss of UGCAUG occurred in the last common ancestor of placental mammals, because an intact UGCAUG element is preserved in four nonmammalian vertebrates. Such a lineage-specific alteration may have contributed to the evolution of PTB SRNs in mammals.

The second example is a 36-n cassetexon from three Muscleblind-like genes, MBNL1, MBNL2, and MBNL3 [Fig. 6B,C; Supplemental Fig. S7C]. We predicted and validated all three paralogous exons as Fox-1/2 targets [Fig. 6B,C]. In the MBNL1 and MBNL2 exons, there are two Fox-1/2-binding sites as putative silencers: one overlapping with the polypyrimidine tract (−14 to −9, denoted as U-I) and the other in the exonic region (12–17, denoted as E-II). Both sites are conserved in almost all vertebrate species we analyzed. Interestingly, the U-I site upstream of the MBNL3 exon is polymorphic in human populations: It overlaps with an A/G SNP [rs3736748] at the fourth position, resulting in two alleles UGC[A/G]UG, although the site is conserved in most vertebrates. In addition, another site (U-III) was created further upstream from the exon [Fig. 6B, Supplemental Fig. S7C]. Interestingly, the allele frequency of the SNP differs radically in different populations, according to HapMap data (x² = 153, P = 6 × 10−34) [The International HapMap Consortium 2007]. Consequently, the Fox-1/2-binding site is intact in most Africans [YRI], but is disrupted in most Asians [HCB/JPT], with Europeans (CEU) somewhere in between. We validated by direct sequencing that the U-I site is intact in HeLa cells [African origin], which is consistent with the increased exon skipping upon Fox-1/2 expression. This example provides a good model to study how genetic variations may affect splicing regulation and result in phenotypic differences among individuals.

In these two examples, the paralogous intronic sequences, especially the Fox-1/2-binding sites, can still be aligned, despite considerable nucleotide substitutions. We found more examples belonging to this category, including another exon pair from MBNL1/2 [Supplemental Fig. S7D], NLGN3/4X/4Y [Supplemental Fig. S7E], and EBP41/412 [Supplemental Fig. S7F]. However, this is not always the case: In two pairs [or trios], one from Fox-1/2 [Supplemental Fig. S7A] and the other from ELAVL2/3/4 [Supplemental Fig. S7G], the intronic sequences, including the putative Fox-1/2-binding sites, are very difficult to align. Since the Fox-1/2 sites in each paralog are significantly conserved across vertebrate species, the creation/loss of putative binding sites occurred very early after gene duplication, and was then fixed in the descendant species. Therefore, gene duplication can be followed by splicing divergence, providing parallel paths for producing genetic diversification.

Discussion
Extensiveness of Fox-1/2 SRNs

In this study, we used the highly conserved and related brain-, heart-, and muscle-specific splicing factors Fox-1 and Fox-2 as a model to predict their RNA targets, and to define and characterize their SRNs. By comprehensive phylogenetic analysis of the specific Fox-1/2 motif in 28 vertebrate species, we predicted thousands of target exons and genes with conserved Fox-1/2-binding sites in vertebrates. We estimate by statistical analysis that ~76% of the predicted targets are bona fide targets, and indeed, more than half of a set of 55 predicted targets could be validated experimentally by manipulating Fox-1/2 expression in HeLa cells. The extent of the SRN is comparable with the gene regulatory networks of certain master transcription factors [Massie and Mills 2008], which is surprising, in light of the very limited number of targets of splicing factors identified to date [for review, see Li et al. 2007].

Our predictions may still represent an underestimate, because we focused only on the conserved components of the Fox-1/2 SRNs that can be predicted with high specificity and sensitivity. Many additional binding sites with a relatively low level of conservation might also be functional, as we observed experimentally [data not shown]. Moreover, in some extreme cases, a functional Fox-1/2-binding site can be thousands of nucleotides away from the regulated exon [Nakahata and Kawamoto 2005], such targets would have escaped our predictions. Furthermore, our predictions included several splicing factors involved in brain- and/or muscle-specific splicing, such as Fox-1/2, PTBP1/2 [PTB/nPTB], CUGBP1/2, NOVA1, and MBNL1/2/3. Therefore, potential regulation of these splicing factors by Fox-1/2 implies that the Fox-1/2 SRNs are not limited to direct targets, but probably include a large number of indirect targets as well. Our study highlights the importance of tissue-specific splicing in diversifying gene expression regulation.

Mechanisms of Fox-1/2-dependent exon activation and repression

The effect of splicing factors on splicing activation or repression often depends on the location and context of the regulatory sequences they bind. This was reported both for ubiquitous splicing factors, such as SR proteins [Ibrahim et al. 2005] and hnRNPs [Hung et al. 2008], as well as for tissue-specific splicing factors [Ule et al. 2006]. Nova-1/2 regulate target exon inclusion and skipping in a predictable way: Downstream intronic binding sites are usually enhancers, whereas upstream and exonic binding sites are usually silencers. A similar mechanism was already proposed for Fox-1/2, based on studies of a few exons, such as ATP5C1 [F1γ] exon 9, fibronectin EIIIB exon, c-src EWS exon 4′ [Jin et al. 2003; Underwood et al. 2005]. The enhancer character of downstream Fox-1/2 sites was confirmed on a larger scale by several studies focusing on brain- or muscle-specific exons [Sugnet et al. 2004; Das et al. 2007]. However, the silencer character of upstream binding sites was not obvious from these studies. Therefore, it was unclear whether there is a general mechanism predictive of Fox-1/2-mediated splicing patterns. Our comprehensive prediction of Fox-1/2 targets, followed by unbiased experi-
mental validation, leads to a conclusive answer: Our analyses of splicing microarrays, and RT–PCR data in primary tissues and in HeLa cells in the presence and absence of Fox-1/2, very consistently demonstrate that downstream intronic Fox-1/2 sites are potent enhancers, whereas upstream intronic sites have the opposite effect.

More experiments are required to understand the underlying mechanisms of how Fox-1/2 interact with the spliceosome to modulate splicing. Putative Fox-1/2-binding sites are generally more enriched in the downstream intron, with a distribution peak ~30 nt from the exon; the upstream intron shows a smaller enrichment with a broader distribution [Yeo et al. 2007; data not shown]. Interestingly, for alternative 5′ splice sites, putative Fox-1/2-binding sites show a strong preferential location in the DIF region, whereas for alternative 3′ splice sites, a more moderate enrichment in the UIF region is seen. The preferential enrichment at particular distances downstream from 5′ splice sites suggests that Fox-1/2 might be more efficient at enhancing 5′ splice site recognition. In contrast, the mechanisms through which Fox-1/2 block exon recognition might be more heterogeneous. For example, in the context of hF17 gene exon 9, Fox-1 binding to a GCAUG element in intron 8 blocks prespliceosomal E-complex formation in intron 9, resulting in skipping of exon 9 [Fukumura et al. 2007]. We found cases [e.g., exons from Muscleblind-like genes] in which the Fox-1/2-binding sites in the UIF region are very close to the downstream 3′ splice site. In these cases, Fox-1/2 likely block the recognition of the intron preceding the alternative exon by interfering with binding of spliceosomal components that recognize the polymeridimine tract, 3′ splice site, and/or branch site. Furthermore, for alternative exons with multiple Fox-1/2-binding sites, these sites are not equivalent in Fox-1/2-dependent activation or repression of exon inclusion. Rather, the sites closer to the splice sites appear to be more efficient than the distal sites, at least with the two minigenes we tested. One interpretation is that Fox-1/2 proteins bound to the proximal sites are more efficient at directly interacting with spliceosomal components.

The potential functional differences between isoforms of Fox-1/2, as well as between the two paralogs, on target recognition and splicing are another interesting question. Fox-2 is reported to autoregulate its expression by repressing the inclusion of exon 6 [Baraniak et al. 2006]. Indeed we predicted this exon as a target; in addition, we also predicted four other exons in Fox-1, and one other exon in Fox-2, as potential targets for autoregulation. Among them are one of the mutually exclusive exons in Fox-1 and its paralogous exon in Fox-2 [Supplemental Fig. S7A]. The Fox-1 exon, dubbed B40, is specifically included in brain, whereas the other mutually exclusive exon, M43, is specifically included in muscle [Nakahata and Kawamoto 2005]. Therefore, Fox-1/2-mediated AS of these two exons might be important for generating different isoforms of Fox-1/2 proteins in different tissues, which may in turn affect target gene splicing differently.

In terms of the comparison between Fox-1 and Fox-2, they have similarly high expression in brain and heart/muscle tissues, but Fox-2 is also widely expressed in other tissues. However, because Fox-1 and Fox-2 recognize the same RNA element [Jin et al. 2003; Auweter et al. 2006; Ponthier et al. 2006], it is impossible to distinguish Fox-1/2 targets through their predicted binding sites, as in the present study. Overexpression and knockdown of Fox-1 or Fox-2 individually or together in HeLa cells suggest that these two proteins have very similar effects in activating or repressing predicted targets. We expect that further insights will be gained by identification of the in vivo targets of Fox-1 and Fox-2 experimentally in the appropriate tissue types, as well as by determination of the mechanisms of action of these factors.

Complex splicing patterns suggest potential combinatorial regulation

Overexpression and knockdown of Fox-1/2 in HeLa cells, followed by RT–PCR analysis, demonstrated that our predictions of Fox-1/2 targets are amenable to more detailed experimental follow-up. However, the observed validation rate was somewhat lower than our statistical estimate, and it was not always clear why some predictions were validated while the others failed. A possible explanation is that the FDR based on permutations might be overoptimistic, due to some type of bias; alternatively, some predicted Fox-1/2 sites may be present within RNA secondary structures that make them inaccessible to Fox-1/2, yet another likely possibility is that Fox-1/2 alone are not always sufficient to affect the splicing pattern of a bona fide target pre-mRNA.

Both splicing microarray analysis and RT–PCR validations suggest the existence of exons with complex splicing patterns that cannot be explained by Fox-1/2 regulation alone. Although the effect of Fox-1/2 on splicing of these exons is not always observable because of the difficulty in identifying the appropriate tissues or developmental stages, in some cases, the requirement for other cooperating splicing factors in Fox-1/2-mediated splicing regulation is readily apparent. Our argument is based on comparison of splicing patterns between brain and muscle/heart, where Fox-1/2 are highly expressed. Using splicing microarray analysis and RT–PCR validations, we identified clusters of cassette exons with inconsistent splicing patterns among these tissues, an indication of combinatorial regulation. This difference is especially pronounced for a cluster of DIF site-containing exons, which are predominantly included in muscle, as expected, but mostly skipped in brain. One good example is the UAP1 exon, with two downstream intronic binding sites, which was also validated by RT–PCR analysis of primary tissues and HeLa cells. Other splicing activators or repressors might be expressed and functional only in brain or heart/muscle, but not in both. A similar argument might hold for HeLa cells, which would help explain the lack of responsiveness of some exons to manipulating Fox-1/2 expression.

A few splicing factors that interact or have the potential to interact with Fox-1/2 have been reported recently. For example, several additional proteins, including...
mutations would cause genetic diseases. Instead, some mutations can readily occur during evolution. Not all of these mutations might have only moderate effects and can therefore be tolerated. Two examples we examined included PTPB, in which a site was likely lost in all mammalian species, and MBNL3, in which a site was likely lost in Asians, while being preserved in Africans. Although further testing is required, these observations suggest that tissue-specific SRNs might show considerable divergence between mammals and nonmammalian vertebrates, as well as among different human populations.

Materials and methods

Compilation of exons and AS events using splicing graphs

We built a database of classified AS events (dbCASE, http://rulai.cshl.edu/dbCASE) using high-quality transcripts [mRNA/EST] and genome alignment (coverage > 85%, identity > 95%), for human, mouse, rat, and other model organisms (Zhang et al. 2007). Exons, introns, and typical types of AS events were detected using splicing graphs. For this study, we used 204,305 human AG-GT internal exons, with annotations of associated AS events.

For each exon, we analyzed exonic sequences and 200-nt U1F/D1F sequences. Multiple alignments of 28 vertebrate species [Miller et al. 2007] were extracted using the mafFrag program obtained from the University of California at Santa Cruz (UCSC) genome browser [http://genome.ucsc.edu].

For each human AS event, we also tried to identify the orthologous AS event in mouse and rat. This was done by mapping the genomic coordinates of the AS region in mouse or rat to the human genome using the tool liftOver obtained from the UCSC genome browser. For example, for cassette exons, the alternative exons and the two flanking exons were used for the mapping. The mapped coordinates were then compared with the corresponding regions of human AS events.

Evaluation of motif site conservation

We focused on Fox-1/2 targets with UGCAUG present in human, unless explicitly mentioned. A BLS approach [Stark et al. 2007] was adapted to measure Fox-1/2-binding site conservation, using the multiple alignment and phylogenetic tree of 28 vertebrate species [Miller et al. 2007]. Briefly, the BLS of each unique Fox-1/2-binding site (in the same alignment columns) is the total branch length of the phylogenetic tree over which the site is conserved, normalized by the total branch length of the tree spanning all species. We allowed no movement of the sites in the assignment of orthologous sites, given the high-specificity and conservation of Fox-1/2-binding sites. However, in some instances, small insertions/deletions interrupt some sites, partly due to artifacts in sequence alignment [e.g., TGCATGG aligned with TGCAT-G], such indels were tolerated. Therefore, our approach is more restrictive than the original approach [Stark et al. 2007], because we sought to trace the history of each individual site.

To determine the significance of motif site conservation, we estimated the null distribution of BLS using 50 random motifs generated by permutations. Random motifs containing CpG or GCAUG were avoided, because CpGs are underrepresented in vertebrate genomes, and the GCAUG element is at least partially functional for Fox-1/2-mediated splicing [Jin et al. 2003]. The same analysis was repeated for each of the random motifs and motif sites to calculate BLS scores. We tried different BLS
thresholds from 0 to 1, with steps of 0.01 to determine an appropriate threshold for Fox-1/2 target prediction. For each threshold, a FDR was calculated by the ratio of the average number of sites with a BLS greater than the threshold for random motifs to that for the Fox-1/2 motif.

**Experimental validation of predicted Fox-1/2 targets**

From the list of predicted targets, we chose to test genes involved in gene expression regulation, such as transcription and RNA processing, and with links to genetic diseases, but others were also selected at random. Human tissue total RNAs were purchased from Clontech. Two shRNAs against human Fox-2 were cloned into the MSCV retroviral vector as described (Dickins et al. 2005). Human Fox-1 cDNA (NM_018723) was cloned into the pWZL-hygro retroviral vector to express Flag-tagged Fox-1 protein.

To generate stable cell pools, HeLa cells were infected with MSCV (expressing the shRNA against Fox-2), or MSCV plus pWZL-hygro (expressing Fox-1 protein) vectors. We replaced the medium 24 h after infection, and 24 h later, infected cells were selected with puromycin (2 µg mL\(^{-1}\)) for 72 h. In the case of double infection, cells were treated with hygromycin for 96 h after selection with puromycin. The effect of knockdown or overexpression was confirmed by Western blotting using antibodies against human Fox-1 (donated by S. Powers and D. Mu) and Fox-2 (Bethyl Laboratories, Inc.), respectively. Total RNA was extracted from the stable cell pools using Trizol reagent (Invitrogen) and treated with DNase I. Reverse transcription was carried out using SuperScript II reverse transcriptase as described by Invitrogen. Semiquantitative PCR using Taq polymerase was performed by adding 0.1 µL of [\(^{32}\)P]-dCTP to each 25-µL reaction. The PCR reactions were run for 20–25 cycles depending on the abundance of the targets. The products were analyzed on a 6% native polyacrylamide gel. The primer sequences used for validation are shown in Supplemental Table S7.

The FMLN3 minigene was cloned into the pcDNA3.1 vector. QuickChange PCR mutagenesis was carried out to generate the mutant constructs. Fugene 6 was used for transfection and RT-PCR analysis was done as above. The PB1 minigene was generated by inserting a PCR fragment containing the cassette exon plus 243-nt U1F and 253-nt DIF sequences, into intron 1 of the human \(\beta\)-globin gene, using BglII and XhoI sites introduced by site-directed mutagenesis.

**Splicing microarrays**

We identified exons, exon junctions, and AS events in the human genome by mapping RefSeqs, mRNAs, ESTs, and transcripts from patent databases to the genome. For each gene, 60-nt probes and 36-nt probes were optimized to monitor exons and exon junctions, respectively, and printed on Agilent arrays. These arrays were used to monitor 47 diverse human tissues and cell lines in dye-swap replicates (Supplemental Table S3). Gene expression levels were estimated from probes monitoring constitutive exons and junctions. For each AS event, a proportional change of isoform abundances, relative to a reference pool, was then estimated by adapting a previous method (Ule et al. 2005b). More detailed information and data availability will be described elsewhere (J.C. Castle, C. Zhang, J.K. Shah, A.V. Bhalla, T.A. Cooper, and J.M. Johnson, in prep.). The raw and processed microarray data, as well as additional information, are available at GEO [http://www.ncbi.nlm.nih.gov/geo, under accession GSE11863] and on our Web site [http://rulai.cshl.edu/Rosetta_AS_supp/index.html](http://rulai.cshl.edu/Rosetta_AS_supp/index.html) for download, and in dbCASE in a searchable form.

**GO term and OMIM phenotype analyses**

The GO term analysis was performed using the online tool DAVID (Dennis et al. 2003). DAVID gives a P-value, before and after multiple-test corrections, based on a modified hypergeometric distribution. We used a background set of 15,040 genes, controlling for the conservation level. More specifically, a gene was included in the control gene set if at least one of its exons has a consecutive hexanucleotide with a BLS greater than a specified threshold [BLS \(\geq 0.22\) for U1F and DIF sequences and BLS \(\geq 0.8\) for exonic sequences]. The OMIM phenotypes and associated genes were downloaded in December, 2007 (McKusick 1998).

**Statistical analysis**

Fisher’s exact test in the software package R [http://www.r-project.org](http://www.r-project.org) was used to evaluate the significance of two-by-two contingency tables.

**Acknowledgments**

We thank Scott Powers and David Mu for providing Fox-1 antibody, John Conboy for critical reading of the manuscript, and Lily Chan for technical help. This work was supported by NIH grant GM74688 [to M.Q.Z. and A.R.K.] and a post-doctoral fellowship from Damon-Runyon Cancer Research Foundation (Z.Z.).

**References**


Erratum

Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2
Chaolin Zhang, Zuo Zhang, John Castle, Shuying Sun, Jason Johnson, Adrian R. Krainer, and Michael Q. Zhang

In the above-mentioned paper, the year for the received date was incorrect. The correct received/accepted dates are as follows:

Received June 6, 2008; revised version accepted July 28, 2008.

We apologize for the error.
Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2

Chaolin Zhang, Zuo Zhang, John Castle, et al.

*Genes Dev.* 2008, **22**:
Access the most recent version at doi:10.1101/gad.1703108

Supplemental Material

http://genesdev.cshlp.org/content/suppl/2008/09/16/22.18.2550.DC1

Related Content

Erratum

*Genes Dev.* October , 2008 22: 2902

References

This article cites 59 articles, 23 of which can be accessed free at:

http://genesdev.cshlp.org/content/22/18/2550.full.html#ref-list-1

Articles cited in:

http://genesdev.cshlp.org/content/22/18/2550.full.html#related-urls

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.