Identification of MEF2-regulated genes during muscle differentiation

James Paris,* Carl Virtanen,* Zhbin Lu, and Mark Takahashi

Microarray Centre, Clinical Genomics Centre, University Health Network, Toronto, Ontario, Canada

Submitted 30 June 2004; accepted in final form 19 October 2004

Paris, James, Carl Virtanen, Zhbin Lu, and Mark Takahashi.
Identification of MEF2-regulated genes during muscle differentiation. Physiol Genomics 20: 143–151, 2004. First published October 26, 2004; doi:10.1152/physiogenomics.00149.2004.—Although a great deal has been elucidated concerning the mechanisms regulating muscle differentiation, little is known about transcription factor-specific gene regulation. Our understanding of the genetic mechanisms regulating cell differentiation is quite limited. Much of what has been defined centers on regulatory cascades and transcription factors. Surprisingly few studies have investigated the association of genes with specific transcription factors. To address these issues, we have utilized a method coupling chromatin immunoprecipitation and CpG microarrays to characterize the genes associated with MEF2 in differentiating C2C12 cells. Results demonstrated a defined binding pattern over the course of differentiation. Filtered data demonstrated 9 clones to be elevated at 0 h, 792 at 6 h, 163 by 1 day, and 316 at 3 days. Using unbiased selection parameters, we selected a subset of 291 prospective candidates. Clones were sequenced and filtered for removal of redundancy between clones and for the presence of repetitive elements. We were able to place 50 of these on the mouse genome, and 20 were found to be located near well-annotated genes. From this list, previously undefined associations with MEF2 were discovered. Many of these genes represent proteins involved in neurogenesis, neuromuscular junctions, signaling and metabolism. The remaining clones include many full-length cDNA and represent novel gene targets. The results of this study provides for the first time, a unique look at gene regulation at the level of transcription factor binding in differentiating muscle.

MEF2; CpG island microarrays; chromatin immunoprecipitation

OVER THE PAST DECADE a great deal has been elucidated concerning the regulation of muscle differentiation. One area of interest has centered on the characterization of muscle transcription factors. Many of these regulatory factors have been identified and can be separated into two major groups: the bHLH (basic helix-loop-helix) and the MADS-box (MCM1, agamous deficiens, serum response factor) families. However, to date, a clear understanding of the genes that these transcription factors act upon has not been characterized. Recently, a number of groups have employed microarray analysis to characterize the expression profile of differentiating muscle (10, 31, 38, 41). These studies provide valuable information by characterizing changes in the steady-state gene expression in a high-throughput fashion. Not surprisingly, large groups of genes were identified to change throughout differentiation. The challenge is to identify the key regulatory gene or cohort of genes from these lists. At the moment, we can only speculate as to whether changes in gene expression were a consequence of primary or secondary events. Answers to this question may come by quantifying the direct associations between key transcription factors and the genes they regulate.

During muscle differentiation, MEF2 stands out as an excellent candidate due to the key role it plays in regulating muscle gene expression (30). The MEF2 regulatory factors belong to the MADS-box family of transcription factors and exist as four isoforms; A, B, C, and D (6, 13, 27, 35). In the terminal stages of muscle differentiation, MEF2 factors have been found to be essential (4). Ornatsky et al. (34) have shown the necessity of this transcription factor through the use of dominant-negative MEF2 mutants effectively blocking muscle differentiation. Furthermore, MADS box binding sites have been identified upstream of many known muscle-specific genes.

To characterize the associations between transcription factors and genomic DNA, a number of technical hurdles must be overcome. First, there is a need for a high-throughput method to analyze the large number of potential associations. Microarrays provide high-volume analysis; however, an appropriate target would be required. Ideally, having the promoter for every gene would allow us to profile the binding pattern across the entire genome, but our ability to predict these sites at the moment is not efficient. By taking advantage of a naturally occurring genomic sequence, we may be able to arrive at some answers. Stretches of cytosines and guanines (>200 bp in length) termed CpG islands have been identified to be located near the 5′ promoter regions of genes (8). Because of the propensity of cytosine nucleotides to be methylated, these are deaminated and converted to a thymidine residue, thus disrupting these clusters. The exception to this occurs when a selective pressure such as transcriptional regulation exists to suppress methylation in these regions. In the mouse it is approximated that there are 37,000 CpG islands, whereas the human genome contains ~45,000 (2). A well-established methodology for the isolation of protein associated DNA already exists through chromatin immunoprecipitation (ChIP). By combining ChIP with CpG microarrays, high-throughput analysis of transcription factor binding can be characterized. Indeed, several groups have recently successfully utilized this methodology to identify the genomic binding pattern of a number of specific transcription factors (25, 44, 45).

It was the goal of the present study to characterize the primary gene associations with a key transcription factor during terminal muscle differentiation. To accomplish this goal, we constructed a mouse CpG island microarray and used ChIP to characterize those genes associated with MEF2 binding. Our findings demonstrate for the first time a unique subset of genes associated with MEF2.

*J. Paris and C. Virtanen contributed equally to this work.
Article published online before print. See web site for date of publication (http://physiogenomics.physiology.org).
Address for reprint requests and other correspondence: M. Takahashi, Microarray Centre, Clinical Genomics Centre, Univ. Health Network, 200 Elizabeth St., MBRC 5R-414, Toronto, Ontario M5G 2C4, Canada (E-mail: markt@uhnres.utoronto.ca).
MATERIALS AND METHODS

Construction of mouse Cpg microarray and arraying procedure. An aliquot of the complete mouse CGI library was obtained from the MRC Rosalind Franklin Centre for Genomics Research (RFCGR; Babraham, Cambridge, UK). The library was plated at the appropriate dilution and grown using blue-white selection. With the use of an automated colony picker (QPix; Genetix, Boston, MA), 7,000 viable colonies were collected and grown. Inserts containing the Cpg islands were amplified from the plasmids by PCR using T7/SP6 primers. Specificity of amplification was confirmed by gel electrophoresis. Purification of the ampiclons was performed using TeleChem filter plates (Sunnyvale, CA) on a Beckman Biomek 2000 robotic workstation (Fullerton, CA). After purification, PCR products were re-arrayed into 384-well polypropylene collection plates from Whatman Polyfiltronics (Rockland, MA). Microarrays were manufactured at the University Health Network Microarray Centre (Clinical Genomics Centre, University Health Network, Toronto, Canada). cDNAs containing the Cpg islands were spotted using the ChipWriter Pro robotic arrayer (Bio-Rad, Waterloo, Canada).

Cell culture. C2C12 murine skeletal muscle cells were obtained from the ATCC (Manassas, VA) and grown on 150-mm plastic dishes (5 per time point) containing Dulbecco’s modified Eagle’s medium supplemented (DMEM) with 10% fetal bovine serum and 1% antibiotics (5 per time point) containing Dulbecco’s modified Eagle’s medium supplemented (DMEM) with 10% fetal bovine serum and 1% antibiotics (5 per time point). C2C12 myoblasts were differentiated into myotubes (100% ethanol at room temperature for 1 h). Sample volumes were then brought up to 100 μl with water, and unincorporated dye was removed by passing through PCR purification columns (Qiagen, Alameda, CA) and eluted with 100 μl water.

MEF2 binding sites for the known genes were amplified following standard PCR parameters (95°C for 2 min, 95°C for 5 min, 60°C for 1 min, 72°C for 1 min) for 36–38 cycles. Primer sequences for each gene are provided in Table 1. All amplicons were run on 2% agarose gels containing 0.4 μg/ml ethidium bromide.

Western blot analysis. Immunoprecipitated samples containing MEF2 antibody or no antibody were run on a 10% SDS-polyacrylamide gel and electrophoresed (semi-dry electrophoretter) for 20 min onto nitrocellulose membranes (Hybond-C, GE Healthcare). Blots were probed with polyclonal antibody directed against MEF2 (1:500, sc-313; Santa Cruz Biotechnology). Donkey anti-rabbit IgG (1:2,500; Santa Cruz Biotechnology) was added where appropriate. Samples were then left to incubate overnight at 4°C with rocking. Protein-DNA-antibody complexes were immunoprecipitated by the addition of an elution buffer (1% SDS, 50 mM NaHCO3) at room temperature for 20 min with shaking. DNA was isolated from the final complex by the addition of 5 M NaCl and 10 μg RNase A (GE Healthcare, Baie d’Urfé, Quebec, Canada) with incubation at 65°C for 5 h. DNA was then recovered by precipitation (100% ethanol at −20°C overnight). Any protein contamination was removed by incubating the sample with protease K (Roche, Mississauga, Canada) at 42°C for 2 h. DNA was then purified by QIAquick PCR purification columns (Qiagen, Alameda, CA) and eluted with 100 μl water.

ChIP and DNA amplification. The protocol as outlined by Iyer et al. (14) was followed with modifications and can be found in detail on our web site (http://www.microarrays.ca). Briefly, proteins were cross-linked to DNA by adding 1% formaldehyde directly to culture medium incubating at 37°C for 10 min. Glycine (125 mM) was then added for 5 min to stop the reaction. Plates were washed twice in ice-cold PBS + 1.0 mM EDTA and kept on ice until all cells were harvested. Cells were collected and washed by passing through a 100 μl cell disruptor to shear the DNA to an average length of ~1,000 bp. Input chromatin concentrations for each time point was normalized by using a Pico Green dsDNA Quantitation Kit (Molecular Probes, Eugene, OR). Samples were first precleared by incubating with 15 μl of S. aureus cells for 15 min at 4°C. BSA (200 μg; Sigma, Oakville, Canada) was added to all samples, and 1 μg of MEF2 antibody (sc-313, lot no. H072; Santa Cruz Biotechnology) was added where appropriate. Samples were then left to incubate overnight at 4°C with rocking. Protein-DNA-antibody complexes were immunoprecipitated by the addition of 20 μl yeast tRNA (Invitrogen) and 10 μl S. aureus cells suspension with 15 min of rocking at room temperature. Cells were sequentially washed twice with dialysis buffer (2 mM EDTA, 50 mM Tris-HCl, pH 8.0) then four times in immunoprecipitation wash buffer (100 mM Tris-HCl pH 8.0, 500 mM LiCl, 1% Igepal CA-630, 1% sodium deoxycholate) for 3 min at room temperature each. Protein-DNA was eluted from the antibody-bead complex by the addition of an elution buffer (1% SDS, 50 mM NaHCO3) at room temperature for 20 min with shaking. DNA was isolated from the final complex by the addition of 5 M NaCl and 10 μg RNase A (GE Healthcare, Baie d’Urfé, Quebec, Canada) with incubation at 65°C for 5 h. DNA was then recovered by precipitation (100% ethanol at −20°C overnight). Any protein contamination was removed by incubating the sample with protease K (Roche, Mississauga, Canada) at 42°C for 2 h. DNA was then purified by QIAquick PCR purification columns (Qiagen, Alameda, CA) and eluted with 100 μl water.

Table 1. Primer sequences used in ChIP reconfirmation experiments

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sense</th>
<th>Antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCK</td>
<td>TAAAGTCGTCCCATCTAGAAAGAG</td>
<td>GTTGGGCTCTAGACGAGACGAG</td>
</tr>
<tr>
<td>MLC</td>
<td>CTGACAGAGCTTCGATCTCACC</td>
<td>TCTTACGGACAGAGGAAGAGAG</td>
</tr>
<tr>
<td>SERCA</td>
<td>CTGGTGTATAGGCTTCATATTG</td>
<td>CTGGGTACCTCCCCCTTGTG</td>
</tr>
<tr>
<td>SERCA</td>
<td>CCCATAGGAGACCATTCGAGC</td>
<td>TGAAGAACCCAGCCTAGGAC</td>
</tr>
<tr>
<td>MYOGENIN</td>
<td>GAAAGCTCTTCAGCTGTCGAG</td>
<td>CTTTTAACCCGCCCTTGCT</td>
</tr>
<tr>
<td>MAP2K2</td>
<td>TGGAGAAGCTCTGGTGAGCTTTGG</td>
<td>GTATCACTATGAGCCTCTTTG</td>
</tr>
<tr>
<td>Mipol</td>
<td>CAGCTGTCAGTTGATTTGCTCATT</td>
<td>GTTTGGTCTACAAGGACGAGAG</td>
</tr>
<tr>
<td>RAGE</td>
<td>TATAGCTCAAAGGCGAGAAATAT</td>
<td>CGAGCCTCTGAGGCTAGAAAG</td>
</tr>
<tr>
<td>Spread2</td>
<td>GATGATGAAAGGACGACGAC</td>
<td>ATACAGATACTACGGTGCACAT</td>
</tr>
<tr>
<td>Riken 8030460</td>
<td>ATCTAGACCCGCGGGAGAGCTATT</td>
<td>TGCGAGCCAGGGAGAGA</td>
</tr>
<tr>
<td>Riken 9430667</td>
<td>TACATCTGAAATGGGTCTCAGGA</td>
<td>CTTTTATTTGTTGTCGTAAT</td>
</tr>
<tr>
<td>Sprouty</td>
<td>GAGACACAGCGGAGGACAGGACT</td>
<td>ATACAAGACTACGGTGCAGACC</td>
</tr>
<tr>
<td>Cleo-2</td>
<td>TTGGCTGCTGCTGAGAATTCTGGTTT</td>
<td>GAAAGGACAGAGACGAGAG</td>
</tr>
</tbody>
</table>

ChIP, chromatin immunoprecipitation.
through a CyScribe GFX purification column and eluting with 60 μl of 0.1 M sodium bicarbonate. Samples were reduced to a final volume of 5 μl in the SpeedVac Plus and used immediately for hybridization.

**Microarray hybridization.** Samples were added to 80 μl of a hybridization mixture containing 100 μl DIG EasyHyb (Roche, Mississauga, Canada) supplemented with 50 μg yeast tRNA and 50 μg salmon sperm DNA. The sample was heated to 65°C for 2 min, cooled, and briefly centrifuged. The entire volume was applied to the microarray and placed in a sealed, humidified hybridization chamber and incubated overnight at 37°C. Slides were then washed (three times for 16 min with 1 × SSC, 0.1% SDS at 50°C, and twice for 5 min in 0.1 × SSC at room temperature) and then centrifuged for 15 min at 640 rpm and scanned.

**Scanning and quantification.** Slides were scanned on a laser fluorescence confocal scanner (ScanArray 4000XL, PerkinElmer). Individual 16-bit TIFF images were obtained by scanning for each of the two fluoros. An overlay image of the two images was created and quantified utilizing the QuantArray (v3.0) program (PerkinElmer). Intensity values for each spot were calculated for both the channels.

**Data warehousing.** Scanned 16-bit TIFF images representing each hybridized microarray slide and the associated quantification data files were entered into the GeneTraffic Microarray Database and Analysis System (Iobion Informatics, La Jolla, CA) and then were analyzed using GeneSpring (Silicon Genetics, Redwood City, CA).

**Confirmation of microarray data by ChIP.** Immunoprecipitated DNA was obtained for each time point as described above with the exception that all triplicates were combined prior to digestion with proteinase K. Primers suitable for PCR were designed for positive control antibodies.

**Quantification utilizing the QuantArray (v3.0) program (PerkinElmer).**

**Muscle differentiation and protein expression.** As a first step, we wished to confirm that the antibody we were using demonstrated specificity for our protein of interest. To accomplish this, we carried out Western blot analysis on the immunoprecipitated samples prior to proteinase K digestion. From this, we were able to identify a single distinct band at the correct molecular weight corresponding to MEF2A (Fig. 1). Although we only observed a single band at the appropriate molecular weight, we cannot exclude the possibility that the antibody was associating with the C and D isoforms as stated by the manufacturer. Therefore, all references to MEF2 will include the A, C, and D isoforms. No band was identified in the no-antibody lane, indicating that our immunoprecipitation using the antibody of choice was specifically isolating MEF2. Next, we wanted to ensure that our model cell system was behaving similarly as to what has been previously reported in the literature. Differentiation was initiated in the C2C12 cells by serum withdrawal and followed along a time course up to 3 days. Whole cell lysates were probed for the expression of both MEF2 and sarcomeric actin (Fig. 2). Expression of sarcomeric actin can be used as a protein marker that is only expressed in

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>LocusLink ID</th>
<th>Percent Repeat</th>
<th>Cpg Island</th>
<th>Consensus Sequence</th>
<th>MSCAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aacs</td>
<td>78894</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acas2l</td>
<td>68738</td>
<td>29.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adcyap1</td>
<td>11516</td>
<td>11.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clcn4-2</td>
<td>12727</td>
<td>8.61</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grasp</td>
<td>56149</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map2k2</td>
<td>26396</td>
<td>19.73</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mcm4</td>
<td>17217</td>
<td>0.34</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mscp1</td>
<td>17763</td>
<td>0.23</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net12</td>
<td>54003</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nr4a2</td>
<td>18227</td>
<td>0</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pak2</td>
<td>NA</td>
<td>52.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rage</td>
<td>26448</td>
<td>27.06</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sast</td>
<td>56527</td>
<td>26.29</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sip1</td>
<td>66003</td>
<td>6.97</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slc12a9</td>
<td>83704</td>
<td>35.85</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slc13a2</td>
<td>20500</td>
<td>14.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spats2</td>
<td>20500</td>
<td>41.27</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spre2</td>
<td>114716</td>
<td>56.91</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tob1</td>
<td>22057</td>
<td>21.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xin</td>
<td>22437</td>
<td>36.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NA, not available.
fused skeletal muscle cells. Indeed, following serum withdrawal, we were first able to detect protein expression of MEF2 at 6 h of differentiation. The appearance of MEF2 correlated well with the appearance of sarcomeric actin. By 1 day, protein expression of MEF2 had reached its peak as had sarcomeric actin. Thus we were confident that we would be able to capture the appearance of MEF2 during terminal differentiation of C2C12 cells during our chosen time course.

**ChIP and PCR amplification of known MEF2 genes.** To confirm that our ChIP assay was behaving properly, we conducted ChIP on differentiating cells using primers designed around transcription factor binding sites for genes known to be regulated by MEF2 (4). From these analyses, we were able to demonstrate that the immunoprecipitated DNA did indeed contain the putative MEF2 binding domains for all of the identified genes associated with our protein of interest (Fig. 3). Together, these findings demonstrate that the MEF2 antibody used in this study was able to immunoprecipitate MEF2-bound DNA specifically.

**Identification of MEF2 regulated genes using ChIP on CpG microarrays.** To conduct a high-throughput analysis of the genes associated with MEF2, we turned to ChIP on CpG island microarray assays. In the past, others have successfully employed these intergenic regions as bait to identify genes associated with specific transcription factors (25, 44, 45). Unfortunately, at the time of this study no mouse array was available. Therefore, we obtained the mouse CpG library generated by Sally Cross and Adriane Bird (9) and currently being housed at the MRC RFCGR. From this library we randomly chose 7,000 clones, which we arrayed onto microarrays. Immunoprecipitated DNA collected from differentiating myoblasts at specific time points was then hybridized to the mouse CpG arrays. The pattern of MEF2 binding at each time point was then determined (Fig. 4). From this analysis, a large number of positive clones were identified at each time point (0 h, 9 clones; 6 h, 792 clones; 1 day, 163 clones; 3 days, 316 clones). Because of the large numbers that would require sequencing, we felt that as a starting point we would begin by defining a subset of these genes. However, to avoid bias on our part, we chose to define those clones that were only found to associate with MEF2 for at least two of the time points.

With the above limitation placed on our search we obtained a total of 291 significant clones. These were sequenced (see Supplemental Table S1, available at the Physiological Genomics web site) and initially screened for overlapping redundancy in the sequence. This reduced the total of remaining clones to 72. We were not surprised by the large number of redundant sequences as the clones were randomly chosen from a library that was not normalized. Of the nonredundant clones, only those containing less than 66% repetitive elements were then considered, reducing the final number of clones to 53. These were then screened against the UCSC assembly to place them on the mouse genome. Of the 53 sequences, 50 were able to be aligned to the genome. Recent evidence demonstrates that regulatory sites may reside 5′, 3′, or within the gene itself (7). With this in mind, to conduct our gene search we took the approach of identifying the first gene(s) located within 20 kb in either direction of each positive clone sequence. From this search, 35 of 50 (70%) of the positive clones had a gene associated with it. Of the putative genes, 20 had some form of annotation associated with them. In addition, when we searched for predicted genes we were able to identify 16 putative genes. Included in this list were a number of RIKEN clones. This group of full-length cDNAs raises the intriguing possibility for novel gene discovery, which we are currently pursuing.

---

1The Supplemental Material for this article (Supplemental Table S1) is available online at [http://physiolgenomics.physiology.org/cgi/content/full/00149.2004/DC1.](http://physiolgenomics.physiology.org/cgi/content/full/00149.2004/DC1)
The temporal binding patterns of MEF2 to the 20 annotated genes were plotted (Fig. 5), from which two distinct groups could be identified. The first group of nine genes had elevated MEF2 binding at 6 h of differentiation, then returned to levels observed at 0 h (Fig. 5A). The second group, consisting of 11 genes, tended to have elevated binding at 6 h that remained high at 1 day but then returned to levels observed at 0 h by 3 days (Fig. 5B). With the exception of two clones, the binding of MEF2 to the regulatory elements of these genes only occurred following the initiation of differentiation. Taken together, these findings demonstrate the dynamic nature of MEF2 binding. Furthermore, when compared with the quantities of MEF2 protein (Fig. 2) in the cells, binding appears to be independent of the quantities of protein available.

From the list of annotated genes (Table 2), a number of functional groups arose. The first group is associated with signaling. This consisted of a number of genes that included Map2k2, Mtcp1, RAGE, SPRED2, PAK2, GRASP, and Tob1. Map2k2, RAGE, PAK2, and SPRED2 are all involved in the MAPK signaling cascade (29, 42). SPRED2 has been shown to inhibit the activation of Raf through its association with and inhibition of ras. By inhibiting the MAPK pathway the mitogenic signaling may be inhibited. Furthermore, two of the genes Tob1 and RAGE may also play a role in the movement of the cells out of mitotic division. Both genes are involved in cell cycle regulation. Tob1 encodes a receptor type tyrosine kinase that interacts with ErbB2 (40). The likely result of this interaction would be the inhibition of cell cycle progression. Temporally, the binding patterns of MEF2 to these genes demonstrated elevated levels by 6 h, then dropped by 1 day back to levels observed at 0 h (Fig. 5A). This pattern was broken by the Map2k2 and PAK2 genes, which demonstrated sustained binding to MEF2 at 1 day of differentiation (Fig. 5B).

Map2k2, also known as Mek2, is an integral member of the MAPK signaling cascade, whereas PAK2 has been associated with activating Raf in this same pathway (20). Elevations in either of these proteins would potentially translate into the activation of a mitotic signal.

The genes Mtcp1 and GRASP are associated with the phosphatidylinositol 3-kinase (PI3-kinase) cascade (22, 32). Mtcp1 (mature T-cell protein 1) was originally identified in T-cells and belongs to a family of binding proteins that have been shown to associate with PKB (22). GRASP encodes a scaffold protein for the phosphoinositide receptor and may play a role in regulating this cascade. It has been well-established that the PI3-kinase signaling pathway plays a key role in the regulation of muscle differentiation (11, 16, 23). Thus Mtcp1 along with GRASP may contribute in promoting the cells to commit to becoming terminally differentiated muscle.

A second functional group consisting of five genes could be separated and were related to neuronal development. Included in this list were Sip1, Nr4a2 (NURR1), Nell, and Sast. The three genes Sip1, Nr4a2 (NURR1), and Nell have all been associated with neurogenesis (15, 21, 39). All three of these genes were found to be associated with MEF2 at elevated levels between 6 h and 1 day (Fig. 5B). Sast encodes a syntrophin associated serine/threonine kinase protein that is located at the neuromuscular junction (24).

The remaining annotated genes identified were related to subgroups that could be associated with the development of a specialized cell type. For example, a third group of genes were found to be involved in ion transport or receptor function. These included Clcn4–2 (18), Slc12a9, and Slc13a2 (12). Others such as Aacs and Acas encode metabolic proteins, a process vital for normal cellular function. Both the neural genes and those representing ion channels point toward the development of an excitable tissue such as skeletal muscle.

Confirmation of positive microarray targets. To reconfirm our findings, we had chosen 10 of our positive clones as targets and conducted ChIP with primers specific to the putative MEF2 binding sites located upstream of each gene. To select potential candidates, each positive clone was required to meet two criteria. First, we searched for the presence of predicted CpG islands within 2 kb of the clone sequence. The location of CpG islands were predicted by searching the genomic sequence for high GC content (roughly ≥50%) over a length of nucleotides greater than 200 bp. Furthermore, nucleotides were scored for the proportion of CG content and had to exceed a ratio greater than 0.6 over the length of the segment. From the list of positive clones located near either a predicted or annotated gene target, 37% demonstrated the presence of a predicted CpG island. However, upon closer examination, those sequences that didn’t demonstrate a putative CpG island were in fact associated with stretches of high GC content. Thus CpG islands may have in fact been present but may not have qualified as such by the UCSC search criteria. The second component of our search required the existence of a putative MEF2 binding site within 2 kb of the target genes. Searching
for these binding sequences is troublesome and difficult, often leading to numerous putative sites or no sites at all. To identify putative MEF2 binding sites, we applied a pattern recognition program and found that 35% of all the target genes demonstrated the presence of a MEF2 binding site. Thus, of the 10 clones, we were only able to reconfirm 7 following PCR amplification (Fig. 6). For two of the clones we were unable to obtain clean PCR products based on our predictions of MEF2 binding regions, and the PCR reaction for the third clone (Clcn4–2) demonstrated no specificity between antibody conditions.

To reconfirm the specificity of the method, we identified five clones that did not demonstrate significant changes in expression over the period of differentiation. These were sequenced, and again we looked for the presence of CpG islands and MEF2 binding sites. Four of the clones demonstrated the presence of a CpG island, and two demonstrated a putative MEF2 binding site. We felt that these two clones could be discounted, as one of the clones did not have a CpG island associated with it nor was it CG rich. The second negative clone did have a predicted CpG island; however, the MEF2 binding site located was ~2,000 bp away from the site of the clone sequence. As our clones are on average 580 bp in length, this site is unlikely to be detected. Taken together, these findings further confirmed the specificity of our immunoprecipitation method.

DISCUSSION

Recently, there has been a great deal of focus placed upon understanding the mechanisms regulating muscle differentiation. This has encompassed aspects ranging from signaling pathways to gene expression profiles. Indeed, these studies have led to many insights but have made it clear that the process of muscle differentiation is a complex and involved process. This becomes especially relevant when considering the role that each of the muscle transcription factors plays. Knockout studies have illustrated the complex interplay between the various muscle transcription factors (5, 33, 36). Although highly informative, these studies lack the acuity to elucidate the direct role that each has on regulating the pattern of gene expression. Thus it was the goal of the present study to
characterize the genes directly associated with a key muscle transcription factor.

Our study provides an initial glimpse into understanding the role that MEF2 contributes to regulating gene expression. To accomplish this, we decided to employ ChIP methods coupled with the high-throughput advantages of microarrays. In this way we were able to define the unique binding signature of a specific transcription factor during muscle differentiation. As will be discussed, the genes identified in the present study were not found to be associated with classic muscle proteins. In fact, we were able to identify 20 genes that, to the best of our knowledge, have not been previously associated with MEF2 regulation. This is the first time that potential gene targets of MEF2 have been isolated in such a fashion.

From our list of clones, we first wished to see if there was a tendency for MEF2 to associate with a particular group of genes. We found that nine were defined as having either a signaling or kinase function of which five are involved with neuromuscular junctions or neural development. This was surprising to us as we expected to find associations with genes encoding classic muscle components, e.g., contractile proteins. However, the selection of clones was random, and the potential exists that we simply had not picked a clone that was located near a gene encoding a contractile protein. Furthermore, we have not sequenced all of the clones on the array. This makes it uncertain whether clones exist on the array that are located near muscle-specific genes. Regardless, our findings define a novel set of genes such as those encoding proteins associated with neurons. MEF2 has been associated with neural tissue and has even been implicated in the development of this tissue (26).

As both neural and muscular tissues share the characteristic of having an excitatory membrane, the potential exists that these genes may also play a similar developmental role in muscle. The presence of genes associated with membrane receptors or ion transporters (Clcn4–2, Slc12a9, and Slc13a2) also suggests a role in establishing an excitatory membrane.

The genes involved in signaling tended to be associated with the MAPK pathway. The functional annotation for some of these genes suggests contradictions exist. For example, the ras associating protein, SPRED2 has been found to inhibit raf activation (42) while PAK2 was shown to activate raf (20). Although this appears to be counterintuitive, the association of MEF2 does not always necessarily correlate to elevations in gene expression. Precedence exists for this, as it has been shown that muscle gene expression is dependent upon MEF2 association with bHLH proteins (reviewed in Ref. 30). Furthermore, MEF2 has been demonstrated to associate with proteins (MITR, HDAC4, and Cabin) that confer repression (28, 46, 49). Thus regulatory elements such as described above may be present at specific sites and times imposing control over MEF2-associated transcription. Our data support this hypothesis, as both Map2k2 and PAK2 are found in a group of genes that are associated with MEF2 for a prolonged period of time (Fig. 5B). This suggests that MEF2 may be acting in an inhibitory role to reduce the expression of these genes during a period in which promitotic signals are not favorable for muscle differentiation to occur. Further analysis will be required to determine whether the expressions of the genes we have identified are altered, e.g., quantitative PCR or Northern blots. Recent evidence suggests that a correlation may in fact exist between steady-state mRNA expression and transcription factor binding. Bergstrom et al. (3) found a high degree of correlation on a subset of data collected from cDNA microarrays, Northern blots, and ChIP analyses. However, a degree of caution is necessary here as the results of this study utilized an expression system retrovirally transfected into MyoD null fibroblasts. Thus the stoichiometry would be heavily in favor of this transcription factor.

To define a biological process such as muscle differentiation, we felt it necessary to utilize a time course model. From previous microarray studies it was clear that gene expression patterns were dynamic. Groups of genes could be found to increase and decrease in a time-dependent manner. The question arises as to what was regulating these changes. Our findings demonstrate that the association of MEF2 with the regulatory sites does not follow a singular on/off pattern (see Fig. 5). Clearly, gene regulation by MEF2 is far more complex than once thought. Of interest, the greatest number of positive clones was found to be elevated in MEF2 binding following 6 h of differentiation. Prior to 6 h, we had only found nine clones to be elevated in MEF2 binding. On closer inspection of these clones, only five were able to be sequenced, and of these, three contained repetitive elements. The remaining two clones were found to be associated with Nr4a2 and a RIKEN clone. Although we were unable to detect MEF2 protein expression at 0 h, we cannot exclude the possibility that a form of MEF2 (e.g., the B isoform) may be involved. Alternatively, the levels of MEF2 may have been too low for us to detect using the
antibody we had chosen. Overall, the number of clones that may be attributed to MEF2 binding is very low, suggesting a limited role for MEF2 at this time point. We were able to detect MEF2 protein expression by the 6 h time point (Fig. 2), which correlated well with elevations in MEF2 binding. However, the greatest protein levels for MEF2 were found to be at the subsequent 1 and 3 day time points. These results suggest that gene regulation by MEF2 is not directly related to its protein level within the cell. Rather, the data define a regulatory network for gene expression that does not follow a linear path in which genes are only activated as the transcription factors become available. Indeed, the MEF2 family of transcription factors have been shown to work in conjunction with other regulatory elements, e.g., MyoD (see Ref. 48 for review). Future work will center on the characterization of genes associated with other muscle transcription factors. Overlap between these various factors could then be searched for, suggesting possible coordinated regulation.

In summary, our findings provide a unique glimpse into the regulatory network associated with transcription factors during a physiological event. We were able to establish a unique cohort of genes not previously identified to be associated with MEF2. From these findings, new insights have arisen into the role that regulatory elements such as MEF2 play in determining cellular fate and outcome. Although the genes identified are limited to those genes related to CpG islands, we believe this methodology will prove extremely valuable as a means of defining the primary associations of genes with transcription factors.

ACKNOWLEDGMENTS

We thank Quyen Tran and Tuyet Diep for assistance in the production of the CpG microarrays and sequencing of the clones.

GRANTS

This work was supported by grants from Genome Canada and the Ontario Research and Development Challenge Fund.

REFERENCES


