research articles Journal of proteome research

Broad Profiling of DNA-Binding Transcription Factor Activities Improves Regulatory Network Construction in Adult Mouse Tissues

Yi-Min Sun,[†,‡,#,§] Yan Zhang,[†,‡,§] Ling-Qin Zeng,[†,‡] Jian-Ping Wu,[‡,#] Li Wei,[‡,#] Ai-Hui Ren,[‡,#] Wei Shao,[†,⊥] Ji-Ying Qiao,[†,‡] Yong-Chao Zhao,[†,‡] Liang Zhang,[‡,#] Keith R. Mitchelson,[†,‡,#] and Jing Cheng*[,†,‡,||]

Medical Systems Biology Research Center, Tsinghua University School of Medicine, Beijing 100084, China, National Engineering Research Center for Beijing Biochip Technology, 18 Life Science Parkway, Beijing 102206, China, CapitalBio Corporation, 18 Life Science Parkway, Beijing 102206, China, The State Key Laboratory of Biomembrane and Membrane Biotechnology, Tsinghua University, Beijing 100084, China

Received June 7, 2008

Molecular systematics involves the description of the regulatory networks formed by the interconnections between active transcription factors and their target expressed genes. Here, we have determined the activities of 200 different transcription factors in six mouse tissues using an advanced mouse oligonucleotide array-based transcription factor assay (MOUSE OATFA). The transcription factor signatures from MOUSE OATFA were combined with public mRNA expression profiles to construct experimental transcriptional regulatory networks in each tissue. SRF-centered regulatory networks constructed for lung and skeletal muscle with OATFA data were confirmed by ChIP assays, and revealed examples of novel networks of expressed genes coregulated by sets of transcription factors. The combination of MOUSE OATFA with bioinformatics analysis of expressed genes provides a new paradigm for the comprehensive prediction of the transcriptional systems and their regulatory pathways in mouse.

Keywords: Microarray • profiling • regulatory network • target gene • transcription factor

Introduction

Studies in mouse tissues have confirmed that transcriptional regulation of genes of related function is highly regulated and that coordinated transcription occurs in genes related to particular tissue-specific functions.[1,2] Transcription of an individual gene is regulated via the binding of activated transcription factor proteins (TFs) to particular *cis*-regulatory DNA elements of the gene promoter and/or to other *cis* gene enhancers. Experimental studies of muscle and liver specific gene transcription suggested that TFs act via regulatory TF modules in which clusters of TF binding sites bring bound TFs into close (physical) contact to confer tissue specificity.[1,3] The modular concept proposes that several coincident bound TFs (and other cofactors) interact together to effect activation of the particular gene. Thus, individual TFs may contribute a combinatorial control to the activation of numerous different genes. Individual TF species may be present in several tissues, and may activate different target genes via interactions with particular gene regulatory modules and by interactions with other TFs with appropriate tissue distributions to define tissue specificity. The temporal regulation of particular processes may further define the roles of TFs.

Current experimental knowledge of tissue-specific TF activities is limited, as the large numbers of target genes and potentially interacting TFs make the task of experimental identification of the activity of any particular TF in any tissue, or in different cell types of a tissue, both technically difficult and extremely time-consuming. In addition, as some TFs may interact at different affinities with sequences varying from the canonical, precise knowledge of the DNA binding specificities of each TFs and the definition of functional enhancer elements at gene promoter regions is also limiting and difficult to confirm. Perhaps contradictorily, this dearth of experimental evidence has stimulated numerous bioinformatics-based approaches to the prediction of tissue-specific TF interactions, inferred from readily measurable differential global gene expression data[4−7] and from likely TF binding sequences present in gene promoter and regulatory regions. Most predictive algorithms have considered that the TFs act in concert and that the functional interactions between the TFs require the close proximity of (neighboring) motifs in genes regulatory regions. For example, the analyses of Yu and colleagues[5,8] and Zhu et al.[9] predict the higher probability of the tissue-specific regulation of numerous genes by particular TFs by involving

* Corresponding author: Dr. Jing Cheng, Medical Systems Biology Research Center, Tsinghua University School of Medicine, Beijing 100084, China. E-mail: jcheng@tsinghua.edu.cn, Telephone Number: +86-10-6277 2239, Fax Number: +86-10-6277 3059.

† Medical Systems Biology Research Center, Tsinghua University School of Medicine.

‡ National Engineering Research Center for Beijing Biochip Technology.

CapitalBio Corporation.

§ These authors contributed equally to this study.

⊥ Current address: Department of Neuroscience and High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

|| The State Key Laboratory of Biomembrane and Membrane Biotechnology, Tsinghua University.

10.1021/pr800417e CCC: $40.75

Published on Web 08/30/2008

both co-occurrence and proximity constraints. With the analysis of bioinformatics data statistically, probability matrices could be constructed for key TFs involved in regulatory hubs and for cellular pathways functioning as a result of these putative TF activities. Although well-established tissue-specific gene expression profiles and verified active TFs were used to train and confirm the modeling algorithms, many of the TFs predicted by such analyses to play roles in particular tissues have not been confirmed experimentally.

We explored methods to provide quantitative assessment of protein−DNA binding specificities that permit some relative measure of TF protein activities in a preliminary proof-of-principle work[10] by an oligonucleotide array-based transcription factor assay (OATFA). We have since made a series of improvements to the original OATFA methods, which we recently reported on platforms capable of simultaneously examining the activities of 93 yeast TFs (YEAST OATFA)[11] and 240 human TFs.[12] Here, we present a MOUSE OATFA platform employing the analysis methods described above which assays some 200 TF activities simultaneously, representing about 10% of the predicted murine TFs. We report a detailed examination of the activities of TFs among several adult mouse tissues. As expected, tissue-specific transcription factor activities could be found in each adult tissue. We also contrasted these experimentally determined TF activities with several bioinformatics-based predictions of organ TF signatures inferred from detailed analysis of published genome-wide gene expression data by examination of differentially expressed gene transcription regulation-targets of the TFs.[4−7] The experimental MOUSE OATFA set identified a majority of the bioinformatically predicted TFs activities in particular tissues, but importantly, many TF activities that were not predicted were also found. Some TFs found active by MOUSE OATFA were also not discernible by examination of the expression of TF-encoding genes in the tissues and, hence, were not identified by direct expression-based prediction analysis. The study illustrated the benefit of MOUSE OATFA analysis, in addition to other global molecular analyses, to yield insight into the molecular interactions that regulate tissue function. Detailed analysis of the pathways associated with the activities of the transcription factor SRF in mouse lung and skeletal muscle revealed its roles in three major pathways involving different genes and different co-interacting transcription factors in each tissue. The MOUSE OATFA analysis also experimentally confirmed the coregulatory roles of key TFs in different tissues.

Materials and Methods

Preparation of Nuclear Extracts of Adult Mouse Tissues. Adult (10−12-week-old) mouse liver, kidney, lung, brain, heart and skeletal muscle were dissected from six male ICR mice and snap-frozen in liquid nitrogen and stored at −80 °C until use. Nuclear extracts were prepared using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Pierce, Rockford, IL) according to the manufacturer's instructions. Protein concentrations were determined using BCA Protein Assay Kit (Pierce). The final nuclear extracts were aliquoted and stored at −80 °C until use.

Probe Sets. The probe design and screening were performed as previously described.[12] All probes were designed based on TRANSFAC Professional r8.2 (http://www.biobase.de), a database of eukaryotic transcription factors, their genomic binding sites and DNA-binding profiles.[13,14] All oligonucleotides were synthesized by the Sangon Corporation (Shanghai, China). Four probe sets for the control of TF-DNA binding (negative control for binding, BNC) were designed to ensure the reliability of microarray analysis. The signal intensities of the BNCs can reveal the levels of nonspecific binding in an individual MOUSE OATFA microarray assay. Four amino-group modified probes (probe 1) for control of hybridization (negative control for hybridization, HNC) were also designed with low similarity to all other PSSMs. The "Total TF-probe library" mixture comprised the 240 duplex TF-capture probes and 4 BNC probes. The sequences of all the 240 probes are shown in Table S1 in Supporting Information.

Microarray Fabrication. The 5′-amine modified probe 1 oligonucleotides were resuspended in distilled water and then diluted with Spotting Solution (CapitalBio, Beijing, China) to a final 20 μM concentration. Each oligo-probe was printed in duplicate on aldehyde-modified glass slides by a SmartArrayer-48 Microarray Spotter (both CapitalBio) under 45% humidity in 22 rows × 24 columns spot configuration of each array. The spot diameter was 120 μm, and the distance from center to center was 200 μm. The printed MOUSE OATFA microarrays were covalently coupled overnight under 70% humidity. After a wash using the 0.2% SDS solution, the microarrays were blocked in the 0.2% $NaBH_4$ buffer for 5 min. The microarrays were then ready for sample hybridization after additional washing twice for 5 min each time using distilled water.

The MOUSE OATFA Assay Method. (1) Binding Reaction. The binding reaction occurred at room temperature for 60 min in a 20 μL binding mixture containing 1× binding buffer, purified TFs or cell nuclear extracts. The TF-capture DNA library was added to the binding mixture which was preincubated at 4 °C for 10 min to reduce possible nonspecific binding. The final concentration of each TF-capture probe was 0.1 nM. As commonly used in the conventional EMSA, specific TF capture-probe interactions were enhanced by the addition of poly dIdC, which reduces nonspecific competition by other proteins that can compete with TFs for binding to their specific TF-binding sequences. The 20 μL of reaction products would be used in the following agarose gel electrophoresis separation.

(2) Separation. The binding mixture with 2 μL of loading buffer (0.25× TBE, 60%; glycerol, 40%) was loaded into separate lanes on 1.5% agarose (Sigma, St. Louis, MO) gel and run at 12 V/cm in chilled 0.5× TBE for 30 min. The gel area that contains the protein−DNA complexes was excised and transferred into a 1.5 mL tube. The DNA was extracted using QIAEX II Gel Extraction Kit (Qiagen, The Netherlands).

(3) Single Primer Amplification (SPA). The recovered DNA, 200 nM Cy5 or Cy3-labeled T7 promoter primer (5′-GGG GTA ATA CGA CTC ACT ATA GGG-3′), 200 μM dNTPs (TaKaRa, Dalian, China) and 2.5 units of Taq polymerase (TaKaRa) were mixed into a 20 μL mixture and amplified for 30 cycles. The SPA products were heated to 45 °C and vacuumed dry.

(4) Hybridization. Two corresponding SPA products with Cy3- or Cy5-label were mixed and hybridized to the MOUSE OATFA microarrays in 3× SSC (450 mM sodium chloride/45 mM sodium citrate, pH 7.0), 0.2% SDS, 5× Denhart's and 25% formamide in a total 12-μL mixture at 42 °C for 16−20 h. The arrays were washed sequentially at 42 °C in 2× SSC plus 0.5% SDS for 10 min and then 0.2× SSC plus 0.1% SDS for 10 min. Immediately after washing, the array slides were spun dry at 1500 rpm for 2 min.

(5) Scanning and Data Analysis. Array images were acquired by a laser confocal scanner LuxScan−10KA (CapitalBio). The signal intensities for each spot were calculated by subtracting local background using LuxScan 3.0 software (CapitalBio).

For each microarray, linear normalization using a per-channel 50th percentile method was adopted to get the ratio for each spot. Data from independent replicate experiments were combined for statistical analysis using Student's *t* test. An average ratio was calculated for each DNA probe on the array from at least three replicates. The DNA-binding activity of a TF was deemed significant if the median ratio was >1.5 and *P* value was <0.05. A hierarchical clustering using average linkage algorithm was performed with the output of dendrograms which were generated by Treeview.

Microwell Colorimetric TF-ELISA Assays. Microwell colorimetric TF-ELISA assays were performed as previously described.[10] Antibodies for AP1 (#sc-1694x), CDP (#sc-6327), C/EBPalpha (#sc-61), FOXO4 (#sc-34903), HNF1 (#sc-8986), HNF4 (#sc-8987), MEF2 (#sc-313x), NF1 (#sc-5567), NF-kappaB (#sc-372x), SRF (#sc-335) and STAT1 (#sc-346) were purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). The binding sequences for the above transcription factors were listed in Table S4 in Supporting Information.

Western Blotting of C/EBPalpha and SRF Proteins. Nuclear extracts were fractionated by 10% denatured SDS-polyacrylamide gel and transblotted onto PVDF membranes using standard techniques. Following a blocking step with 5% nonfat milk for 2 h at room temperature, membranes were incubated with C/EBPalpha and SRF polyclonal antibody (Santa Cruz) at 1:1000 dilution overnight at 4 °C. Membranes were washed three times in 0.5% TBST and incubated with a 1:2000 dilution of horseradish peroxidase coupled secondary antibody. The immunocomplexes were detected by using SuperSignal chemiluminescence reagent (Pierce).

Chromatin Immunoprecipitation. The ChIP experiments were performed as previously described[15] with minor modifications. Tissues were homogenized in prechilled glass Dounce homogenizer with 2 mL ice-cold PBS. After a washing step, the resuspended cells were cross-linked with 1% formaldehyde for 10 min at room temperature, with occasional swirling. Glycine was added to a final concentration of 0.2 M and the incubation was continued for an additional 5 min. Cells were collected and sequentially washed with ice-cold PBS twice, cell lysis buffer (10 mM Tris-HCl [pH 7.5], 10 mM NaCl, 3 mM $MgCl_2$, 0.5% IGEPAL, 1 mM PMSF) three times, and resuspended in 2.5 mL SDS lysis buffer (10 mM Tris-HCl [pH 7.5], 150 mM NaCl, 3 mM $MgCl_2$, 1 mM $CaCl_2$, 1% SDS, 1 mM PMSF). Cells were then sonicated on ice followed by centrifugation for 10 min at 15 000*g*. Supernatants were collected and diluted in IP dilution buffer (20 mM Tris-HCl [pH 8.0], 2 mM EDTA, 1% Triton X-100, 150 mM NaCl, protease inhibitors). Chromatin immunoprecipitation was performed for 6 h or overnight at 4 °C with specific antibodies followed by the addition of protein A (or G) magnetic beads for an additional 2 h. The beads were washed sequentially for 10 min in wash buffer I (20 mM Tris-HCl [pH 8.0], 2 mM EDTA, 1% Triton X-100, 150 mM NaCl, 1 mM PMSF), wash buffer II (20 mM Tris-HCl [pH 8.0], 2 mM EDTA, 1% Triton X-100, 0.1% SDS, 500 mM NaCl, 1 mM PMSF) and wash buffer III (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 0.25 M LiCl, 0.5% IGEPAL, 0.5% deoxycholate). Precipitates were then washed three times with TE buffer and extracted twice with elution buffer (25 mM Tris-HCl [pH 7.5], 10 mM EDTA, 0.5% SDS) by heating for 30 min at 65 °C. To reverse the cross-links, samples were incubated with 1.0 mg/mL proteinase K at 65 °C for at least 6 h. DNA fragments were purified for further PCR analysis. The primer sequences for PCR analysis of the regulated genes are listed in Table S7 in Supporting Information.

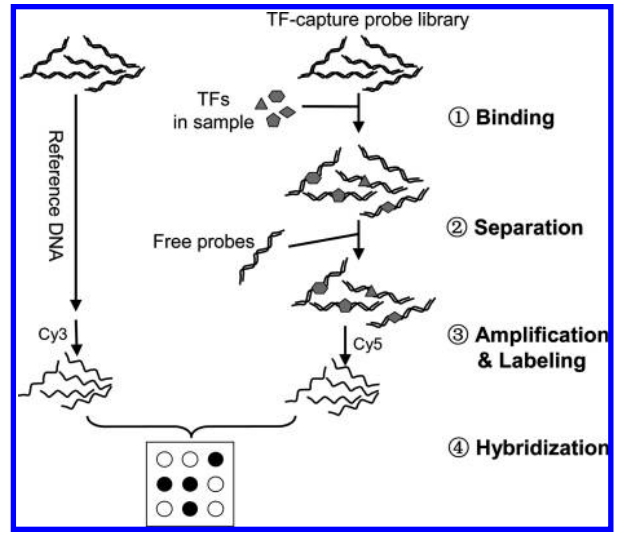

Figure 1. The MOUSE OATFA single sample assay (SSA) procedure. The TF capture DNA library was directly amplified and labeled as the reference DNA to be hybridized with labeled sample DNA.

Results

Analysis of TF Activity Using MOUSE OATFA. The experiments described here employed a single sample assay (SSA) format (Figure 1) whereby TF signals are corrected for residual unbound capture probe signals. Use of the MOUSE OATFA platform involves four basic steps: (1) Binding, a set of duplex DNA binding oligonucleotides (TF-probe mixture) were incubated with a nuclear extract to allow the formation of DNA/protein complexes; (2) Separation, the DNA/protein complexes were separated from any residual unbound DNA probes by agarose gel electrophoresis; (3) Amplification and labeling, the DNA probes were then extracted from the TF protein complexes and the DNA quantitatively PCR amplified using a 5′-fluoro-tagged single strand primer; (4) Hybridization, finally probes were hybridized to the profiling microarray to provide a semiquantitative estimation of relative TF activities via the estimation of the amount of bound probe DNA.

Among the mammalian TRANSFAC data, each transcription factor typically has a number of different binding matrices, and each matrix has interactions with a number of different TFs. Thus, MOUSE OATFA should be considered to be a high-throughput technology to guide the rapid identification of the most likely TF activities in a tissue or cell. To simplify these binding interactions, the 240 binding matrices can be classified into 72 subfamilies (Supplementary Table S2 in Supporting Information). The validation of performance of MOUSE OATFA platform was described in detail in the Supporting Information under **"Performance of MOUSE OATFA platform"**.

MOUSE OATFA Analysis of TF Activity in Six Adult Mouse Tissues. We undertook the comparative analysis of the intrinsic activities of DNA-binding TFs across a panel of six normal adult mouse tissues (liver, kidney, lung, brain, heart and skeletal muscle) to identify tissue-specific active TF profiles which are shown in Figure 2A. Several observations can be made from these experiments. First, when hierarchical clustering was used, the same types of tissue from different individuals clustered together. It was reasonable that heart and (skeletal) muscle also clustered together because of the similarity of many of their biological processes and functions. The cluster analysis

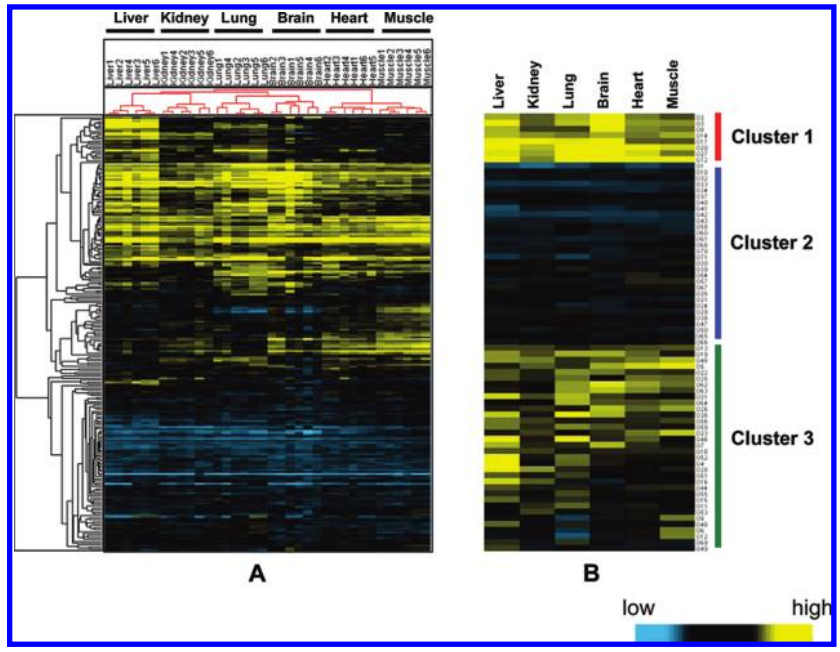

Figure 2. The hierarchical clustering of the activities of 200 different TFs in six mouse tissues. (A) The entire data set comprising 6 replicates of each tissue. (B) Selected representative MOUSE OATFA signals of subsets of TFs. Cluster 1: The TFs were active in each tissue. Cluster 2: The TFs were not detected in each of the tissues (below detection limits). Cluster 3: The TFs were active in some of the six tissues.

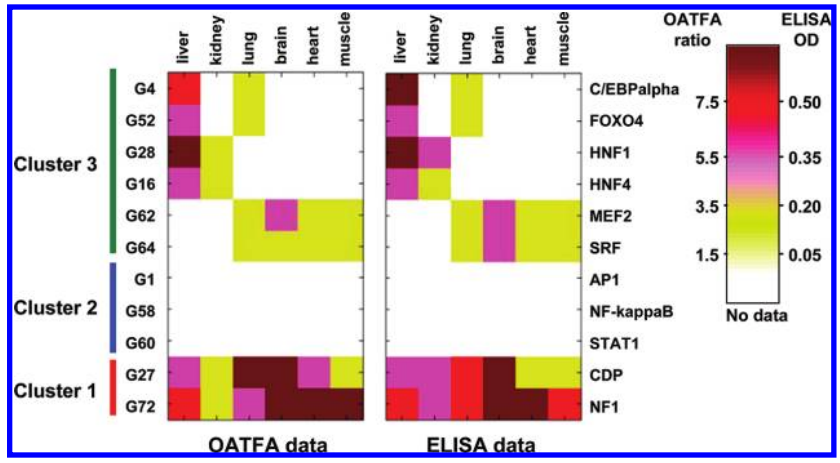

Figure 3. Comparison of MOUSE OATFA signals of 11 TFs with their ELISA estimations in the tissues. Representative subsets of TFs were indicated (left side). Cluster 1: The TFs were active in each tissue. Cluster 2: The TF activities were not detected in each of the tissues (below detection limits). Cluster 3: The TFs were active in some of the six tissues. The name code for individual TFs are indicated (right side).

also indicated a close functional similarity between liver and kidney. Both findings are similar to the conclusions of a cluster analysis of global mRNA expression of various mouse tissues.[16] Second, the 72 groups of TFs could be divided into three categories based on their activities in these 6 tissues (Figure 2B and Table S3 in Supporting Information), being either active in each of the tissues (Cluster 1), not active in any of the examined tissues (Cluster 2) or active in some tissues (Cluster 3). Cluster 1 includes such TF proteins as the ubiquitously expressed SP1 (G20) zinc finger protein and c-Myc (G8); Cluster 2, comprising TFs which were found not to be active in any of the adult tissues studied here, included regulatory transcription factors such as glucocorticoid receptor (G10) and members of the STAT family (G60), which are typically kept inactivated or are only weakly activated and below the detection limit of our methods. Tissue-specific TF activity (Cluster 3) is an area of intense interest and has been the subject of numerous bioinformatics-based approaches for the prediction of tissue-specific TF interactions[4−7] inferred from readily measurable tissue-specific gene expression data of prospective TF targets and the TF binding sequences present in their gene promoter and regulatory regions.

We selected a number of tissue-specific TF activities for independent determination by antibody-based TF-ELISA analysis (Figure 3). We examined 11 transcription factors in the six different tissues, representing two TFs from category 1 (NF1, CDP), three TFs from category 2 (STAT1, NF-kappaB, AP1) and six TFs from category 3 (SRF, MEF2, HNF4, HNF1, FOXO4, C/EBPalpha). The semiquantitative estimates of the activity of each of the 11 TFs determined by MOUSE OATFA agreed closely with their quantitative TF-ELISA analysis. This data also demonstrated a number of novel TF activities that have not been previously reported, one example being FOXO4 (also called AFX), which is a member of the forkhead FOXO

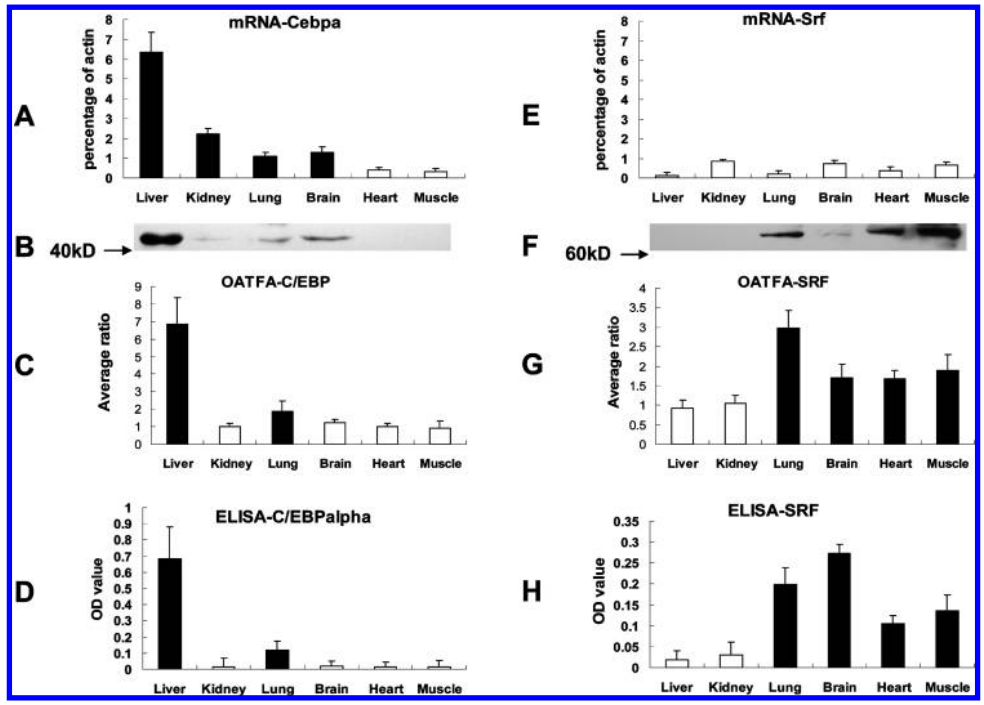

Figure 4. The amounts of mRNA and protein and the relative activities of C/EBPalpha and SRF. Left panels: C/EBPalpha was strongly expressed (mRNA, GNF SymAtlas) in liver, but markedly less so in other tissues, particularly in heart and muscle. C/EBPalpha activity was strongest in liver, then lung, and was below cutoff in the other tissues. C/EBPalpha protein levels were highest in liver, but were detectable in brain, lung, and least in kidney. Right panels: levels of SRF mRNA expression (GNF SymAtlas) were virtually undetectable in each of the tissues, but SRF activity was highest in lung, then in heart, muscle and brain. SRF protein levels were highest in muscle and heart and followed by lung, but were just detectable in brain. The estimations of TF activity by MOUSE OATFA and ELISA agreed closely, but determination of the TF protein levels or the expressed TF gene mRNA levels may be at variance in individual tissues.

transcription factor family, and plays important roles in cell cycle progression, apoptosis, oxidative stress and DNA repair.[17] Despite the absence of direct experimental data, Yang and colleagues predicted that FOXO4 could participate in regulation of eye- and liver-specific gene expression by interaction with PBX1A and FOXO1A, respectively. From our MOUSE OATFA experiments, FOXO4 showed highest activity in normal mouse liver and next highest in lung, both of which were confirmed by ELISA assays. HNF1 and HNF4, members of the hepatocyte nuclear factor family, were reported to regulate tissue-specific gene expression in liver and kidney.[5] We also confirmed the tissue-specific signatures of these TFs by both MOUSE OATFA and ELISA assays (Figure 3).

A further advantage of MOUSE OATFA can be seen by contrasting it to technologies used to study the abundance of proteins in cells, such as two-dimensional electrophoresis and protein microarrays.[18] This can be illustrated by considering the CCAAT enhancer binding protein alpha (C/EBPalpha), a member of basic leucine zipper (bZIP) transcription factor family which regulates the expression of numerous liver-specific genes[19] and processes during lung maturation.[20] The MOUSE OATFA analysis determined that the highest activity of C/EBPalpha occurred in liver followed by lung, while the other four tissues showed no detectable activity, which were quantitatively confirmed by independent ELISA analysis (Figure 3). These quantitative estimations of C/EBPalpha activities determined by ELISA analysis and MOUSE OATFA in the six tissues are also contrasted with two other types of quantitative data—expressed mRNA gene data and Western blot determination of protein (Figure 4). The ELISA and OATFA estimations of C/EBPalpha activities agreed closely (Figure 4D,C), and also agreed with expressed mRNA levels (Figure 4A); however, Western blotting analysis (Figure 4B) indicated that C/EBPalpha protein was present at greater levels in brain than in lung and absent (or below detection limits) from muscle and heart, in contrast to the other assay methods. These findings illustrate that the estimation of protein by Western blotting does not necessarily reflect active TFs, rather it measures total protein regardless of activity.

The activities of SRF in each tissue by ELISA analysis and MOUSE OATFA (Figure 4H,G) were also contrasted with expressed mRNA gene data (Figure 4E) and Western blot determination of protein (Figure 4F). Although the expressed mRNA for SRF was almost undetectable in liver, kidney and lung, SRF protein level was readily detectible by Western blot in muscle (highest), heart and lung (least), and was just detectible in brain. In contrast, the ELISA determination of SRF activity (which correlated well with MOUSE OATFA activities, see Figure 3) showed the highest TF activity in brain, followed by equivalent high levels of activity in lung, heart and muscle. These contrasting assays, which are used frequently to indicate potential TF activity, illustrate well that TF activities are neither necessarily reflected in the absolute protein levels nor necessarily in the expressed mRNA levels.

MOUSE OATFA array data agreed closely with the known biological data concerning each tissue, across the full range of the surveyed TFs, with high levels of agreement with reported tissue-specific TF activities and good agreement with the predicted tissue specific TF activities. Details of the comparisons between the OATFA results and the reported or predicted tissue-specific TF activities are given in the Supporting Information, under "Correlation of OATFA result with published tissue-specific signatures". Notably, although the current MOUSE OATFA platform only measures the activity of some 10% of the

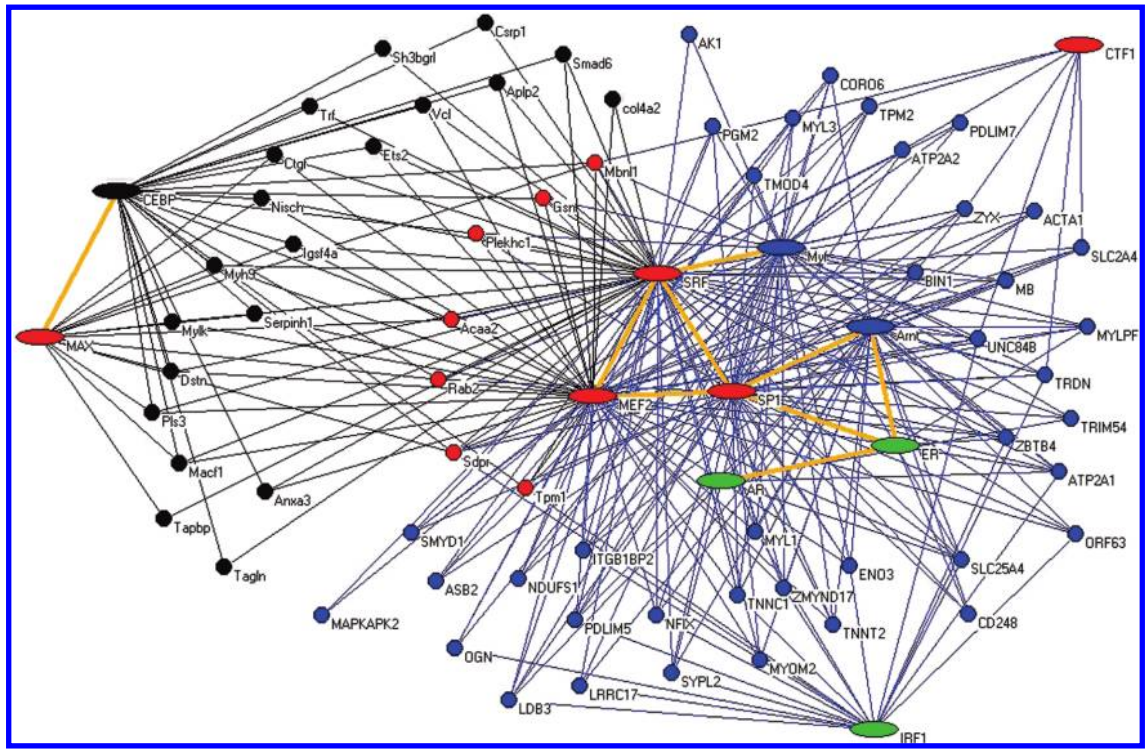

Figure 5. SRF-centered regulatory networks in lung and muscle. Kamada-Kawai interaction network of the transcription factor SRF and 66 of its target genes and 10 coregulatory TFs found in skeletal muscle and lung as predicted by oPOSSUM analysis. Oval symbols represent coregulatory TFs and the lines linking TFs and target genes indicate transcription regulatory relationships. Muscle-specific expressed genes are in blue, lung-specific expressed genes are in black and genes expressed in both tissues are in red. Broad (beige) lines linking TFs indicate known protein to protein interactions.

predicted TFs of mouse, the tissue-specific signatures of TF activities could provide support for predicted functional interaction among several of the TFs. For example, HNF1 and HNF4 activities cluster together closely and have high activities in liver and kidney, consistent with predictions that HNF1 interacts with HNF4.[5,21] Similarly, HNF1 and C/EBPalpha, which also reportedly interact,[22] were both highly active in the liver. Generally, an individual TF was found active in several particular tissues and rarely in only one tissue. Considering that only six tissues were examined, it is likely that a majority of TFs are active in more than one mouse tissue, interacting with different coregulatory partners in different tissues to coordinate the expression of different tissue-specific target genes.

SRF-Centered Regulatory Networks in Lung and Muscle. The assembly of regulatory interactions linking transcription factors to their target genes in the tissue can be viewed as a directed graph, here called a transcriptional regulatory network in which the regulators and targets represent the nodes and the regulatory interactions are the edges. The transcriptional regulatory networks could be constructed by integration of the mRNA expression data and the transcription factors regulating them, as predicted and experimentally validated by MOUSE OATFA (Figure S2 upper section in Supporting Information).

Several TFs, such as MEF2 (G62) and SRF (G64), determined by MOUSE OATFA to be active in skeletal muscle, were also active in lung (Figure 3). Experimental studies of tissue-specific gene transcription by Wasserman and colleagues have suggested that TFs act via regulatory TF modules, in which clusters of TF binding sites bring bound TFs into close (physical) contact to confer tissue specificity.[1,3] The modular concept suggests that particular TFs may interact to effect activation of different genes in different tissues through interactions with other particular tissue-specific TFs in each tissue. To explore this idea further, an oPOSSUM analysis[4] revealed the interactions (Figure S2, lower section for flowchart in Supporting Information) that involve the transcription factor SRF and 66 of its target genes in both skeletal muscle and lung, and the relationship with a further 10 coregulatory TFs predicted by the oPOSSUM analysis, each confirmed active by OATFA in those tissues, which can be illustrated in a Kamada-Kawai free network (Figure 5) drawn using the Pajek 1.21 program.[23] In addition to SRF, the core TFs comprised ER, AR and IRF1 (green ovals) and Arnt and Myf (blue ovals) in skeletal muscle and C/EBP in lung (black ovals), with SRF, MEF2, MAX, SP1 and CTF1 active in both tissues (red ovals). Target genes whose expression is regulated by combinations of TFs are also indicated for muscle (blue circles and connecting lines), lung (black circles and connecting lines) and both tissues (red circles) joined to their appropriate regulatory TFs. The transcription factors ER, AR and IRF1 (green ovals) are newly identified in skeletal muscle by this analysis and were also validated by MOUSE OATFA. The broad pale lines interconnecting TFs represent known protein-to-protein interactions. The drawing of a free network of SRF and other coregulatory TFs and their target genes for both skeletal muscle and lung tissues clearly illustrates the concept of regulatory TF modules. Although a number of TFs and target genes function in both tissues (red symbols), the other members of this set of SRF and coregulatory TFs and target genes form two clusters of either lung specific (black symbols) or skeletal muscle specific (blue and green symbols) functionalities. The MOUSE OATFA results also confirmed each of the 11 TFs predicted by the oPOSSUM analysis for the Kamada-Kawai free network (Table S3 in Supporting Information), validating the core of the inferred

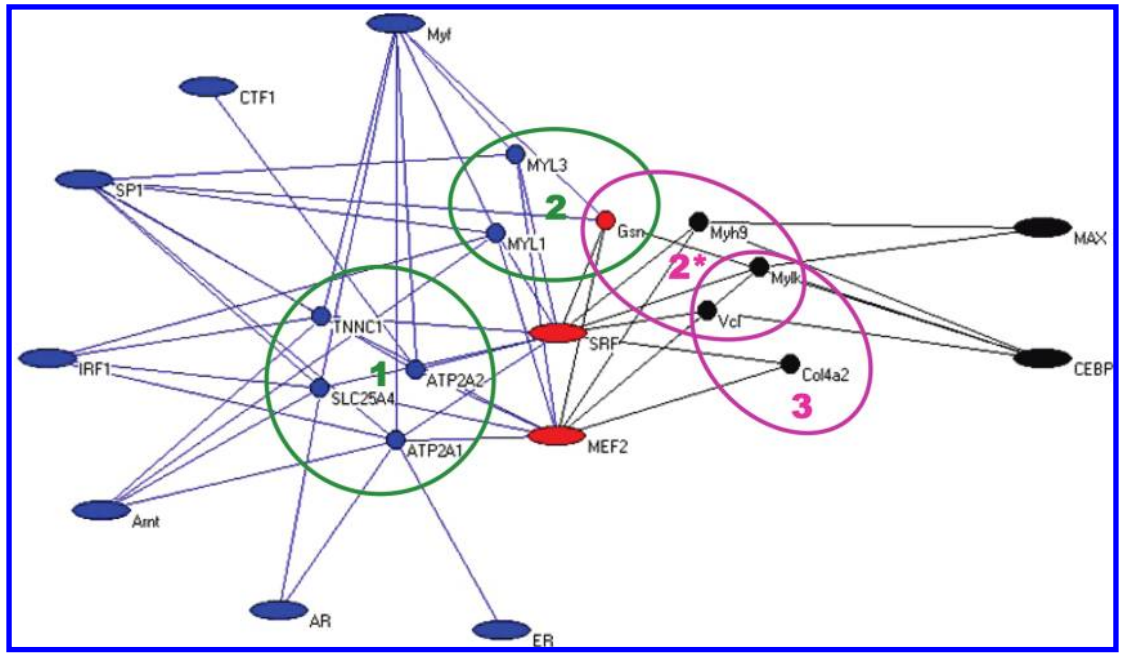

Figure 6. Pathway analyses of the target genes of SRF in lung and muscle. Kamada-Kawai interaction pathways of the transcription factor SRF and 11 of its target genes and 10 coregulatory TFs found in skeletal muscle and lung as predicted by oPOSSUM analysis. Circle 1 indicates the expressed genes involved in the 'calcium signaling pathway' in muscle. Circles 2 and 2* indicate the expressed genes involved in the 'regulation of actin cytoskeleton pathway' in muscle and lung, respectively. Circle 3 indicates the expressed genes involved in the 'focal adhesion pathway' in lung. Oval symbols represent coregulatory TFs and the lines linking TFs and genes indicate transcription regulatory relationships.

relational network. The direct benefit of the validation of the activity of each of the TFs by MOUSE OATFA is that such interaction networks may be drawn with greater confidence, and offers more realistic predictions for further testing of the biological networks active in specific tissues.

The pathways predicted by the MAS software (Supplementary Figure S3, column 1, in Supporting Information) (http://bioinfo.capitalbio.com/mas, and see the Supporting Information for more details) to be active in the six tissues (columns 2−7) at a significant *p*-value of $<10^{-6}$ were compared to the SRF-regulated pathways present in lung and muscle (columns 8 and 9). Three SRF-regulated pathways were detected at this significance level, the 'regulation of actin cytoskeleton pathway' in both lung and muscle, and the 'focal adhesion pathway' in lung and the 'calcium signaling pathway' in muscle, respectively. The 'regulation of actin cytoskeleton pathway' can direct many essential cellular processes, including cell−matrix formation, cell−cell adhesion and cell motility. The 'regulation of actin cytoskeleton pathway' can also direct changes in the shape of cells and was also predicted active in heart, but not in brain and kidney. Elements of each of these pathways have been reported in these tissues; thus, we illustrate the relationships between the transcription factor SRF and the 10 coregulatory TFs with the 11 target genes involved in these 3 pathways in both skeletal muscle and lung in a Kamada-Kawai free network (Figure 6). Again the concept of regulatory TF modules can be seen in regard to the genes predicted active in three different pathways in lung and muscle, respectively. In skeletal muscle, *Myl3*, *Myl1* and *Gsn* function in the 'regulation of actin cytoskeleton pathway' (circle 2), while *Gsn*, *Myh9*, *Vcl* and *Mylk* function in this pathway in lung (circle 2*). In addition, *Vcl* and *Mylk* are predicted to function with *Col4a2* in the 'focal adhesion pathway' pathway (circle 3) in lung. The 'calcium signaling pathway' (circle 1) in muscle, involving *Tnnc1*, *Slc25a4*, *Atp2a1* and *Atp2a2* genes, was independent of the other pathways. The active TFs in muscle and lung, which regulate expression of these genes, are indicated along the outer circle linked to their target genes by lines. Different genes appear to be involved in different cascades of the same signaling pathway and are coregulated by some TFs in common, and also by tissue-specific TFs in each of the two tissues, further illustrating the functional modularity of the coregulation of these pathways.

In lung, a majority of the genes predicted to be regulated by the SRF-centered networks have not been reported in that tissue, and those which had been reported were usually validated experimentally for a single TF. We therefore chose 7 target genes (*Acaa2*, *Myh9*, *Mylk*, *Serpinh1*, *Tagln*, *Tpm1*, *Vcl*) expressed in lung and predicted to be coregulated by SRF, C/EBPalpha and MEF2 for further validation. The predicted TF binding sites are close to the transcription initiation site of each gene and three genes (*Myh9*, *Mylk*, *Vcl*) are core elements of the MAS predicted significant pathways 2* and 3 shown in Figure 6. A majority (16/19) of the oPOSSUM/ MOUSE OATFA predictions were validated by the ChIP data (Figure 7). Several of these coregulatory functions in lung were novel. The regulation of *Vcl* by SRF[24] and *Tagln* by SRF[25] was known previously. Here, the combination of MOUSE OATFA/oPOSSUM/ChIP indicates that the *Myh9*, *Mylk* and *Vcl* genes are coregulated by SRF, C/EBPalpha and MEF2, and Acaa2 and Tagln are coregulated by SRF and C/EBPalpha but not by MEF2, both of which have not been reported previously. oPOSSUM also predicted that Tpm1 expression is regulated by all 3 TFs, but only the regulation by SRF was confirmed by ChIP. The presence of MEF2 was confirmed in lung by antibody assay (ChIP), after initial indication by MOUSE OATFA. We can see here that the combination of MOUSE OATFA assay with subsequent oPOSSUM analysis also greatly simplifies the

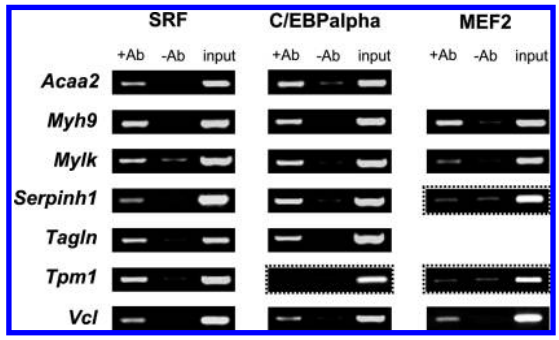

Figure 7. ChIP validation of the coregulation of target genes by TFs in mouse lung tissue. In each image, the 3 bands show, (left) the ChIP of DNA with either SRF, or C/EBP or MEF2 antibody, (middle) mock IP control with normal IgG, and (right) the total input DNA control. Each ChIP sample was PCR amplified with the target gene-specific primers indicated in the left-hand column. The dashed rectangles indicate predicted TF interactions which were not confirmed by ChIP.

identification of previously undiscovered coregulation of specific genes by particular TFs.

Discussion

With the completion of mouse genome,[26] whole-genome profiles of mouse tissue-specific gene expression[16] have led to the development of predicted networks of transcriptional control based on the analysis of differential gene expression and the regulatory TF modules associated with those genes.[8,9,27] These reports led us to compare the activities of 200 TFs determined by a new MOUSE OATFA active transcription factor profiling platform to the numerous predictions of probable TF activities extracted by bioinformatics analysis of coincident expression of target genes in six adult mouse tissues and with advanced analysis using oPOSSUM software which also considered the cobinding of multiple TFs at gene regulatory promoter elements.[4]

We used MOUSE OATFA validation of TF activities to explore the predicted regulatory pathways functioning in individual tissues and to develop regulatory networks of the TFs and their target genes with each TF activity experimentally determined, rather than predicted bioinformatically. There are about 70 transcription factors from vertebrates in the oPOSSUM prediction output, of which some 40 TFs are represented on the current version of MOUSE OATFA. If these 40 TFs are subjected to oPOSSUM analysis, all are predicted to regulate a cluster of target genes regardless of the *Z*-scores in each of the 6 tissues, yet from experimental MOUSE OATFA analysis with each tissue, less than half of the 40 TFs were found to be active (see Table 1). Although the binding sites of all 40 TFs occur in the promoter regions of genes as seen from the oPOSSUM predictions, one cause of some of the discrepancy between the predictions and the experimental MOUSE OATFA results might be that particular TFs are silent in some adult tissues and only active during particular stages of tissue development. The TFs which were active in tissues included "universal TFs" (Cluster 1) found in all tissues, as well as TFs active in particular adult tissues (see Supplementary Table S3 in Supporting Information). Since we could demonstrate independent validation of many of the active TFs found by MOUSE OATFA, although the complexity of inferred transcriptional regulatory network is decreased, we believe this analysis increases the reliability of active TF identification (see Table 2).

Table 1. The Signatures of Active TFs in the Six Tissues[a]

TFs	liver	kidney	lung	brain	heart	muscle
CREB1	3.11	3.87	−5.07	10.96	6.72	6.08
CTF1	−4.60	5.66	6.56	6.21	1.04	5.61
MAX	2.23	4.19	−3.32	0.80	−0.43	0.55
PPARG	4.63	−0.71	−2.70	−1.42	4.91	0.38
SP1	−2.37	−6.92	−6.81	2.15	−9.45	0.37
USF1	1.97	0.94	−5.57	−1.25	0.32	5.32
AR						9.39
Arnt						3.52
CEBP	8.47		5.15			
E2F1	−2.71	−4.11		−2.53	−5.93	−7.26
ER	14.06	5.98		2.82	11.10	5.12
FOXD1	0.29					
HNF1	10.21	11.37				
HNF4	14.19	10.89				
IRF1	−4.34	−2.41				−5.80
MEF2			−2.93	0.90	16.39	23.20
Myf				−1.38	−3.69	2.68
NF-Y			−3.57			
RXR	1.14			−2.07		
SRF			20.98	−3.38	18.90	24.20

[a] The active TFs were selected from oPOSSUM analysis in six tissues and confirmed experimentally by MOUSE OATFA. The numbers are *Z*-scores generated by the prediction software.

Table 2. Comparison of the Regulatory Networks Predicted by oPOSSUM with and without Inclusion of OATFA Data

		number of TFs	number of genes	number of TF-gene interactions
Liver	oPOSSUM	40	369	4586
	oPOSSUM + OATFA	**14**	**308**	**1368**
Kidney	oPOSSUM	40	328	3953
	oPOSSUM + OATFA	**11**	**245**	**837**
Lung	oPOSSUM	40	192	2923
	oPOSSUM + OATFA	**10**	**171**	**636**
Brain	oPOSSUM	40	519	8001
	oPOSSUM + OATFA	**12**	**441**	**1759**
Heart	oPOSSUM	40	276	3791
	oPOSSUM + OATFA	**11**	**230**	**842**
Muscle	oPOSSUM	40	290	4168
	oPOSSUM + OATFA	**14**	**248**	**1217**

Contrasting the new observation of SRF activity in lung with its well-established activity in skeletal muscle, we used oPOSSUM to develop novel networks in both tissues which involved a total of 66 target genes and 10 coregulatory TFs (see Figure 5). These networks illustrate the different interactions SRF undertakes with different sets of coregulatory TFs in lung and muscle, resulting in expression of different sets of tissue-specific gene targets. Protein-to-protein interactions are known to occur between particular TFs involved in these networks, such as the SP1/MEF2 pair[28] and the C/EBP/MAX pair,[29] yet potentially novel TF-TF interactions might be expected between several TFs seen to coregulate the numerous target genes. For example, interactions might be predicted to occur in lung between C/EBPalpha, MAX and SRF or MEF2, which could be tested. Using molecular annotation system (MAS) analysis, we further investigated the regulatory pathways predicted to be active in the six tissues (significant at p-value $<10^{-6}$) and compared them to SRF-regulated pathways predicted to occur in lung and muscle. Three pathways were identified, the 'regulation of actin cytoskeleton pathway' occurring in lung and muscle, the 'focal adhesion pathway' in lung and the 'calcium signaling pathway' in muscle, respectively. oPOSSUM analysis of the regulatory TFs and their target genes represented in these three pathways (see Figure 6) in lung and muscle again revealed that different combinations of validated TF activities regulate different sets

of target genes in each tissue, even for the same pathway as shown in the 'regulation of actin cytoskeleton pathway'.

Interestingly, the expression of a number of these coregulated genes has been reported to be directly related with human lung diseases, or to be affected by drug treatments. *Mylk* which resides in the center of 'regulation of actin cytoskeleton pathway' and 'focal adhesion pathway' (Figure 6) shows significant association with both asthma and sepsis.[30] *Serpinh1*, also called heat shock protein 47 (*Hsp47*), is a collagen-specific molecular chaperone that has been shown to play a major role during the processing and/or secretion of procollagen. The expression of *Serpinh1* is typically up-regulated during pathological conditions in which collagen deposition occurs in the lung. LPS-stimulated acute respiratory distress syndrome (ARDS) in rat is accompanied by strongly induced expression of *Serpinh1* mRNA as lung damage progresses,[31] and the inhibition of *Serpinh1* expression improves bleomycin-induced pulmonary fibrosis.[32] Although we have not considered any pathological conditions during our study of tissue-specific TFs in mouse, we suggest that the examination of coregulatory TF activities could be a productive avenue for initial investigation of the regulatory pathways and mechanisms affecting pathological states.

Although a number of bioinformatics approaches have been developed that infer both dynamic TF networks and TF activities from time-course microarray gene expression studies,[33,34] or from parallel ChIP protein profiling analyses,[35,36] the work described here is the first to systematically investigate the coordinated network of transcriptional activation and gene expression responses of a biological system by combined high-throughput TF activity profiling, microarray expression assay and bioinformatics analysis across several mammalian tissues. Despite the discrepancy in the large-scale analysis of gene expression microarrays (10 000 genes) compared to MOUSE OATFA (200 TFs) in respect to data points, remarkably MOUSE OATFA provided additional, verifiable identifications of TF activities not seen by expression profiling alone. The current MOUSE OATFA platform examines the activities of only 200 TFs, yet it represents the largest scale of TF profiling currently available.

In summary, our analysis shows that the pathway activities inferred from gene transcriptional evidence and bioinformatics modeling alone may be not considered definitive, and that additional independent evidence such as tissue-specific TF activities should be taken into account. Here, the inclusion of TFs shown active by MOUSE OATFA in lung tissues into network and pathway modeling exercises were subsequently confirmed by independent ChIP analyses for a large number of the predictions. These confirmed predictions including a number of novel pathways and networks. Our analyses caution the inference of networks of TF functionalities and gene targets by bioinformatics alone, and suggest that direct measurements of TF activities should be included into such modeling exercises. The TF binding data are illustrative of the power of the MOUSE OATFA platform to provide unique information concerning the presence of active TFs in individual tissues, as the expression of their TF-encoding genes does not always provide an accurate representation of their likely activities. We further suggest that two levels of genome-wide information are ideally required for informed bioinformatics-based development of network models and active tissue-specific pathways. Gene expression data should be used to identify global transcription (changes) and parallel analyses of TF activities by MOUSE OATFA to provide guidance to TF activity predictions. Although MOUSE OATFA is highly reproducible and provides parallel high-throughput analyses, being based on consensus TF binding matrices it should be considered a rapid guide technology with some limit to the precise identification of unique TFs. It is well-recognized that many mammalian TFs can bind to several closely related sequences and that members of families of TFs may bind with very similar gene regulatory sequences (see Supplementary Table S2 in Supporting Information). Following the identification of putative TF activities by MOUSE OATFA, the identity of TFs of interest should ideally be confirmed directly by independent antibody-based detection methods such as ELISA or by ChIP.

In a review of 144 human developmental disorders in which the function of the causative gene had been identified, 49 (34%) were due to mutated TF genes, a number nearly double that of the next largest class.[37] Several comparative analyses of the inferred active regulatory pathways have also been used to predict likely cell responses during disease and infection.[38,39] The ability to identify TF tissue signatures, and perturbations in these signatures, has the potential to indicate early disease responses that might be used in a predictive capacity. We suggest that the MOUSE OATFA platform offers a new concept for the identification of biomarkers for disease and infection states. The ability to identify inappropriate reactivation of developmentally regulated TF functions, or inappropriate TF activity levels during progressive and degenerative diseases might provide new potential targets for drug treatments, markers for disease diagnosis and for prognosis of the outcomes of human diseases.

Conclusions

We first developed a sensitive and reliable mouse oligonucleotide array-based transcription factor assay (MOUSE OATFA) based on our previous work with human transcription factors. The DNA-binding activities of about 200 different TF proteins in six mouse tissues (liver, kidney, lung, brain, heart and skeletal muscle) were then determined and validated. Using oPOSSUM software, we combined the TF signatures from MOUSE OATFA and public mRNA expression profiles from each tissue to construct experimental transcriptional regulatory networks, finding novel networks of expressed genes coregulated by sets of TFs in each tissue. As an example of that analysis, newly proposed SRF-centered regulatory networks in lung and muscle were tested using ChIP to confirm the involvement of particular TFs. Here, some 16/19 predicted target genes were confirmed as coregulated, including several TF-gene interactions for genes encoding proteins associated with asthma and sepsis. The combination of MOUSE OATFA with bioinformatics analysis of expressed genes gives deep insight into complex transcriptional systems and their regulatory pathways in mouse.

Acknowledgment. This work was supported by the grant of the National Hi-Tech Program of China (No. 2006AA020701).

Supporting Information Available: Performance of the MOUSE OATFA platform. Identification of tissue-specific genes and tissue-expressed genes. Correlation between the OATFA results with published tissue-specific signatures. Comparative OPOSSUM analysis. Molecular Annotation System (MAS). The rank order of relative TFs activities in LPS stimu-

lated RAW264.7 cells (Figure S1). Flowchart of transcriptional regulatory network construction (Figure S2). Predictive MAS analysis of active pathways (Figure S3). Sequences of the 240 probes on the MOUSE OATFA microarray (Table S1). List of the distribution of the TF capture probes into functional groups (Table S2). Details of the TF activities in the six tissues (Table S3). The oligonucleotide probe sequences in the TF-ELISA (Table S4). Correlations between the MOUSE OATFA results and the literature (Table S5). The oPOSSUM analytical results of the genes in the six tissues (Table S6). PCR primer sequences of the regulated genes in ChIP assays (Table S7). List of tissue-specific and tissue-expressed genes (Table S8). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wasserman, W. W.; Fickett, J. W. Identification of regulatory regions which confer muscle-specific gene expression. *J. Mol. Biol.* **1998**, *278*, 167–181.

(2) Lein, E. S.; Hawrylycz, M. J.; Ao, N.; Ayres, M.; Bensinger, A.; Bernard, A.; Boe, A. F.; Boguski, M. S.; Brockway, K. S.; Byrnes, E. J.; et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **2007**, *445*, 168–176.

(3) Krivan, W.; Wasserman, W. W. A predictive model for regulatory sequences directing liver-specific transcription. *Genome Res.* **2001**, *11*, 1559–1566.

(4) Ho Sui, S. J.; Mortimer, J. R.; Arenillas, D. J.; Brumm, J.; Walsh, C. J.; Kennedy, B. P.; Wasserman, W. W. oPOSSUM, identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic Acids Res.* **2005**, *33*, 3154–3164.

(5) Yu, X.; Lin, J.; Zack, D. J.; Qian, J. Computational analysis of tissue-specific combinatorial gene regulation, predicting interaction between transcription factors in human tissues. *Nucleic Acids Res.* **2006**, *34*, 4925–4936.

(6) Smith, A. D.; Sumazin, P.; Zhang, M. Q. Tissue-specific regulatory elements in mammalian promoters. *Mol. Syst. Biol.* **2007**, *3*, 73.

(7) Barrera, L. O.; Li, Z.; Smith, A. D.; Arden, K. C.; Cavenee, W. K.; Zhang, M. Q.; Green, R. D.; Ren, B. Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome Res.* **2008**, *18*, 46–59.

(8) Yu, X.; Lin, J.; Zack, D. J.; Qian, J. Identification of tissue-specific cis-regulatory modules based on interactions between transcription factors. *BMC Bioinf.* **2007**, *8*, 437.

(9) Zhu, Z.; Shendure, J.; Church, G. M. Discovering functional transcription-factor combinations in the human cell cycle. *Genome Res.* **2005**, *15*, 848–855.

(10) Shao, W.; Wei, H. J.; Qiao, J. Y.; Zhao, Y. C.; Sun, Y. M.; Zhou, Y. X.; Cheng, J. Parallel profiling of active transcription factors using an oligonucleotide array-based transcription factor assay (OATFA). *J. Proteome Res.* **2005**, *4*, 1451–1456.

(11) Zhao, Y. C.; Shao, W.; Wei, H. J.; Qiao, Y. J.; Lu, Y.; Sun, Y. M.; Mitchelson, K.; Cheng, J.; Zhou, Y. X. Development of a novel oligonucleotide array-based transcription factor assay platform for genome-wide active transcription factor profiling in *Saccharomyces cerevisiae*. *J. Proteome Res.* **2008**, *7*, 1315–1325.

(12) Qiao, J.; Shao, W.; Wei, H.; Sun, Y.; Zhao, Y.; Xing, W.; Zhang, L.; Mitchelson, K.; Cheng, J. Novel high-throughput profiling of human transcription factors and its use for systematic pathway mapping. *J. Proteome Res.* **2008**, *7*, 2769–2779.

(13) Wingender, E.; Chen, X.; Hehl, R.; Karas, H.; Liebich, I.; Matys, V.; Meinhardt, T.; Pru, M.; Reuter, I.; Schacherer, F. TRANSFAC, an integrated system for gene expression regulation. *Nucleic Acids Res.* **2000**, *28*, 316–319.

(14) Matys, V.; Fricke, E.; Geffers, R.; GÖssling, E.; Haubrock, M.; Hehl, R.; Hornischer, K.; Karas, D.; Kel, A. E.; Kel-Margoulis, O. V. TRANSFAC®, transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **2003**, *31*, 374–378.

(15) Shang, Y.; Hu, X.; DiRenzo, J.; Lazar, M. A.; Brown, M. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **2000**, *103*, 843–852.

(16) Zhang, W.; Morris, Q. D.; Chang, R.; Shai, O.; Bakowski, M. A.; Mitsakakis, N.; Mohammad, N.; Robinson, M. D.; Zirngibl, R.; Somogyi, E. The functional landscape of mouse gene expression. *J. Biol.* **2004**, *3*, 21.

(17) Yang, H. L.; Zhao, R. Y.; Yang, H. Y.; Lee, M. H. Constitutively active FOXO4 inhibits Akt activity, regulates p21 Kip1 stability, and suppresses HER2-mediated tumorigenicity. *Oncogene* **2005**, *24*, 1924–1935.

(18) Kolkman, A.; Slijper, M.; Albert, J. R. H. Development and application of proteomics technologies in *Saccharomyces cerevisiae*. *Trends Biotechnol.* **2005**, *23*, 598–604.

(19) Lee, Y. H.; Sauer, B.; Johnson, P. F.; Gonzalez, F. J. Disruption of the C/EBPalpha gene in adult mouse liver. *Mol. Cell. Biol.* **1997**, *17*, 6014–6022.

(20) Martis, P. C.; Whitsett, J. C.; Xu, Y.; Perl, A. K. T.; Wan, H. J.; Ikegami, M. C/EBPalpha is required for lung maturation at birth. *Development* **2006**, *133*, 1155–1164.

(21) Stoffel, M.; Duncan, S. A. The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13209–13214.

(22) Wu, K. J.; Wilson, D. R.; Shih, C.; Darlington, G. J. The transcription factor HNF1 acts with C/EBPalpha synergistically activate the human albumin promoter through a novel domain. *J. Biol. Chem.* **1994**, *269*, 1177–1182.

(23) Portales-Casamar, E.; Kirov, S.; Lim, J.; Lithwick, S.; Swanson, M. I.; Ticoll, A.; Snoddy, J.; Wasserman, W. W. PAZAR, a framework for collection and dissemination of cis-regulatory sequence annotation. *Genome Biol.* **2007**, *8*, R207.

(24) Gineitis, D.; Treisman, R. Differential usage of signal transduction pathways defines two types of serum response factor target gene. *J. Biol. Chem.* **2001**, *276*, 24531–24539.

(25) Du, K. L.; Ip, H. S.; Li, J.; Chen, M.; Dandre, F.; Yu, W.; Lu, M. M.; Owens, G. K.; Parmacek, M. S. Myocardin is a critical serum response factor cofactor in the transcriptional program regulating smooth muscle cell differentiation. *Mol. Cell. Biol.* **2003**, *23*, 2425–2437.

(26) Waterston, R. H.; Lindblad-Toh, K.; Birney, E.; Rogers, J.; Abril, J. F.; Agarwal, P.; Agarwala, R.; Ainscough, R.; Alexandersson, M.; An, P.; et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* **2002**, *420*, 520–562.

(27) Hallikas, O.; Palin, K.; Sinjushina, N.; Rautiainen, R.; Partanen, J.; Ukkonen, E.; Taipale, J. Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **2006**, *124*, 47–59.

(28) Grayson, J.; Bassel-Duby, R.; Williams, R. S. Collaborative interactions between MEF-2 and Sp1 in muscle-specific gene regulation. *J. Cell. Biochem.* **1998**, *70*, 366–375.

(29) Zada, A. A.; Pulikkan, J. A.; Bararia, D.; Geletu, M.; Trivedi, A. K.; Balkhi, M. Y.; Hiddemann, W. D.; Tenen, D. G.; Behre, H. M.; Behre, G. Proteomic discovery of Max as a novel interacting partner of C/EBPalpha, a Myc/Max/Mad link. *Leukemia* **2006**, *20*, 2137–2146.

(30) Gao, L.; Grant, A. V.; Rafaels, N.; Stockton-Porter, M.; Watkins, T.; Gao, P.; Chi, P.; Munoz, M.; Watson, H.; Dunston, G.; et al. Polymorphisms in the myosin light chain kinase gene that confer risk of severe sepsis are associated with a lower risk of asthma. *J. Allergy Clin. Immunol.* **2007**, *119*, 1111–1118.

(31) Hagiwara, S.; Iwasaka, H.; Matsumoto, S.; Noguchi, T.; Yoshioka, H. Coexpression of HSP47 gene and type I and type III collagen genes in LPS-induced pulmonary fibrosis in rats. *Lung* **2007**, *185*, 31–37.

(32) Hagiwara, S.; Iwasaka, H.; Matsumoto, S.; Noguchi, T. Antisense oligonucleotide inhibition of heat shock protein (HSP) 47 improves bleomycin-induced pulmonary fibrosis in rats. *Respir. Res.* **2007**, *8*, 37.

(33) Pournara, T.; Wernisch, L. Factor analysis for gene regulatory networks and transcription factor activity profiles. *BMC Bioinf.* **2007**, *8*, 61.

(34) Sayyed-Ahmad, A.; Tuncay, K.; Ortoleva, P. J. Transcriptional regulatory network refinement and quantification through kinetic modeling, gene expression microarray data and information theory. *BMC Bioinf.* **2007**, *8*, 20.

(35) Gao, F.; Foat, B. C.; Bussemaker, H. J. Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinf.* **2004**, *5*, 31.

(36) Galbraith, S. J.; Tran, L. M.; Liao, J. C. Transcriptome network component analysis with limited microarray data. *Bioinformatics* **2006**, *22*, 1886–1894.

(37) Boyadiev, S. A.; Jabs, E. W. Developmental biology, Frontiers for clinical genetics. *Clin. Genet.* **2000**, *57*, 253–266.

(38) Chen, J.; Xu, H.; Aronow, B. J.; Jegga, A. G. Improved human disease candidate gene prioritization using mouse phenotype. *BMC Bioinf.* **2007**, *8*, 392.

(39) Vahtola, E.; Louhelainen, M.; Merasto, S.; Martonen, E.; Penttinen, S.; Aahos, I.; Kytö, V.; Virtanen, I.; Mervaala, E. Forkhead class O transcription factor 3a activation and Sirtuin1 overexpression in the hypertrophied myocardium of the diabetic Goto-Kakizaki rat. *J. Hypertens.* **2008**, *26*, 334–344.

PR800417E