

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Regulation of gene expression programs by serum response factor and megakaryoblastic leukemia 1/2 in macrophages

Permalink

<https://escholarship.org/uc/item/8cc7d0t0>

Author

Sullivan, Amy Lynn

Publication Date

2009

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Regulation of Gene Expression Programs by Serum Response Factor and
Megakaryoblastic Leukemia 1/2 in Macrophages

A dissertation submitted in partial satisfaction of the requirements for the
degree Doctor of Philosophy

in

Biomedical Sciences

by

Amy Lynn Sullivan

Committee in charge:

Professor Christopher K. Glass, Chair

Professor Stephen M. Hedrick

Professor Marc R. Montminy

Professor Nicholas J. Webster

Professor Joseph L. Witztum

2009

Copyright

Amy Lynn Sullivan, 2009

All rights reserved.

The Dissertation of Amy Lynn Sullivan is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2009

DEDICATION

To my husband, Shane, for putting up with me through all of the long hours, last minute late nights, and for not letting me quit no matter how many times my projects fell apart.

To my son, Tyler, for always making me smile and for making every day an adventure.

To my gifted colleagues, for all of the thought-provoking discussions, technical help and moral support through the roller-coaster ride that has been my graduate career.

To my family and friends, for all of your love and support. I couldn't have done it without you!

EPIGRAPH

If at first you don't succeed, try, try, again.

-T.H. Palmer's Teacher's Manual (1840)

TABLE OF CONTENTS

Signature Page	iii
Dedication	iv
Epigraph	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
Acknowledgements	x
Vita	xi
Abstract of the Dissertation	xii
Chapter 1 Introduction	1
Introduction	2
Regulation of the actin cytoskeleton	3
Regulation of macrophage function by Rho GTPases	5
Regulation of transcription by the Serum Response Factor	8
Activation of SRF by the Ternary Complex Factors	10
Activation of SRF by the myocardin family of coactivators	12
Differential regulation of SRF by the TCFs and myocardin/MKL factors	14
Animal models of SRF, TCF, and myocardin/MKL function	15
Chapter 2 SRF and MKL1/2 Regulation of Macrophage Gene Expression	21

Introduction	22
Materials and Methods	26
Results	36
Discussion	50
Acknowledgements	56
Chapter 3 Future Directions and Conclusion	79
Introduction	80
Understanding SRF association with CArG-less DNA	80
Elucidating the mechanism of SRF regulation from distal sites ...	88
Control of cell-type and non-cell-type specific gene expression by SRF	90
Assaying the requirement of MKL and SRF for macrophage function through mouse models	92
Conclusion	96
References	97

LIST OF FIGURES

Figure 1	Mechanisms of SRF transcriptional regulation	20
Figure 2	MKL1/2 and SRF knockdown in samples used for microarray analysis	57
Figure 3	MKL and SRF regulate gene expression in macrophages	59
Figure 4	Genomic location annotation and motif analysis of SRF ChIP-seq peaks in primary macrophages	66
Figure 5	SRF associates with different transcription factors depending on genomic location	67
Figure 6	Motif distribution relative to the center of the SRF ChIP-seq peak	68
Figure 7	Motif distribution relative to the SRF motif	69
Figure 8	Distal SRF/PU.1 peaks are associated with the H3K4me1 enhancer mark	70
Figure 9	Heat map analysis of ChIP-seq data in primary macrophages	71
Figure 10	Motif analysis of SRF ChIP-seq peak sequences in undifferentiated and differentiated PUER cells	74
Figure 11	The H3K4me1 enhancer mark is acquired only at sites that gain both SRF and PU.1	75
Figure 12	SRF, PU.1, and MKL regulate hematopoietic-specific cytoskeletal gene expression in macrophages	76
Figure 13	Confirmation of SRF, PU.1 and MKL1/2 knockdown in primary macrophages	77

LIST OF TABLES

Table 1	Top genes downregulated by MKL1/2 knockdown	58
Table 2	Full results of gene ontology analysis for genes regulated by MKL	60
Table 3	Top genes downregulated by SRF knockdown	61
Table 4	Full results of gene ontology analysis for genes regulated by SRF	62
Table 5	Top genes commonly downregulated by MKL or SRF knockdown	63
Table 6	Full results of gene ontology analysis for genes regulated by both MKL and SRF	64
Table 7	Top SRF ChIP-seq peaks in primary macrophages	65
Table 8	Top SRF ChIP-seq peaks in undifferentiated PUER cells	72
Table 9	Top SRF ChIP-seq peaks in differentiated PUER cells	73
Table 10	Transcription factors that are downregulated after SRF knockdown	78

ACKNOWLEDGEMENTS

I would like to recognize Professor Chris Glass for all of his help and support as the chair of my committee. He exercised extraordinary patience throughout the development of my projects and always had an open door when I needed guidance.

I also would like to acknowledge all of the current and past members of the Glass lab for all of the constructive criticism regarding my projects and for the seemingly endless amount of technical help.

Chapter 2, in its entirety, is currently being prepared for submission for publication of the material. Current co-authors include Sven Heinz, Chris Benner and Chris Glass. The dissertation author was the primary investigator and author of this material.

My graduate work was supported by the Cancer Cell Biology Training Grant, an American Heart Association pre-doctoral fellowship, and National Institute of Health grant HL088083.

VITA

EDUCATION

- 2009 Ph.D. in Biomedical Sciences, University of California, San Diego, San Diego, CA
- 2000 B.A. in Biology-Chemistry, Whitman College, Walla Walla, WA

RESEARCH EXPERIENCE

- 2000-2002 Research Assistant, Oregon Health Sciences University, Advisors: Dr. Soren Impey, Dr. Richard Goodman
- 1999 Summer Research Student, Johns Hopkins University School of Medicine, Advisors: Dr. Nicole Stricker, Dr. Richard Huganir
- 1998 Summer Research Student, Whitman College, Advisor: Dr. J. Charles Templeton

PUBLICATIONS

G. Pascual, **A.L. Fong**, S. Ogawa, A. Gamliel, A.C. Li, V. Perissi, D.W. Rose, T.W. Willson, M.G. Rosenfeld, C.K. Glass. A SUMOylation-Dependent Pathway Mediates Transrepression of Inflammatory Response Genes by PPAR-gamma. *Nature*. 437 (2005) 759-763.

J.S. Arthur, **A.L. Fong**, J.M. Dwyer, M. Davare, E. Reese, K. Obrietan, S. Impey. Mitogen- and Stress-Activated Protein Kinase 1 Mediates cAMP Response Element-Binding Protein Phosphorylation and Activation by Neurotrophins. *J. Neurosci*. 24 (2004) 4324-4332.

D. Sawka-Verhelle, L. Escoubet-Lozach, **A.L. Fong**, K.D. Hester, S. Herzig, P. Lebrun, C.K. Glass. PE-1/METS, an Antiproliferative Ets Repressor Factor, is Induced by CREB-1/CREM-1 during Macrophage Differentiation. *J. Biol. Chem*. 279 (2004) 17772-17784.

S. Impey, **A.L. Fong**, Y. Wang, J-R. Cardinaux, D.M. Fass, K. Obrietan, G.A. Wayman, D.R. Storm, T.R. Soderling, R.H. Goodman. Phosphorylation of CBP Mediates Transcriptional Activation by Neural Activity and CaM Kinase IV. *Neuron*. 34 (2002) 235-244.

ABSTRACT OF THE DISSERTATION

Regulation of Gene Expression Programs by Serum
Response Factor and Megakaryoblastic Leukemia 1/2 in
Macrophages

by

Amy Lynn Sullivan

Doctor of Philosophy in Biomedical Sciences

University of California, San Diego, 2009

Professor Christopher K. Glass, Chair

Macrophages are key players in the regulation of the innate and adaptive immune responses, tissue homeostasis and wound healing. In response to inflammation or injury, macrophages are actively recruited to affected sites where they perform specialized functions, such as phagocytosis of invading pathogens or dead cells, which are key for host recovery and survival. Many studies have shown that members of the Rho family of GTPases are important mediators of macrophage recruitment and function. These proteins activate the signaling cascades that are necessary to control

the dynamic changes in the actin cytoskeleton that are required for directed movement and protrusion of the extracellular membrane. In several non-macrophage cell types, it has also been described that Rho GTPases can stimulate the expression of differentiation and cytoskeletal genes. This gene regulation was shown to be mediated by the DNA-binding transcription factor, serum response factor (SRF) and its coactivators megakaryoblastic leukemia (MKL) 1 and 2. Based on this data, the goal of this study was to elucidate the gene programs that are regulated by SRF and MKL1/2 in macrophages. Using mRNA expression profiling of primary macrophages, we discovered that both MKL1/2 and SRF regulate both general and hematopoietic-specific genes. In order to determine which of these genes are direct targets of the SRF pathway, we performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) analysis. Our results showed that, in general, SRF binding is not restricted to the proximal promoter, but occurs primarily at distal sites (further than +/- 500bp relative to the TSS) in close association with the macrophage and B-cell specific factor, PU.1. In particular, SRF and PU.1 were found to be localized to distal sites of several hematopoietic-specific target genes. Subsequent siRNA knockdown experiments showed that both SRF and PU.1 are required for full expression of these genes, providing insights into how cell-specific programs of SRF-dependent gene expression are achieved.

Chapter 1

Introduction

Introduction

Macrophages are responsible for a broad range of functions regarding tissue homeostasis and host defense. During development, they are responsible for the clearance of apoptotic cells resulting from normal tissue remodeling and have been shown to be required for the proper development of tissues such as bones, adipose, mammary glands, and the pancreas (β -cells)[1]. Macrophages also help maintain tissue health by regulating normal cell turnover and by performing constant surveillance for invading bacteria, viruses, or other pathogens.

In response to infection or injury, macrophages participate in the initiation and amplification of inflammatory responses through the release of chemotactic and pro-inflammatory cytokines that recruit neutrophils and circulating monocytes to the site of infection. Recruitment of monocytes requires cell adhesion to activated endothelium and directed migration into tissues in response to chemotactic signals. Once in the inflamed or injured tissue, monocytes terminally differentiate into macrophages that perform numerous functions, including phagocytosis of pathogens, dead cells, and cellular debris. In addition to playing roles in innate immunity through killing of pathogens, phagocytosed peptides can also be processed inside of the phagolysosome and subsequently presented to T-cells to regulate adaptive immune responses.

All of the functions that have so far been described are highly dependent on the proper regulation of the actin cytoskeleton. Through the regulated polymerization and depolymerization of actin, macrophages can carry out their specialized functions, such as directed movement in response to a chemotactic gradient, pseudopod extension and closure during phagocytosis, fusion of the phagosome with the lysosome for peptide processing, and clustering of MHC molecules on the cell surface for presentation to T-cells.

In addition to activation or repression of cytoskeletal binding proteins, regulating the expression of cytoskeleton-associated genes also impacts cell motility and function. Serum response factor (SRF) and its coactivators from the myocardin family of transcription factors (myocardin, megakaryoblastic leukemia (MKL) 1, and MKL2) have been shown in several cell types to be key regulators of cytoskeletal and contractile gene expression in both a cell-type and non-cell-type dependent manner[2-4]. The focus of this study is to identify and understand the regulation of SRF and MKL-dependent gene expression in the macrophage, with particular emphasis on the regulation of hematopoietic-specific cytoskeletal target genes.

Regulation of the actin cytoskeleton

Actin is the most abundant protein in eukaryotes and was originally described as a scaffold protein important for the maintenance of cell shape and for the directed movement of cells in response to extracellular signals[5].

Globular actin (G-actin) monomers are able to polymerize into actin filaments (F-actin) which grow in a polarized fashion. G-actin monomers, bound to ATP, are incorporated into actin filaments at the plus end (or barbed end), which is oriented toward the plasma membrane and is the fast growing end of the filament. Over time, the ATP that is bound to each actin subunit is hydrolyzed to ADP, which destabilizes the filament and causes depolymerization of the ADP-actin subunit from the end of the filament. Thus, local regulation of the rate of actin polymerization and depolymerization (actin treadmilling) is critical for net movement in any particular direction[5].

Interestingly, *in vivo* observations of lamellipodia formation showed that the actual rate of actin polymerization at plus ends is around one hundred times faster than the rate of actin polymerization *in vitro*[6]. Since actin depolymerization is the rate limiting step in the actin treadmilling process, actin filaments will polymerize to the point where there is a balance between actin polymerization and depolymerization and then maintain that rate of actin treadmilling *in vitro*. In order to get the polymerization rate observed *in vivo*, intracellular G-actin concentrations must be maintained locally at very high levels with the assistance of actin binding proteins that have the ability to significantly alter the balance between polymerization and depolymerization[7].

Many actin binding proteins have been identified that are crucial for regulating cellular functions requiring cytoskeletal rearrangement. For example, general factors such as actin depolymerizing factor (ADF)/cofilin

and the filament capping protein, gelsolin, are able to increase the steady state of G-actin that is available for polymerization at the plasma membrane to facilitate growth of actin filaments [8]. Ubiquitously expressed actin binding proteins are also able to control the directionality of actin filament growth by nucleating actin branches (Arp2/3 complex) in response to extracellular stimuli[9].

In addition to general cytoskeletal factors, cells may harbor cell-type specific proteins that allow them to carry out their differential functions. In the hematopoietic lineage, several cell-type specific actin binding proteins have been identified that affect cell function and immune response. For example, coronin 1a (Coro1a) is a hematopoietic-specific actin binding protein that has been shown to play a role in the negative regulation of phagosome-lysosome fusion. Retention of Coro1a on the phagosome by phagocytosed mycobacteria prevents phagosome-lysosome fusion and serves as a mechanism for host immune evasion [10]. Mouse knockout models of another actin binding protein, Lymphocyte/Leukocyte specific protein 1 (LSP1), have also shown defects in leukocyte migration and chemotaxis in response to various cytokines[11]. Taken together, actin binding proteins serve as key targets for the integration of cellular signals and subsequent changes in the rate and directionality of actin dynamics, which have important consequences for both cell-type and non-cell-type specific cellular functions.

Regulation of macrophage function by Rho GTPases

Cytoskeletal rearrangements often occur following the activation of intermediate signaling proteins downstream of membrane receptor binding to extracellular molecules (e.g. cytokines, growth factors, etc.). Some of the best characterized of these signaling intermediates belong to the Rho family of guanosine triphosphatases (GTPases). The Rho family consists of more than 20 members, but the most studied are those proteins in the Rho, Rac, and cdc42 subfamilies[12]. Rho proteins are activated through binding to guanine nucleotide exchange factors (GEFs) that catalyze the exchange of bound GDP to GTP. Conversely, binding of GTPase activating proteins (GAPs) increases the rate of GTP hydrolysis to GDP, resulting in Rho protein inactivation. Binding to GTP causes a conformational change that allows Rho to bind to and induce activation of its effector molecules, such as Rho-associated kinase (ROCK), rhotekin, mDia and myosin light chain phosphatase[13]. It is these effectors that in turn activate cytoskeletal reorganization through alterations in actin binding protein function.

Macrophage migration to sites of inflammation is a multi-step process that leads to the activation of Rho, Rac, and cdc42[12]. Macrophages are able to sense and migrate toward the increasing gradient of chemotactic factors, such as Ccl2 and MCSF, which are released from the site of infection. Cdc42 has been shown to be essential for this process because knockout and dominant negative macrophage model systems for this factor have shown increases in migration speed, but preferential movement in the direction of the

chemotactic gradient was lost[14-15]. Additional experiments in mice showed that *Rac1*^{-/-}, *Rac2*^{-/-} and *Rac1/2*^{-/-} macrophages have abnormal cell morphology at the leading edge and are incapable of matrigel invasion, but do not show any defects in cell migration[16]. In the absence of Rho, proper membrane ruffling and lamellipodia formation occurs, but the macrophages are unable to move because of defects in actomyosin-mediated contractility that prevents detachment at the rear of the cell[17].

Once macrophages reach the site of inflammation, one of their primary jobs is to phagocytose invading pathogens and cellular debris, both for clearance and for antigen presentation to T-cells[18]. Phagocytosis is mediated by activation of either the Fc-receptor or the complement receptor (CR) on the macrophage cell surface. Fc-receptor activation by IgG opsonized particles is followed by actin mobilization that results in pseudopod extension and membrane ruffling around the particle to be internalized. In contrast, CR-mediated phagocytosis, does not require membrane protrusion and particle engulfment, but rather complement opsonized particles sink into the membrane[19] The actin dynamics required for Fc-receptor mediated internalization have been shown to be dependent on the activities of Rac and cdc42, but not Rho, while CR-mediated phagocytosis relies on Rho, but not Rac or cdc42[18-21].

Following phagocytosis, vacuoles of internalized particles (phagosomes) fuse with a series of endosomes and lysosomes in a process

called phagosome maturation. During this process, the internalized particles (e.g. dead or dying cells, bacteria, etc.) are digested so that they can be bound to MHC molecules and localized to the cell surface for presentation to T-cells. By performing FRET assays in *Rac1/2-/-* deficient macrophages, Wang *et. al.* showed that these GTPases are required for proper phagosome maturation after phagocytosis of opsonized bacteria, while another related study showed that Rho regulates phagosome maturation of apoptotic cells, but not opsonized targets[22-23]. Further studies in dendritic cells using dominant negative and constitutively active forms of Rho, Rac, and cdc42 showed that both Rho and cdc42 are involved in antigen presentation processes [24]. Collectively, it can be concluded that Rho family members, through the activation and repression of actin binding proteins, are essential for regulating much of the actin dynamics required for macrophage recruitment and response.

Regulation of transcription by the Serum Response Factor

In addition to controlling cell function through the modification of actin binding proteins, activation of Rho family members can also indirectly result in an increase in transcriptional activity through the DNA-binding protein, serum response factor (SRF)[25-40]. SRF is highly evolutionarily conserved and has been found in most, if not all, animal, plant, and fungus species[3]. It is also a founding member of the MADS (MCM1, Agamous, Deficiens, SRF) box family of transcription factors. These factors contain a highly conserved, N-terminal,

MADS box domain that mediates DNA binding, dimerization, and protein-protein interactions and many of these proteins have been shown to have important roles in both cell-type and non-cell type specific expression of immediate early, growth and development genes[41].

SRF was originally cloned through its binding to a specific sequence in the *c-fos* promoter that was required for gene induction in response to serum stimulation[42-43]. This element (AGGATGTCCATATTAGGACATCT) was termed the serum response element (SRE)[44]. In independent studies, Minty and Kedes identified an evolutionarily conserved sequence in the human, mouse, and chicken *cardiac α -actin* promoters that they termed the CCArGG (or CArG) box (CC(A/T)₆GG)[45]. It was later described that the CArG sequence forms the core of the *c-fos* SRE and that this motif is conserved in the promoters of additional muscle-specific, contractile and growth-related genes (e.g. *α -myosin heavy chain*, *cardiac* and *skeletal myosin light chain 2*, *cardiac troponin*, *JunB*, *Egr1*)[45-48].

SRF binds as a homodimer to the CArG box[49]. *In vitro* binding assay and crystal structure analyses of the SRF/CArG interaction suggests that very little deviation in CArG sequence is allowed for sufficient SRF binding. The 5'-CC and 3'-GG direct the binding of each SRF monomer, so the six base pair spacing between the two is essential to maintain the appropriate conformation for homodimerization. Due to these spatial constraints, there is also very little flexibility in the nucleotide composition in the A/T core region[50-53]. Such a

well defined binding site suggests that SRF is a potent transcription factor whose binding must be strictly regulated in order to maintain proper target gene transcription.

Activation of SRF by the Ternary Complex Factors

Although the CArG box is essential for SRF binding, the genomic context also seems to be important for SRF regulation of target genes. Chang *et. al.* showed that swapping of the flanking sequences (15 nucleotides), but not the CArG boxes, of cell type and non-cell type specific promoters (e.g. *SM22 α* and *c-fos*, respectively) results in gene expression patterns corresponding to the flanking sequences used[54]. Consistent with the requirement for CArG-flanking sequences, SRF has been shown to interact in complex with other DNA-binding transcription factors to mediate cardiac and smooth muscle cell specific transcription. For example, SRF was shown to complex with the muscle-cell specific factor, MyoD, and the general transcription factor SP1 to regulate expression of the *cardiac α -actin* gene[55]. Additional complexes have been observed between SRF and other factors such as GATA-6, GATA-4, Nkx-3.2, Nkx2-5, Nkx2-3, Barx1b, and MEF2 on the promoters of smooth muscle genes[56-59].

Despite the large number of proteins that have been identified as SRF interactors on target gene promoters, the best studied of these cofactors are the ternary complex factors (TCFs) (namely, ELK1, SAP1, Net). These proteins are members of the Ets family of transcription factors that represent a

large family of proteins with a conserved Ets DNA binding domain. In general, Ets factors regulate genes involved in differentiation, development, proliferation, and transformation[60].

TCFs bind to an Ets motif (GGA(AT)) that is found adjacent to the CArG motif (collectively called the SRE) in the promoters of immediate early/growth genes[2, 44, 61-62]. Surprisingly, even though Ets factors bind directly to SRF, the actual distance between the CArG box and the Ets motif can vary significantly. Gel shift studies by Treisman, *et. al.* showed that SRF and Ets factors (Elk1 or Sap1), can form ternary complexes when the Ets site is as far as 27 base pairs away from the start of the CArG motif[61]. This study also went on to show that in most cases, the orientation of the Ets site can be inverted (TTCC instead of GGAA) and still maintain the ability to form ternary complexes, which attests to the high flexibility of the SRF/Ets interaction.

In addition to direct DNA binding, TCFs cooperatively bind to the DNA binding domain of SRF through their B-box domain and are targets of the MAP kinase signaling cascades[63-65]. Phosphorylation at the c-terminus leads to a conformational change that increases TCF's affinity for DNA and also facilitates the recruitment of coactivators, such as CREB binding protein/p300, steroid receptor coactivator 1 and activating signal cointegrator 1, to activate SRF dependent transcription of target genes following growth factor stimulation[66-68].

Activation of SRF by the myocardin family of coactivators

During further characterization of the *c-fos* SRE, Hill *et. al.*, discovered that serum induced, Rho-dependent activation of *c-fos* expression was dependent on the SRF motif, but not on the associated Ets motif[40]. This result was somewhat surprising considering previous work showing the requirement of the Ets motif for activation of *c-fos* downstream of growth factor signaling[69]. Subsequent studies of gene expression in muscle cells have resulted in the discovery of a dichotomy in SRF gene regulation. Using specific inhibitors of the ERK-MAPK (U0126) and the actin-Rho-ROCK (latrunculin B, cytochalasin D, Y27632) pathways in luciferase reporter assays, several groups have shown that proliferation and activation of growth gene promoters is dependent on the ERK-MAPK pathway, while activation of smooth muscle cell specific differentiation gene promoters occurs downstream of actin polymerization and the Rho-ROCK signaling pathway[25, 34, 70]. The activation of immediate early/growth genes was further shown to be dependent on the activation of TCFs in conjunction with SRF at target gene promoters, while the actin-Rho-Rock dependent genes have recently been shown to be under the control of the myocardin family (myocardin, MKL1, MKL2) of transcriptional coactivators[30, 71-73].

Myocardin was originally identified in a bioinformatic screen for novel, cardiac-restricted genes and was found to be a very potent regulator of *SM22* gene expression[74]. Megakaryoblastic leukemia 1 (MKL1/MRTF-A/MAL) was

originally cloned in a survey of predicted human cDNAs, but was later given its name through the analysis of a t(1;22) translocation event that occurs between MKL1 and RNA binding motif protein 15 (RBM15/OTT) that causes acute megakaryoblastic leukemia in infants[75-76]. Through homology searches of a genomic database using myocardin and MKL1, a third member of the myocardin family was identified and named MKL2[72]. Further tissue analysis confirmed that myocardin expression is restricted to the myocardium and vascular system, while expression of MKL1 and MKL2 are expressed in most cell types[30, 72-74, 77].

All three proteins of the Myocardin family have very homologous domain structures. The N-terminal RPEL domain is required for actin binding to MKL1/2, while the basic and glutamine-rich domains mediate myocardin/MKL1/2 binding to the DNA binding domain of SRF. Myocardin and MKL1/2 also harbor a SAP (Saf, Acinus, Pias) DNA binding domain, but in contrast to TCFs, a consensus binding sequence and/or independent DNA binding activity has never been described. Each family member is also able to homo- or heterodimerize with each other through a leucine zipper domain that is located just N-terminal of the transactivation domain[73, 78-79].

Despite the high structural homology and common gene targets of the myocardin and MKL proteins, they have been shown to have very different mechanisms of activation. In several studies, myocardin was shown to be strictly nuclear while MKL1 showed both nuclear and cytoplasmic localization

and MKL2 exhibited primarily cytoplasmic localization[30, 80]. Miralles *et. al.* suggested that differences in localization of the myocardin and MKL factors is due to a lack of conservation of the second RPEL motif in myocardin which prevents high affinity binding to actin[30]. Further work with MKL1 and MKL2 in NIH3T3 cells led to the novel model of MKL regulation described below.

In serum starved, quiescent cells, there is an abundance of free G-actin in the cytoplasm that binds to MKL through its RPEL domains and prevents it from entering the nucleus. After stimulation with serum or other growth factors and subsequent Rho activation, actin polymerization is induced, which depletes the free G-actin pool, and allows MKL to translocate to the nucleus, bind SRF, and activate transcription[27, 30, 81]. Once target gene activation is complete, MKL is exported back into the cytoplasm through the nuclear exporter Crm1 in another actin-dependent process that has yet to fully be elucidated[82].

Differential regulation of SRF by the TCFs and myocardin/MKL factors

As described earlier, SRF has been shown to regulate a diverse array of gene programs, namely growth, differentiation, and cell migration/adhesion. So how does SRF effectively control such a broad range of genes? The first clue to the mechanism of SRF differential activation was identified when it was found that SRF dependent transcription could be initiated downstream of both the Erk-MAPK and Rho GTPase pathways (Fig. 1). Further exploration of the signaling molecules downstream of these pathways led to the discovery that

the TCFs and myocardin/MKL1/2 were direct coactivators of SRF-mediated transcription. Because the TCFs are able to directly bind to a defined DNA sequence and myocardin/MKL factors are not, target genes could be directly grouped into those genes that contain SREs (TCF growth gene targets) and those genes that contain only CArG boxes (myocardin/MKL differentiation and cytoskeletal gene targets)[2]. In addition, interaction studies went on to prove that coactivator binding to SRF was mutually exclusive, meaning that TCF and myocardin/MKL directly compete for SRF complex formation, thus providing an additional mechanism of control for SRF target gene transcription[4, 83].

Animal models of SRF, TCF and myocardin/MKL function

Despite the ubiquitous expression of SRF, TCF and myocardin/MKL factors, target gene regulation and function has only been elucidated in a limited number of cell types. In most cases, this is because of the early lethality of the systemic knockout mouse generated, but in other cases, it is because either no phenotype has yet been observed or because only the obvious phenotypes have been studied so far. SRF $-/-$ mouse development fails at gastrulation and they do not form any mesodermal layer [84]. Similarly, myocardin and MKL2 knockout mice were also found to be embryonic lethal, but from defects in smooth muscle cell (SMC) differentiation and vascular development, respectively[85-87]. Surprisingly, MKL1 mice are viable, but they were shown to have defects in mammary myoepithelial cell development that results in premature involution and the inability to effectively lactate[88-

89]. Recent studies have also reported that MKL1 deficient mice have decreased numbers of platelets in the peripheral blood resulting from decreased numbers of mature megakaryocytes in the bone marrow (data not shown)[90]. However, the systemic consequences of decreased platelet counts in the MKL1 $-/-$ mouse has not been fully elucidated.

Mouse knockout models of the TCF factors have also been generated with a variety of results. Introduction of a hypomorphic allele of the TCF factor Net, is lethal due to asphyxiation caused by dilated lymphatic vessels which results in the build up of chyle in the thoracic cavity[91-92]. In the case of SAP-1, systemic knockout results in a decrease in thymocyte positive selection with concomitant decreases in CD4 and CD8 single positive cells and peripheral T-cell numbers[92]. Surprisingly, Cesari *et. al.*, generated Elk-1 $-/-$ mice and found that they are viable with no obvious defects, which the authors conclude is a result of compensation by other TCF factors[93].

Because of the early lethality of several of these mouse model systems, further study would require the use of conditional knockout strategies and target gene deletion using cell type specific CRE transgenes. So far, SRF is the only factor for which conditional knockout models have been developed[94]. The most extensive research on SRF has been done in muscle and neuronal cell systems, but recent publications have also identified key roles for SRF in keratinocytes, liver, and T- and B-cell development and function. Deletion of cardiac SRF using CRE recombinases under the control

of the *α-myosin heavy chain* or the *Nkx2.5* gene promoters resulted in embryonic lethality due to improper chamber formation and abnormal beating myocyte development, respectively[95-96]. Conditional knockout studies have also shown that SRF expression in the liver, skin, and skeletal muscle are important for survival in developmental models[97-99]. Interestingly, post-natal models of SRF deletion in both liver and skin resulted in viable mice, but had defects in regeneration and cell adhesion, respectively[97, 100].

Additional studies in neurons, T-cells and B-cells showed that loss of SRF is not lethal, but that SRF deficiency in neurons resulted in cells with distinct defects in synaptic activity induced gene expression and migration, while targeted lymphocyte deletion resulted in the loss of single positive thymocytes and peripheral T-cells as well as the loss of marginal zone B-cells[101-104].

As described above, development of SRF^{-/-} mice fails at gastrulation and they do not form any mesodermal layer[84]. Interestingly, SRF^{-/-} ES cells are able to grow and differentiate *in vitro*, but they show defects in cell spreading, adhesion, and migration, indicating that the early lethality of the knockout may actually be due to the improper migration of cells of the developing tissues, rather than defects in proper cell proliferation and differentiation[105-107]. Gene expression analysis using conditional knockout models of SRF in different muscle subtypes, keratinocytes, and neurons in mice has confirmed the role of SRF in regulating cytoskeletal and contractile genes that are essential for proper development, targeting, and function of

these tissues[96-97, 99, 102]. The role of SRF in the regulation of cytoskeletal factors is also highly conserved because defects in migration and proper cell targeting and morphology have been seen in inactivation and knockdown models of SRF in *Dictyostelium*, *Caenorhabditis elegans*, and *Drosophila melanogaster* [108-111]. SRF-dependent expression of these cytoskeletal/contractile genes seems to be mediated by the myocardin/MKL family of coactivators because knockout and conditional inactivation models of these factors yield migration, targeting, and contractile phenotypes similar to models of SRF loss of function[99, 112].

Considering the fact that macrophage function is highly dependent on proper regulation of the actin cytoskeleton and that the MKL family of co-activators are capable of conveying information regarding the status of the actin cytoskeleton to the nucleus, we decided to explore the roles of MKL1/2 and SRF in the regulation of macrophage gene expression. Here, we demonstrate that MKL family members and SRF function in a coordinate manner to regulate the expression of both general and hematopoietic-specific genes involved in diverse biological processes, including the structure and function of the macrophage cytoskeleton. Using genome-wide location analysis, we find that SRF binding to hematopoietic specific genes is not restricted to the proximal promoter region, but occurs at distal sites in close proximity to the macrophage and B-cell specific transcription factor, PU.1. *PU.1* expression is required for proper macrophage and B-cell development

and current studies in the lab suggest that PU.1 mediates cell type specific expression by binding to distal enhancer sites[113-114]. The functional studies presented here further demonstrate that PU.1 is required for the activation of SRF-dependent, hematopoietic-specific, cytoskeletal target genes, providing insights into how cell-specific programs of SRF-dependent gene expression are achieved.

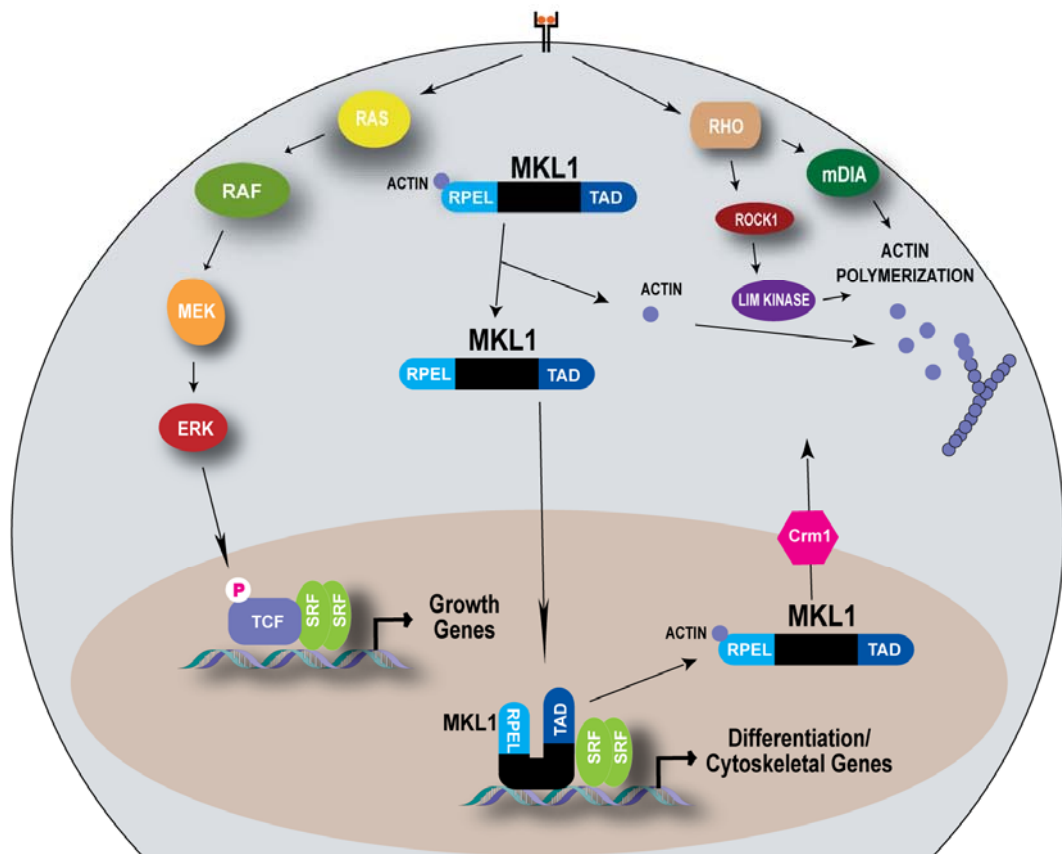


Figure 1 Mechanisms of SRF transcriptional regulation. There are two primary pathways that can be activated in response to extracellular stimuli (e.g. growth factors, cytokines, etc.) that can result in activation of SRF-dependent transcription. Growth genes are generally activated downstream of the Ras-MAP kinase pathway through interaction with the TCF family of coactivators. Activation of differentiation and cytoskeletal genes occurs downstream of Rho-GTPase dependent actin polymerization, which results in the translocation of cytoplasmic MKL to the nucleus and allows for MKL/SRF interaction on target genes.

Chapter 2

SRF and MKL1/2 Regulation of Macrophage Gene Expression

Introduction

Cells of the monocyte/macrophage lineage are key regulators and effectors of acquired and innate immune responses and participate in diverse aspects of tissue homeostasis[1, 115-117]. These roles require the acquisition of both general and specialized functions of the actin cytoskeleton that are necessary for directed migration, adhesion, phagocytosis and antigen-presentation. Numerous studies have shown that many of the cytoskeletal rearrangements relating to cell motility, adhesion and phagocytosis are regulated by the Rho family of GTPases[12, 14, 19-20, 118]. During leukocyte migration, lamellipodium extension at the leading edge, as well as receptor clustering and contraction at the trailing edge are induced by a complex balance of actin polymerization and branching mediated by signaling through RhoA, Rac, and cdc42 [119]. Rac and cdc42 have also been shown to regulate the actin rearrangements related to phagocytic uptake via the Fc-Receptor, while RhoA-dependent mechanisms are required for complement receptor-mediated internalization of opsonized targets[120].

In addition to broadly expressed components of the cytoskeleton, such as actin itself, macrophages utilize a number of cell-restricted factors to enable or regulate specialized aspects of cytoskeleton-dependent processes. For example, *coronin 1a*, is specifically expressed in cells of the hematopoietic lineage and functions as a negative regulator of phagosome/lysosome fusion in macrophages[10]. The mechanisms that enable the coordinated expression

of genes required for both the general and specialized functions of the macrophage cytoskeleton remain poorly understood.

Based on its known function in other cell types, the Serum Response Factor (SRF) transcription factor is likely to play important roles in regulating the expression of cytoskeletal genes in macrophages. Although deletion of the SRF gene in mice results in embryonic lethality at the time of gastrulation[84], SRF^{-/-} embryonic stem cells are able to grow and differentiate *in vitro*, but show defects in cell spreading, adhesion, and migration [105-107]. Conditional knockouts of SRF in muscle and brain of mice have confirmed roles of SRF in regulating cytoskeletal and contractile genes that are essential for proper development and function of these tissues[95-96, 99, 102, 104]. The roles of SRF in the regulation of cytoskeletal factors are also highly conserved because defects in migration and proper cell targeting and morphology have been seen in inactivation and knockdown models of SRF in *Dictyostelium*, *C. elegans*, and *Drosophila melanogaster* [108-111]. These results are consistent with a recent genomic analysis that estimates that nearly half of the SRF target genes are related to the cytoskeleton and contractile apparatus[3].

SRF is a member of the MADS (MCM1, Agamous, Deficiens, SRF) box family of transcription factors that contain a highly conserved, N-terminal, MADS box domain that mediates DNA binding, dimerization, and protein-protein interactions[41]. SRF regulates gene expression by binding as a homodimer to a sequence termed the CArG box (CC(A/T)₆GG), which has

been primarily associated with growth (e.g. c-fos, junB) and muscle-specific differentiation and cytoskeletal genes (e.g. SM22 α , SM α -actin)[2, 49, 121]. The ability of SRF to regulate such diverse programs of gene expression is currently attributed to the ability of SRF to interact with different cofactors. The best studied of these cofactors are the ternary complex factors (TCFs) (ELK1, SAP1, Net), which are members of the Ets family of transcription factors. The TCFs bind to an Ets motif (GGA(A/T)) that is found adjacent to the CArG motif (collectively called the serum response element (SRE)) in the promoters of immediate early/growth genes[2, 44, 61-62]. In addition to direct DNA binding, TCFs interact with the DNA binding domain of SRF and can be phosphorylated by MAP kinases to facilitate the recruitment of coactivators, such as cyclic-AMP response element binding protein (CREB) binding protein (CBP), to activate SRF dependent transcription of target genes following growth factor stimulation[63-66].

More recently, SRF has also been shown to interact with members of the myocardin family of transcription factors (myocardin, MKL1/MRTF-A/MAL, MKL2/MRTF-B)[30, 72, 74, 78-79, 122]. Interestingly, the MKLs regulate gene expression downstream of RhoA/actin signaling[27, 30, 79]. This novel mechanism of regulation involves the sequestration of MKL in the cytoplasm of quiescent cells through its binding to free G-actin. After cellular stimulation and Rho activation, actin polymerization is induced, which depletes the free G-actin pool, and allows MKL to translocate to the nucleus, bind SRF, and

activate transcription[27, 30, 80-82, 123]. Interaction studies have shown that MKL does not bind to DNA, but does bind directly to SRF through its DNA binding domain, indicating that MKL and TCFs compete with each other for binding to SRF, thus providing a mechanism for regulation of differential programs of gene expression[30, 83]. Further work suggests that SRF-dependent expression of these cytoskeletal and contractile genes is mediated by the myocardin/MKL family of coactivators because knockout of MKL in *Drosophila* and inactivation in mouse skeletal muscle using a dominant negative MKL1 transgene yield phenotypes similar to those in models of SRF loss of function [99, 112].

The observation that members of the MKL family of co-activators are capable of conveying information regarding the status of the actin cytoskeleton to the nucleus, where this information is interpreted by SRF, led us to explore the roles of these proteins in the regulation of macrophage gene expression. Here, we demonstrate that MKL family members and SRF function, in a coordinate manner, to regulate the expression of both general and hematopoietic-specific genes involved in diverse biological processes, including the structure and function of the macrophage cytoskeleton. Using genome-wide location analysis, we find that SRF binding to hematopoietic specific genes is not restricted to the proximal promoter region, but also occurs at distal sites in close proximity to the macrophage and B-cell specific transcription factor, PU.1. Functional studies further demonstrate that PU.1 is

required for activation of these genes, providing insights into how cell-specific programs of SRF-dependent gene expression are achieved.

Materials and Methods

Cell Culture

Primary, thioglycollate elicited macrophages were isolated from 6-8 week old, male, C57BL/6 mice (Harlan) by peritoneal lavage, 3 days after intraperitoneal injection of thioglycollate broth. Macrophages were plated in growth medium containing DMEM (4.5g/L glucose) (Cellgro) with the addition of 10% heat-inactivated FBS (Hyclone) and 100U penicillin/streptomycin (Invitrogen). After 3 hours, cells were washed in PBS and fresh growth medium was replaced. PU.1 ^{-/-} and PUER cells were cultured as described (Singh paper) and differentiated with 100nM 4-hydroxy-tamoxifen (Sigma) for the indicated times.

siRNA Transfection

Primary macrophages were plated at 7.5×10^5 cells per well of a 24-well plate in growth medium without antibiotics overnight. Cells were transfected with non-specific (NS) control or SMARTpool siRNAs (Dharmacon) using Deliver X transfection reagent (Panomics) according to the manufacturer's instructions. siMKL1/2 samples were transfected with a 1:1 mix of MKL1 and MKL2 SMARTpool siRNAs. Cell samples for microarray analysis were harvested 48 hours post-transfection. Cell samples transfected with the PU.1 siRNA showed visible cytotoxicity at longer incubations (~48 hours), so

cells treated in parallel with the PU.1 siRNA were harvested at 30 hours when no cytotoxicity was evident.

Expression Array Profiling

Total RNA from primary, elicited macrophages was purified using the RNeasy kit (Qiagen) with the inclusion of the RNase-free DNase digestion step according to the manufacturer's instructions. Non-specific and siSRF transfected RNA samples were labeled with the TotalPrep RNA Labeling Kit (Ambion) and hybridized to a MouseRef-8 Expression Bead Chip (Illumina) according to manufacturer's instructions. Non-specific control and siMKL1/2 transfected RNA samples were amplified and labeled using the Quick Amp Labeling kit (Agilent) and hybridized to the mouse 44K Whole Genome Microarray (Agilent) according to the manufacturer's instructions. Slides were scanned according to manufacturer's specifications and quantified using BeadArray (Illumina) or Feature Extraction (Agilent) software. Genes were considered present if every replicate for at least one condition (i.e. NS, SRF, MKL1/2) were above the threshold value cutoff of 100 and 64 (based on the histograms of the expression values) for the Agilent and Illumina arrays, respectively. Target gene expression was considered to be significantly changed if the difference between the expression in the control and target siRNA treated samples was below a 0.01 false discovery rate (FDR).

Gene Ontology Analysis

Gene ontology analysis of microarray and was performed using the web-based DAVID Functional Annotation tool (<http://david.abcc.ncifcrf.gov/home.jsp>)[124-125]. Significantly changed genes were compared to background sets composed of the total complement of genes that were represented on the microarray. For the comparison of the SRF and MKL data sets, the background set was defined as the common set of genes that were represented on both microarrays. Gene ontology terms were considered significant if they had a p-value and Benjamini value less than 0.05

Quantitative, Real-Time PCR Analysis (qPCR)

cDNA was prepared from total RNA templates using Superscript III (Invitrogen) reverse transcriptase according to the manufacturer's instructions. Purified CHIP DNA samples were used directly for analysis. qPCR was performed using DNA template, 50ng of each primer, and 2X SYBR-Greener Master Mix (Invitrogen) in 10 μ l reactions on a Step One Plus or 7300 Real Time PCR system (both from Applied Biosystems). cDNA primers were designed using Primer3 and are as follows: GAPDH F- 5'-aatgtgtccgctcgtggatct-3' R-5'-catcgaagggtggaagagtgg-3', SRF F-5'-tggcaccagtgtctgctagt-3' R-5'- acatgaatggcctgcaca-3', PU.1 F-5'-cctcagtcaccagggttctctac-3' R-5'-catcagcttctccatcagacac-3', LSP1 F-5'-gatgcgaggaacaggaagag-3' R-5'-aggctgatgagtgtctgctg-3', Thbs1 F-5'-caggattcactggctcacag-3' R-5'-ttgcactcacagcggtacat-3', Flt1 F-5'-

ctcagacaagtcacaaacctggag-3' R-5'-gggaactcatctgggtccataa-3', Coro1a F-5'-
cagcgtggatgggctacat-3' R-5'-ccgactttctaggcactgtcat-3', Lcp1 F-5'-
agctaaattctccctggttg-3' R-5'-ccttctgtccacctccgata, Nrp1 F-5'-
aaccttggtggaattgctgt-3' R-5'-cctggagatgttcttgcacc, Becn1 F-5'-
agttgccgttatactgttctgg-3' R-5'-ttctccacgtccatcctgta, Pira3 F-5'-
tctccatgagtgctgacc-3' R-5'-ttctccgttggtgatcctg-3'.. ChIP primers were made
to amplify the peak identified by ChIP-seq and are as follows: LSP1 F-5'-
gtgtgtgagcgagcctaa-3' R-5'-aattgcctctcgttcagat-3', Coro1a F-5'-
tctgagcctgctgttctca-3' R-5'-cttcaacccgacaaccactt, Lcp1 F-5'-
ggaagtggaaggtggttct-3' R-5'-gtggtcacaaggcaggaagt-3'.

Chromatin Immunoprecipitation (ChIP) Assay

ChIP assays were performed as previously described[113]. Briefly, 20 million cells were fixed by adding formaldehyde to the cell growth medium to a final concentration of 1%. After a 10 minute incubation at room temperature, the reaction was quenched with glycine (125mM final) for 10 minutes. Cells were scraped and washed three times in cold PBS. Cell pellets were resuspended in swelling buffer (10mM HEPES/KOH, pH 7.9, 85mM KCl, 0.5% NP-40, 1mM EDTA, 1X Complete protease inhibitors (Roche), 1mM PMSF, 5uM E64d (Biomol)) and incubated on ice for 5 minutes. Cells were pelleted and resuspended in 500µl nuclear lysis buffer (50mM Tris/HCl, pH 7.4, 1% SDS, 0.5% Empigen BB, 10mM EDTA, 1X Complete protease inhibitors, 1mM PMSF, 5uM E64). Cell suspensions were sonicated on wet ice 6 times 10

seconds using 13 watts of output power on a Misonix 3000 sonicator. Debris was pelleted by centrifuging samples for 10 minutes at 18,000 x g. Cleared supernatant was recovered and diluted 2.5 times with dilution buffer (20mM Tris/HCl, pH 7.4, 100mM NaCl, 2mM EDTA, 0.5% Triton X-100, 1X Complete protease inhibitors, 1mM PMSF, 5uM E64). Samples were precleared with 50uL blocked CL4B sepharose (washed twice with TE, blocked with 10ug/mL glycogen, 0.5% BSA for at least 1h, washed twice with TE and resuspended in TE to 50% slurry) for 2h at 4°C. The supernatant was recovered and 5% was kept as an input sample. Immunoprecipitations were performed using 2.5ug of antibody at 4°C, overnight, with rotation. Antibody complexes were recovered using 50ul of Immunopure protein A Agarose (Pierce) (blocked as above, overnight at 4°C). After 1h beads were pelleted and transferred to 0.45um filter cartridges (Millipore) using wash buffer I (WBI) (20mM Tris/HCl, pH 7.4, 150mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA, 1X Complete protease inhibitors, 1mM PMSF, 5uM E64). The cartridges were spun for 2min at 2,200 x g at 4°C to remove buffer. The beads were washed one more time with WBI, followed by 2 washes each with WBII (20mM Tris/HCl, pH 7.4, 500mM NaCl, 1% Triton X-100, 2mM EDTA, 1X Complete protease inhibitors, 1mM PMSF, 5uM E64), WBIII (10mM Tris/HCl, pH 7.4, 250mM LiCl, 1% NP-40, 1% Na-Deoxycholate, 1mM EDTA, 1X Complete protease inhibitors, 1mM PMSF, 5uM E64) and TE. Chromatin was eluted twice with 100uL elution buffer (0.1M NaHCO₃, 1% SDS) for 20 min and 10 min, respectively, and

incubated overnight at 65°C to reverse the crosslinks. Samples for were then incubated with RNase A (0.33mg/mL) for 2 hours, followed by Proteinase K (0.5mg/mL) for an additional 2 hours. Chromatin was purified using Qiaquick PCR purification kit (Qiagen) according to the manufacturer's instructions. SRF (sc-335), PU.1 (sc-352), C/EBP β (sc-150) and control antibodies (sc-2027) were purchased from Santa Cruz Biotechnology. Antibodies against H3K4me1 (ab8895) and H3K4me3 (ab8580) were from Abcam.

High Throughput Sequencing (ChIP-seq)

Purified ChIP DNA (10-50ng) was adapter ligated and PCR amplified according to manufacturer's instructions (Illumina). Amplified fragments were then run for 36 cycles on an Illumina Genome Analyzer according to the manufacturer's instructions. Sequence tags returned by the Illumina Pipeline (first 23-25bp) were mapped to the mouse genome using the mm8 assembly (NCBI build 36). Only those tags that mapped uniquely to the genome were considered for further analysis. Peaks were visualized by preparing custom tracks for the UCSC Genome Brower (<http://www.genome.ucsc.edu/>) in a manner similar to that previously described[126].

Identification and Annotation of ChIP-seq Peaks

ChIP-seq peaks were identified using the HOMER software suite (<http://biowhat.ucsd.edu/homer/>), which was developed by our lab to facilitate ChIP-seq analysis, using the method that has been described previously[113, 127]. Briefly, the position of ChIP-sequencing tags was adjusted 3' of its

mapped position by 75bp, which corresponds to half of the recommended fragment length for Illumina sequencing. To eliminate clonally amplified peaks, only one tag from each unique position was considered for analysis. Peaks were identified by searching for clusters of tags within a 200bp sliding window, while requiring that adjacent clusters be at least 1kb away from each other. The threshold for valid peaks was determined to be the value at which the false discovery rate was 0.001, as determined by peak finding using randomized tag positions. For comparison, all experiments were normalized to represent 5 million total mapped tags. To increase confidence in our analysis, peaks were considered significant if they had more than 2 times the threshold number of peaks. Also, for the comparison of peaks from multiple experiments the genomic location of the peak is defined as the average of the genomic positions of all of the peaks located within 100bp of each other. For each combined peak position, the tag count for each experiment is the sum of the tags located within 200bp of the center of the combined peak. SRF/PU.1 peaks were defined as those peaks where the center of the PU.1 peak is located within 100bp of the center of the SRF peak. H3K4me1 enhancer regions were defined and analyzed as previously described[113]. Peaks were associated with genes by identifying the nearest RefSeq TSS. Peak genomic location annotation was determined using the locations of RefSeq genes. For the annotation analysis, the promoters are defined as -500bp to +500bp relative to the TSS.

HOMER *de novo* Motif Analysis

HOMER was used to identify enriched sequence motifs in ChIP-seq peaks as previously described[113, 127]. Briefly, peak sequences (+/- 200bp from the center of the peak) were compared to 50,000 random, genomic sequences that were generated to match the size and GC-content of the peak sequences (to remove the sequence bias introduced by CpG islands). Motifs of 8, 10, and 12 base pairs were identified by screening all oligos for enrichment in the target set compared to the background set, with an allowance of 2bp mismatches to increase the sensitivity of the method. The top 200 oligonucleotides of each length with the lowest P-values were then converted into probability matrices and heuristically optimized to maximize hypergeometric enrichment of each motif in the given data set. As optimized motifs were found they were removed from the data set to facilitate the identification of additional motifs. Sequence logos were generated using WebLOGO (<http://weblogo.berkeley.edu/>).

Immunoprecipitation and Western Blot Analysis

siRNA transfected, primary, thioglycollate-elicited macrophages were rinsed with cold PBS and lysed 48 hours post-transfection as follows. MKL1/2 and control samples were lysed with equal volumes of 100°C, 3X SDS sample buffer (187.5mM Tris, pH 6.8, 12% w/v SDS, 30% glycerol, 150mM DTT, 0.06% bromophenol blue) supplemented with 50 μ M MG132 (Biomol) and 50 μ M E64d (Biomol). The addition of MG132 and E64d to these extracts is

essential due to the high susceptibility of MKL to degradation by cathepsins. Equal volumes of each sample were used for SDS-PAGE and western blot analysis.

NS and SRF siRNA treated samples were lysed with 1mL cold RIPA buffer (40mM Tris, pH7.4, 150mM NaCl, 1% Triton X-100, 0.2% SDS, 0.5% Na-deoxycholate, 1X complete protease inhibitors, 1mM PMSF, 5 μ M E64, 5 μ M MG132). After a 10 minute incubation on ice, samples were cleared at 18,000xg for 10 minutes at 4°C. The supernatant was recovered and the protein concentration was measured using a Protein Assay Reagent (Biorad). 160 μ g of protein from each sample was incubated with SRF antibody (1 μ g) overnight at 4°C. Each sample was incubated for 1 hour with 20 μ l of Immunopure Protein A Agarose before washing 3 times with cold, RIPA buffer. After the last wash, the beads were resuspended in 3X SDS sample buffer. All protein samples were run on 4-12% bis-tris SDS-PAGE gel (Invitrogen), and transferred to a PVDF membrane (Immobilon-P, Millipore). Western blots were performed by blocking with 5% milk in PBS + 0.1% Tx-100 (PBST), washing 3 times with PBST, and incubating with antibodies against SRF (sc-335, Santa Cruz Biotechnology), MKL2 (sc-47282, Santa Cruz Biotechnology), actin (A4700, Sigma), or MKL1 antibody (see below) diluted 1:1000 (1:500 for MKL2) in 5% BSA in PBST for 2 hours. The membrane was washed 3 times in PBST and then incubated with the appropriate HRP or alkaline phosphatase-conjugated secondary antibodies, diluted 1:4000 in 5% milk in

PBST, for 1 hour followed by three more washes in PBST. For alkaline phosphatase detection, membranes were incubated in assay buffer (20mM Tris, pH 9.8, 1mM MgCl₂) for 2 minutes, then with 1X CDP-star substrate (Applied Biosystems) for 2 minutes before exposure to film (Phenix Research Products). For HRP detection, membranes were incubated with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) for 2 minutes before exposure to film. For the immunoprecipitated samples, 3% of input from each sample was subjected to SDS-PAGE and transferred to PVDF as described above. Total protein was visualized using Ponceau S Stain (Sigma).

Generation of MKL1-specific antibodies

Amino acids 449-609 of mouse MKL1 (MRTF-A NCBI Accession # AF532597) were PCR cloned in frame with the Histidine tag of the pET28a vector (Novagen) using the Hind III and Xho I restriction enzyme sites. The construct was transformed into BL21 (DE3) ecoli, grown in a large scale culture to an OD₆₀₀ of 0.4-0.7 and induced for 3 hours with 1mM IPTG. Bacterial pellets were resuspended in Buffer A, pH 7.6 (50mM Sodium Phosphate, 5% glycerol, 300mM NaCl, 2mM b-mercaptoethanol, 0.1% Tx-100) + 1mM PMSF and sonicated 35 x 1 second on power 6 (Branson Sonifier 250). The sample was clarified at 20,000xg for 20 minutes at 4°C and supernatant was recovered. The tagged protein was recovered by incubating with NiNTA agarose (Qiagen) for 4 hours at 4°C with rotation. The beads

were washed with 35mLs of Buffer A + 20mM imidazole and eluted with increasing concentrations of imidazole (50mM, 75mM, 100mM, 125mM, 200mM, 300mM, 500mM). Elutions containing the most intact protein were pooled and dialyzed against PBS. Guinea pigs were injected with 0.5mg of purified His-tagged protein and sacrificed 7 weeks later to recover serum. The serum for western blot analysis was cleaned up using ammonium sulfate precipitation. 0.5mL of saturated $(\text{NH}_4)_2\text{SO}_4$ was added dropwise to 1mL of serum and incubated on ice for 4 hours. After centrifuging at 3000 x g for 30 minutes at 4°C, the supernatant was recovered and 0.6mL of $(\text{NH}_4)_2\text{SO}_4$ was used to precipitate the antibody overnight at 4°C. The sample was centrifuged as before and the antibody pellet was recovered and resuspended in 1mL of PBS, followed by 2 rounds of dialysis against PBS.

Results

SRF and MKL1/2 regulate overlapping programs of gene expression in macrophages

We initially evaluated the expression of SRF and MKL family members in primary, thioglycollate-elicited and bone marrow-derived macrophages. Microarray experiments and optimized quantitative PCR (qPCR) assays demonstrated the presence of *SRF*, *MKL1* and *MKL2* transcripts in both types of macrophages, while *myocardin* (cardiac-restricted) mRNA was not expressed (Fig. 2 and data not shown). To identify MKL1/2 (MKL) regulated genes in macrophages, we performed expression array analysis of RNA from

primary, thioglycollate-elicited macrophages that were transfected with either a combination of siRNAs targeted against MKL1 and MKL2 or a non-specific (NS) control siRNA. MKL siRNA transfection resulted in significant reductions in both *MKL1* and *MKL2* mRNAs (~83% and ~77%, respectively), as well as an almost complete loss of these factors at the protein level (Fig. 2A and B). Expression array results showed that knockdown of MKL led to the alteration of 2263 basally transcribed genes, compared to the NS-treated control, using a false discovery rate (FDR) cutoff of 0.01. The genes that are most significantly decreased after MKL1/2 knockdown are shown in Table 1. Gene ontology (GO) analysis was performed on this gene list using the web-based DAVID Functional Annotation Tool to search for enrichment of specific functional annotations associated with the genes whose expression were affected by knockdown of MKL[124-125]. This analysis resulted in the enrichment of GO terms for apoptosis, cytoskeleton and organelle organization, as well as cell development and small GTPase signal transduction (Fig. 3A, full results in Table 2).

Because the MKL target genes should only represent a subset of all SRF target genes, we performed an independent expression array analysis to identify SRF-dependent genes in macrophages. siRNA knockdown of SRF in primary macrophages resulted in an approximately 65% reduction in *SRF* mRNA expression, and a significant reduction in SRF protein levels as determined by qPCR and immunoprecipitation followed by western blot

analysis, respectively.(Fig. 2C and D). Expression array analysis of these samples resulted in significant changes in the basal expression of 1093 target genes compared to NS control siRNA treated samples (FDR<0.01). Examples of the most downregulated genes identified in this analysis are listed in Table 3. GO analysis of this SRF-dependent set of genes resulted in the enrichment of terms for various metabolic processes as well as organelle, cell, and cytoskeleton organization and biogenesis (Fig. 3B, full results in Table 4).

To identify the subset of SRF-dependent genes that are also MKL-dependent, the genes that were expressed on both array platforms were compiled and evaluated for regulation by both SRF and MKL. Table 5 shows the top genes that are commonly downregulated by SRF and MKL. GO analysis of genes whose expressions were significantly altered (1216 genes) by both SRF and MKL knockdown (FDR<0.01) resulted in a combined list of the terms that were enriched in the individual analyses for MKL and SRF (Fig. 3C, full list in Table 6). qPCR validation of the expressions of several of the target genes identified by the microarray was also performed as shown in Figure 3D. *Thrombospondin (Thbs1)*, *fms-related tyrosine kinase-1 (Flt1)*, and *coronin1a (Coro1a)* all showed a strong dependence on both SRF and MKL for their expression. In contrast, *neuropilin (Nrp1)*, *beclin1 (Beclin1)*, and *paired Ig-like receptor 3 (Pira3)*, were found to be dependent on SRF, but not dependent on MKL, for their full expression (Fig. 3D).

Recent work has shown that in addition to regulation of SRF, MKL coactivators are also able to regulate activation of the SMAD family of proteins. Based on these findings we hypothesized that we would identify an SRF-independent subset of MKL-dependent target genes. Although such a subset was suggested by the microarray analysis, we were unable to validate these targets because qPCR analysis of several independent siRNA experiments showed that these targets were also SRF dependent (data not shown). We believe that this inconsistency is probably due to differing sensitivities of the gene probes on each of the microarray platforms used in these studies. Because SRF-independent, MKL-dependent target genes are not readily obvious in our analyses, we conclude that MKL1/2 appear to primarily function in an SRF-dependent manner in resting macrophages.

Identification of direct gene targets of SRF in macrophages

De novo motif analysis of the promoters of the MKL and SRF dependent target genes identified by the microarray experiments failed to detect any significant DNA sequence enrichment in the target gene promoters for CArG boxes or any other factor motifs (data not shown). These results suggest that many of the target genes identified in the microarray analysis are either secondary targets of SRF and MKL, or that the critical regulatory regions of these genes are not located in promoters. In order to identify the direct targets of SRF, we performed chromatin immunoprecipitation (ChIP) for SRF in primary macrophages followed by high throughput sequencing of the

enriched DNA fragments (ChIP-seq). Once the SRF ChIP-seq peaks were identified and mapped to the genome, they were associated with genes based on the nearest transcription start site (TSS). Analysis of the recovered sequence tags resulted in the identification of 2480 significant (greater than two times the threshold number of tags) peaks, or regions of specific SRF binding, across the mouse genome. Table 7 lists those SRF peaks with the highest tag counts in this experiment. An example of the SRF ChIP-seq data for chromosome 3, compared to input control, is shown in Figure 4A as a University of California, Santa Cruz (UCSC) Genome Browser image[128]. Attempts were also made to perform ChIP-seq for MKL1, but the enrichment failed to meet the standards required for sequencing analysis.

In order to determine whether SRF has any location preference in binding to the genome, we annotated the genomic positions of significant SRF peaks (>12tags). These peaks were classified (based on the locations of genes identified by RefSeq) to be located in promoters (-500bp to + 500bp relative to the TSS), exons, introns, or intergenic regions (sequences excluding promoters, exons, and introns). The highest percentage of SRF peaks were found to be located in intergenic regions (42.3%), while a smaller percentage of peaks were found in promoters (26.8%) and introns (29.2%). SRF peaks were rarely identified in exons (1.7%) (Fig. 4B).

De novo motif analysis of the total set of SRF peak sequences was performed to determine the factors that may be functionally cooperating with

SRF in macrophages. Surprisingly, the most enriched sequence identified was the motif for the macrophage and B-cell specific Ets factor, PU.1 (log p = -1241), while the second most enriched motif was for SRF (CArG box) (log p = -878) (Fig. 4C). Significant association with motifs for the ubiquitous transcription factors SP1 and CREB were also identified (log p = -202 and -175, respectively) (Fig. 4C). In addition, a composite PU.1/IRF motif was identified, consistent with the ability of PU.1 to form ternary complexes with IRF factors on these elements[129-130].

In an effort to identify those genes that are direct targets of SRF, we compared the genes regulated by the SRF and MKL siRNAs to the SRF ChIP-seq data to see which genes were associated with SRF peaks (Fig. 4D and E). We observed that 25-28% of the genes that were downregulated by SRF siRNAs were also associated with SRF peaks and that most of these peaks were found at distal sites (exons, introns, or intergenic regions) (Fig. 4D). Overall, 28.4% (84 out of 295) of the SRF downregulated genes have peaks, most of which only have peaks located at distal sites (56 out of 84 or 66.6%), followed by those genes with peaks only in promoters (18 out of 84 or 21.4%) and those genes that have peaks at both promoter and distal sites (10 out of 84 or 12%). We performed similar analysis comparing those genes that were significantly downregulated by MKL siRNA treatment to the SRF ChIP-seq data and found that 44% of the genes that were most highly downregulated (>2 fold) were also associated with SRF peaks. Those genes that were

slightly less downregulated (1.99-1.3 fold) were only 20-25% associated with SRF peaks (Fig. 4E). When considering all of the MKL siRNA downregulated genes with SRF peaks (150 genes), the majority of those genes had peaks that were located in distal regions (101 genes, 67.3%), while genes with peaks in promoter regions (38 genes, 25.3%), and those with peaks at both promoter and distal sites (11 genes, 7.3%), were identified much less. The high propensity for target genes to be associated with distal SRF binding suggests a broader range of SRF transcriptional control than the proximal promoter regulation that has traditionally been associated with SRF function.

SRF binds primarily to distal regions in macrophages

To assess whether the location of SRF binding has any bearing on its association with other transcription factors, the significant SRF peak gene list was separated into those peaks associated with target gene promoters (-500bp to +500bp) and those peaks associated with distal sites (all areas, excluding promoters) for motif analysis. SRF peak sequences (+/- 200bp from the center of the peak) in promoter regions were shown to be highly enriched in CREB, SP1, GABP and NFY binding sites in addition to SRF binding sites (Fig. 5A) in primary macrophages. When the same *de novo* motif analysis was performed on those SRF peaks located distal to target genes, we found that these sequences were only enriched for PU.1 and SRF DNA sequence motifs (Fig. 5B).

In order to get a better idea of where SRF is binding in the primary macrophage genome (relative to the TSS) as well as to determine which of these peaks are associated with the transcription factor binding sites identified in the motif analyses, we compiled the peak location and motif data into scatter plots. In each scatter plot, the blue diamonds represent the location and tag count of each SRF peak (Fig. 5C-F). These plots qualitatively show that there is an increase in the number of strong SRF peaks (high tag count) as the distance from the TSS increases. (Fig. 5C-F, blue diamonds, compare the density of peaks with high tag counts (16 tags ($\log_2 4$) to 256 tags ($\log_2 8$)), located either closer or further than 1024 base pairs (bp) ($\log_2 10$) from the TSS). Quantitatively, there are approximately 2.6 fold fewer SRF peaks containing greater than 16 tags located within 1kb (+/- 1kb) of the TSS, than are located outside of this region (527 peaks compared to 1414 peaks, respectively).

The location of those peaks containing the most enriched DNA sequence motifs from the ChIP seq data set (Fig. 5A and B), were also included in the scatter plot analysis (Fig. 5C-F, pink squares). The most striking result is that despite the fact that SRF binding motifs are significantly enriched in the data set, relatively few peaks (23.8%) actually contain consensus SRF motifs (Fig. 5C). What is also clear from this scatter plot representation is that SRF and PU.1 binding motifs are most enriched in those peaks located greater than 1024bp ($\log_2 10$) from the TSS (Fig. 5C and D),

while the densest cluster of SP1 and CREB motifs are found in those peaks located less than 1024bp from the TSS (Fig. 5E and F). Collectively, this data suggests that multiple mechanisms of SRF gene regulation may exist in macrophages based on binding location and differential cofactor association.

Additional analysis was performed on the SRF peak sequences to determine if there is any spatial relationship between SRF binding (presumably the center of the peak) and the potential binding sites of the identified associated factors. These results showed that out of all of the SRF peak sequences, the SRF motifs are most frequently located within 25bp of the center of the peak (bin size = 25 base pairs), while the locations of associated factor binding sites (PU.1, SP1, and CREB) have a much broader distribution (100-200bp) from the center of the peak (Fig. 6). In order to evaluate the spatial location of associated motifs only in those peak sequences that also contain SRF motifs, we repeated the motif location analysis as in Figure 6 using peaks that were centered on the SRF motif (Fig. 7). These results show that the PU.1 motif can be found at any location within approximately -100bp to +100bp from the SRF motif. In slight contrast, SP1 and CREB motifs are located within the same window as PU.1 with relation to the SRF motif, but cannot be found right at the SRF motif location (Fig. 7). These results suggest that there are no strictly defined spatial location requirements for SRF association with any of the factors identified in the motif analysis, other than their co-enrichment in 100bp regions.

SRF associates with PU.1 at distal sites

To confirm that the SRF association with PU.1 motifs at distal sites also correlates with actual PU.1 binding at distal sites, we performed ChIP-seq analysis of PU.1 in primary macrophages. Figure 8A shows a scatter plot representation of the combined SRF and PU.1 ChIP seq peaks that were determined to be significantly SRF-positive (>12 tags, blue dots), PU.1-positive (>14tags, pink dots), and SRF-PU.1-positive (>12 and >14 tags, respectively, yellow dots). Because there was a high correlation of SRF sites that also have PU.1, we tested to see if SRF and PU.1 directly interact. Co-immunoprecipitation experiments failed to show direct binding of these two factors, which is consistent with the results of Watson *et. al.* and Shore and Sharrocks who also failed to observe ternary complex formation between SRF and PU.1[48, 64] (data not shown).

Recent data has shown that PU.1 located in distal regions of the genome is highly associated with chromatin modifications ascribed to transcriptional enhancers and is much less associated with chromatin marks associated with proximal promoters[113]. In order to determine the SRF associations with chromatin marks with and without PU.1, we analyzed the association of SRF with PU.1, the highly associated PU.1 factor, C/EBP, and histone 3-lysine4 (H3K4) promoter and enhancer methylation marks (trimethyl (me3) and monomethyl (me1), respectively) using heat map analysis of the ChIP seq data for each factor/mark in primary macrophages [113, 131-133].

White regions denote areas of low tag counts, while red shading denotes areas with high tag counts (>15tags) for the factor/histone mark indicated. After generating the heat map data for each SRF peak, the data was clustered in order to identify the different patterns of factors/marks that were associated with SRF peaks (Fig. 8B-D, full heat map in Fig. 9). Figure 8B represents those SRF peaks that are bound to promoter regions because there is a close proximal association of SRF in those peaks with the H3K4me3 mark, which has been shown to be highly associated with poised or active promoters[134-135]. This association also occurs in the absence of PU.1 and C/EBP, which have been described to be primarily localized at distal genomic regions[113]. In the case where PU.1 is bound with SRF, there is enrichment of the H3K4me1 enhancer mark directly adjacent to the location of factor binding, while the H3K4me3 promoter mark is relatively absent (Fig. 8C). Surprisingly, C/EBP binding was not observed to be associated with the SRF/PU.1 peaks, which is consistent with the absence of enriched C/EBP motifs in the distal peak motif analysis (Fig. 8C and 5B). Putative enhancer association of SRF also seems to be PU.1 dependent because in the absence of PU.1 binding, SRF peaks show very little defined association with the H3K4me1 enhancer mark (Fig. 8D).

In order to fully assess the requirement for PU.1 and SRF binding at distal enhancer sites, we used the tamoxifen-inducible PU.1 cell line, PUER, for CHIP-seq analysis. PUER cells are *PU.1* *-/-* hematopoietic stem cells that

have been transduced with PU.1 that has been fused to the ligand binding domain of the estrogen receptor[136]. In the absence of stimulation (when PU.1 is not active), PUER cells are in suspension and are similar to myeloid progenitor cells. After treatment with tamoxifen, PU.1 is activated which causes the cells adhere and differentiate to become more like macrophages[137]. To analyze the requirement of PU.1 for SRF binding, CHIP-seq was performed for SRF, PU.1, and H3K4me1 in untreated and tamoxifen treated PUER cells. The most significant peaks observed for SRF in undifferentiated and differentiated PUER cells are listed Tables 8 and 9, respectively, while the PU.1 and H3K4me1 data sets have been described previously[113]. As a control, motif analysis was performed on the SRF peak sequences recovered from each cell line to verify that the factor associations were similar to those observed in primary macrophages. This analysis showed that SRF motifs were enriched at both proximal and distal sites and that motifs for the same associated factors observed in primary macrophages were also enriched in these data sets (compare Fig. 5A and B with Fig. 10A and B). Importantly, distal SRF association with the PU.1 motif only occurred in PUER cells following 24h of tamoxifen treatment, which suggests that this system is a good model to test the dependence of SRF on PU.1 activation for recruitment to distal sites (Fig. 10A and B).

For a global view of the association of SRF, PU.1, and H3K4me1 at distal sites after PU.1 activation, heat maps were generated and clustered,

using the ChIP-seq data, as indicated, for those regions containing SRF peaks that acquire PU.1 after 24 hours of tamoxifen stimulation (Fig. 11A and B). For this analysis, distal peak sites (-3kb to +3kb relative to the TSS) were used to minimize the H3K4me1 signal associated with promoter regions (Fig. 8B) and each column of data was centered on the location of SRF peaks in PUER cells after 24 hours of stimulation as described for Figure 8B-D. In addition, the relative intensities of the H3K4me1 mark in the peak data represented by the heat maps was further quantified using histogram analysis. For the histograms in Figure 11C and D, the number of sequence tags mapped for the H3K4me1 mark were counted using 5bp bins and then graphed as a function of their location, relative to the center of the SRF distal peaks (excluding those peaks +/-3kb relative to the TSS) as indicated.

SRF peaks that were gained after 24 hours of stimulation (24h compared to 0h) that also gained PU.1, showed the largest increase in H3K4me1 signal directly adjacent to the SRF/PU.1 binding site and represent the largest subset of SRF peak locations that acquire PU.1 (Fig. 11A and C). An additional, but much smaller subset of locations where SRF is bound in both stimulated and unstimulated cells and PU.1 binding increases with stimulation showed no change in the H3K4me1 signal observed (Fig. 11B and D). Although there is no change in the H3K4me1 signal when only PU.1 and not SRF is acquired after tamoxifen treatment, the overall H3K4me1 signal is significantly lower than that observed for the genomic locations that acquire

both factors after treatment, thus representing a basal level of H3K4me1 signal (compare scales on Fig. 11C and D). When SRF is present in the absence of PU.1 in both treated and untreated cells, the H3K4me1 mark is present, also with low intensity (similar to Fig. 11D), and little to no change in signal is observed (data not shown). Collectively, this data suggests that significant acquisition of the H3K4me1 mark at SRF sites only occurs when SRF and PU.1 are both acquired in a differentiation dependent manner.

SRF and PU.1 regulate the expression of hematopoietic-specific cytoskeletal genes

Through the current data analysis, we discovered that several of the most significantly altered genes after knockdown of SRF or MKL in macrophages were cytoskeletal genes whose expression has only been reported in hematopoietic cells[11, 138-139]. Further analysis of the ChIP-seq data on several of these SRF target genes (*lymphocyte/leukocyte specific protein 1 (LSP1)*, *Coronin 1a (Coro1a)*, and *lymphocyte cytosolic protein 1 (Lcp1)*) showed that there are distal, but not promoter, SRF and PU.1 peaks associated with these genes (Fig. 12A). In order to confirm that the association between SRF and PU.1 is functional and not just correlative, the expression and regulation of these target genes was assessed. In the absence of PU.1 activation in PUER cells (0h), *LSP1*, *Coro1a*, and *Lcp1* mRNAs were detected at very low levels, but following macrophage differentiation induced by PU.1 activation with tamoxifen, transcript levels were dramatically upregulated as

shown by qPCR analysis (Fig. 12B). These results suggest that expression of these genes is PU.1 dependent.

To determine whether SRF and PU.1 are both capable of altering the expression levels of these cytoskeletal target genes in primary macrophages, siRNA knockdown experiments were performed. qPCR analysis of primary, thioglycollate-elicited macrophages transfected with either PU.1 or SRF siRNAs, showed decreased levels of *LSP1*, *Coro1a* and *Lcp1* mRNAs compared to cells transfected with a NS siRNA control (Fig 12C). This knockdown effect was specific because neither the SRF nor the PU.1 siRNAs significantly altered the expression of *PU.1* or *SRF*, respectively (Fig. 13). In addition to confirming the SRF dependence of *LSP1*, *Coro1a*, and *LCP1* expression in primary macrophages, we also confirmed their dependence on the MKL coactivators. Figure 12D shows that after knockdown of MKL1 and MKL2 in primary macrophages, mRNA expression of these hematopoietic-specific cytoskeletal gene targets is also diminished. Taken together, this data suggests that SRF and PU.1 functionally associate to regulate the expression of macrophage-specific, cytoskeletal target genes and that these genes are also targets of the MKL coactivators.

Discussion

SRF is one of the most well described transcription factors to date, but despite its ubiquitous expression, relatively few cell types have been used as model systems for studying SRF gene regulation. In this study, we extend the

current knowledge of SRF gene regulation by analyzing the role of SRF and its Rho-dependent coactivators, MKL1 and MKL2, in the control of macrophage gene expression programs. Of particular interest with regard to macrophage function, are the large number of cytoskeletal and cellular organization genes that are both SRF- and MKL-dependent as identified by the GO analysis of siRNA treated primary macrophages (Fig. 3, Fig 12). Given that ES cell, muscle and neural specific knockout of SRF results in major cytoskeletal defects such as improper cell spreading, localization and contractile function of the target cells, we hypothesize that targeted deletion of SRF or MKL in the macrophage would also result in major disruption of macrophage function, in particular, pathogen clearance and wound healing, and should be the focus of future study[95, 102, 104, 106].

Knockdown of MKL and SRF using siRNAs in primary macrophages, clearly confirmed that out of the SRF-dependent genes identified, only a subset of those genes are actually MKL dependent. Recent work has shown that in addition to SRF, MKL can also function as a coactivator of the SMAD family of transcription factors, so it was somewhat expected to also reveal a subset of genes that were MKL, but not SRF, dependent[140-141]. Based on the microarray studies, there were candidate genes for this subset, but attempts to confirm these targets using qPCR analysis showed that they were also SRF-dependent (data not shown). This discrepancy is presumably caused by differences between probes and the differing sensitivities of the

microarray platforms used for these studies. Based on these results, we believe that MKL is primarily functioning through SRF to regulate the expression of basally expressed genes and that a SMAD-dependent subset of MKL target genes may result only after extracellular stimulation.

Microarray analysis is informative in that it can provide a global view of how gene expression is changing in response to a stimulus (i.e. growth factors, siRNA knockdown, etc.), but it is limited because the measurements are only as good as the probes on the array and because it cannot make the distinction between direct, secondary, tertiary, etc. targets of the change. In order to identify the genome-wide, direct targets of SRF, CHIP-sequencing analysis was performed. Surprisingly, much of the SRF bound to the genome is actually outside of the proximal promoter region (~73%) (Fig. 4B). This result was unexpected, because many of the SRF target genes that have previously been described have SRF motifs (CArG boxes) in the proximal promoter region (i.e. *c-fos*, *SM22 α* , *SRF*, etc.)[44, 142-144]. Additionally, bioinformatic studies have suggested that SRF binding mainly occurs at proximal sites because almost all functional CArG boxes were found to be located within 4kb of the TSS[121, 145]. In contrast, the current CHIP-seq analysis shows that only 39% of significant SRF peaks are located in this 8kb region and out of those peaks, only 23.5% contained CArG boxes. Out of the remaining 61% of significant SRF peaks that are located outside of this 8kb region, 26.2% have SRF motifs (data not shown). Based on this data, we

hypothesize that SRF can regulate gene activation from not only proximal promoter sites, but also from extremely distal regions through both CArG-dependent and yet to be defined CArG-independent mechanisms.

In an effort to determine how many of the ChIP-seq peaks are also associated with genes that are SRF or MKL dependent, we compared the SRF ChIP-seq data to the SRF and MKL target genes identified in the microarray analysis. Based on our analysis, less than 30% (28.4% and 23.5%, respectively) of the SRF and MKL dependent genes were actually associated with SRF peaks. This low level of correlation could be due to off-target effects of the siRNAs used, or could be representative of secondary or tertiary effects on gene expression. SRF is known to regulate the expression of other transcription factors, which may, in turn, affect the expression of additional target genes (Table 10). Because the regulation of these additional targets occurs in an indirect manner, they would be SRF- or MKL-dependent, but would not be associated with SRF peaks.

A particularly surprising result of our ChIP-seq data analysis is the identification of a large number of SRF peaks that do not contain SRF motifs. One possibility for this result is that SRF may directly bind to non-CArG DNA sequences. To test this hypothesis, supershift analysis of SRF peak sequences that do not contain SRF motifs was performed, but these sequences failed to show any SRF complex formation, while SRF complex formation on the *β-actin* CArG was readily observed (data not shown). A

second possibility is that because the DNA/protein complexes are fixed for ChIP analysis, fragments of DNA that are not directly bound to the factor being studied, but are instead associated through protein/ DNA complex interactions, may still be enriched. The current data is consistent with this second model in that distal SRF binding is primarily associated with PU.1 as well as monomethylated H3K4, which has been described in other cell systems to be a mark for transcriptional enhancers. One of the ways that enhancers have been shown to exert their effects is through looping of the DNA to allow factors from distal regions to interact in complex with factors present at promoters[146]. This type of association may explain why most of the CArG containing peaks occur at distal sites while those peaks that do not contain SRF motifs are highly associated with motifs for proximal promoter factors, such as CREB, SP1, and NFY. In fact, it was found to be relatively rare that SRF and CREB, SP1, or NFY motifs actually occur within the same peak sequence, which further indicates an indirect association of SRF with non-SRF motif-containing sequences (data not shown). It is, however, still possible that the association of SRF with these non-CArG containing sites is a result of the non-specific association of some factor during the ChIP procedure. In this regard, it is important to note that SRF ChIP-seq data from Jurkat cells used to validate a ChIP-seq analysis program also resulted in the enrichment of SP1 sites, in addition to CArG boxes, in the SRF peak sequences. Taken together, we believe that it is highly possible that the significant enrichment of proximal

promoter sequences occurs because of a complex association between SRF bound to distal locations and transcription factors bound to proximal promoter sites.

Previous work in the lab has shown that the acquisition of PU.1 at distal sites during macrophage differentiation is highly associated with H3K4me1 at these sites[113]. Subsequently, several of these sequences were shown to have macrophage-specific enhancer function. Considering this data and the fact that a subset of SRF target genes overlap with PU.1 regulated programs, we propose here that SRF mediates two different types of transcriptional regulation. The first type is regulation of ubiquitously expressed genes, such as *β-actin* and *egr1*, through SRF motifs located in the proximal promoter regions of these genes. The second type of regulation is that of cell-type specific genes, such as the ones described here, which is mediated through distal association of cell-type specific factors, such as PU.1, at regions with enhancer activity (Fig. 12). In support of this model, SRF has previously been shown to bind to a distal enhancer region for the muscle cell specific transcription factor, *MyoD*, in smooth muscle cells, but in macrophages, *MyoD* is not expressed and SRF binding to this site was not observed[66]. In combination with the current work, additional studies exploring the genome-wide location of SRF in other cell types should result in a more complete understanding of how SRF can mediate broad programs of gene expression in both cell-type and non-cell-type specific manners.

Acknowledgements

I would like to thank Sven Heinz for the PU.1, C/EBP, H3K4me1, and H3K4me3 ChIP-seq data sets in primary macrophages and PUER cells. I would also like to thank Chris Benner for generation of the heat maps and for much consultation regarding the use of the HOMER software suite that he developed. Labeling and hybridization of the expression microarray samples, consolidation of the microarray data, and sequencing of the ChIP samples was performed by James Sprague, Roman Sasik, and Colleen Ludka, respectively, of the UCSD Biogem Core facility. I would also like to thank Lynn Bautista for her assistance with preparation of the figures.

Chapter 2, in its entirety, is currently being prepared for submission for publication of the material. Current co-authors include Sven Heinz, Chris Benner and Chris Glass. The dissertation author was the primary investigator and author of this material.

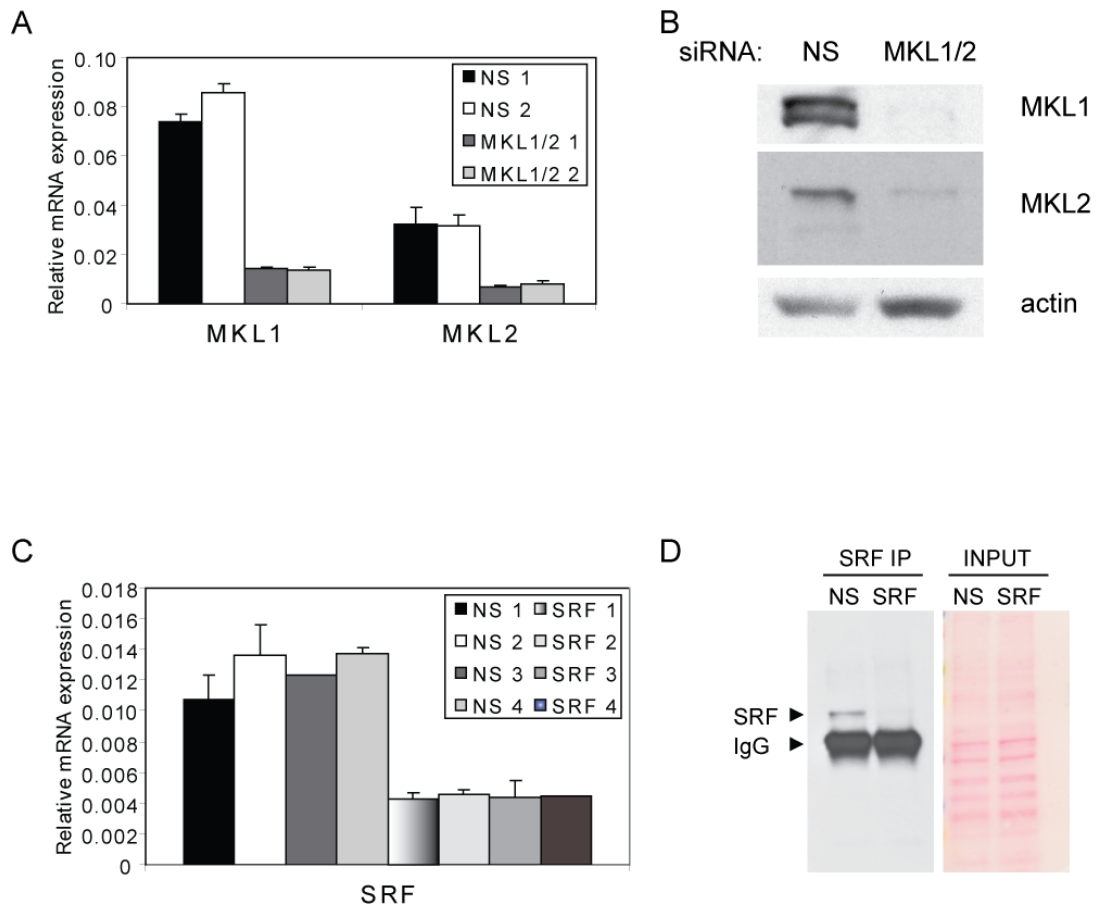


Figure 2 MKL1/2 and SRF knockdown in samples used for microarray analysis. (A) qPCR analysis of the relative mRNA expression of MKL1 and MKL2 (relative to GAPDH) in primary macrophages 48 hours post-transfection of non-specific (NS) control or MKL1/2 combined siRNAs. Each of the four samples that were used for the microarray analysis is shown. (B) Western blot analysis of MKL1 and MKL2 expression in primary macrophages 48 hours post-transfection of NS control or MKL1/2 combined siRNAs. Actin is shown as a protein loading control. (C) qPCR analysis of microarray samples from primary macrophages treated with NS and SRF siRNAs as in (A). (D) Western blot analysis of SRF immunoprecipitates in primary macrophages after treatment with NS control or SRF siRNAs for 48 hours (left). Input samples (3%) were also run on SDS-PAGE, transferred to PVDF and stained with Ponceau S to confirm that equal amounts of protein were used for the immunoprecipitation (right).

Table 1 Top genes downregulated by MKL1/2 knockdown. Microarray results showing the genes whose expressions are most decreased, compared to NS control, after treatment with MKL1/2 siRNAs in primary macrophages.

Accession #	Symbol	Description	fold decrease
NM_145581	Siglecf	sialic acid binding Ig-like lectin F	5.23
NM_011580	Thbs1	thrombospondin 1	4.60
NM_010228	Fit1	FMS-like tyrosine kinase 1	3.34
NM_176912	Gpr77	G protein-coupled receptor 77	3.16
BC050941	Mkl1	MKL (megakaryoblastic leukemia)/myocardin-like 1	3.09
NM_007748	Cox6a1	cytochrome c oxidase, subunit VI a, polypeptide 1	3.07
NM_009898	Coro1a	coronin, actin binding protein 1A	2.65
NM_019391	Lsp1	lymphocyte specific 1	2.55
NM_010050	Dio2	deiodinase, iodothyronine, type II	2.51
AK088706	E430024C06Rik	RIKEN cDNA E430024C06 gene	2.50
NM_008607	Mmp13	matrix metalloproteinase 13	2.49
NM_007502	Atp1b3	ATPase, Na ⁺ /K ⁺ transporting, beta 3 polypeptide	2.39
NM_009876	Cdkn1c	cyclin-dependent kinase inhibitor 1C (P57)	2.35
NM_007725	Cnn2	calponin 2	2.26
NM_011338	Ccl9	chemokine (C-C motif) ligand 9	2.25
NM_023596	Slc29a3	solute carrier family 29 (nucleoside transporters), member 3	2.23
NM_011299	Rps6ka2	ribosomal protein S6 kinase, polypeptide 2	2.22
NM_153560	C230093N12Rik	RIKEN cDNA C230093N12 gene	2.16
NM_133792	Lypla3	lysophospholipase 3	2.13
NM_029422	Tm7sf4	transmembrane 7 superfamily member 4	2.12
NM_133656	Crk	v-crk sarcoma virus CT10 oncogene homolog (avian)	2.08
NM_010828	Cited2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	2.06
BC016624	Actb	actin, beta, cytoplasmic	2.06
NM_008538	Marcks	myristoylated alanine rich protein kinase C substrate	2.05
NM_024166	Chchd2	coiled-coil-helix-coiled-coil-helix domain containing 2	2.03
AK164477	Arhgap19	Rho GTPase activating protein 19	2.02
NM_008037	Fosl2	fos-like antigen 2	2.02
NM_181405	Rnpepl1	arginyl aminopeptidase (aminopeptidase B)-like 1	2.02
NM_010240	Ftl1	ferritin light chain 1	2.01
NM_009272	Srm	spermidine synthase	2.00
XM_142195	LOC245651	similar to coiled-coil-helix-coiled-coil-helix domain containing 2 (predicted)	2.00
XM_980293	Agmat	agmatine ureohydrolase (agmatinase)	1.98
NM_008049	Ftl2	ferritin light chain 2	1.98
NM_013855	Abca3	ATP-binding cassette, sub-family A (ABC1), member 3	1.98
NM_016917	Slc40a1	solute carrier family 40 (iron-regulated transporter), member 1	1.97
NM_011660	Txn1	thioredoxin 1	1.97
NM_009502	Vcl	vinculin	1.97
NM_133970	D8Erd354e	DNA segment, Chr 8, ERATO Doi 354, expressed	1.96
NM_018824	Slc23a2	solute carrier family 23 (nucleobase transporters), member 2	1.95
NM_018729	Cd244	CD244 natural killer cell receptor 2B4	1.95
NM_138313	Bmf	Bcl2 modifying factor	1.95
NM_028304	4933435A13Rik	RIKEN cDNA 4933435A13 gene	1.93
NM_175145	Tmem127	transmembrane protein 127	1.92
NM_009101	Rras	Harvey rat sarcoma oncogene, subgroup R	1.92
NM_007381	Acadl	acyl-Coenzyme A dehydrogenase, long-chain	1.90
NM_007392	Acta2	actin, alpha 2, smooth muscle, aorta	1.90
NM_177687	Crebl2	cAMP responsive element binding protein-like 2	1.88
AK003800	Tceal7	transcription elongation factor A (SII)-like 7	1.87
BC046640	2610307O08Rik	RIKEN cDNA 2610307O08 gene	1.86
BC058094	Solh	small optic lobes homolog (Drosophila)	1.86

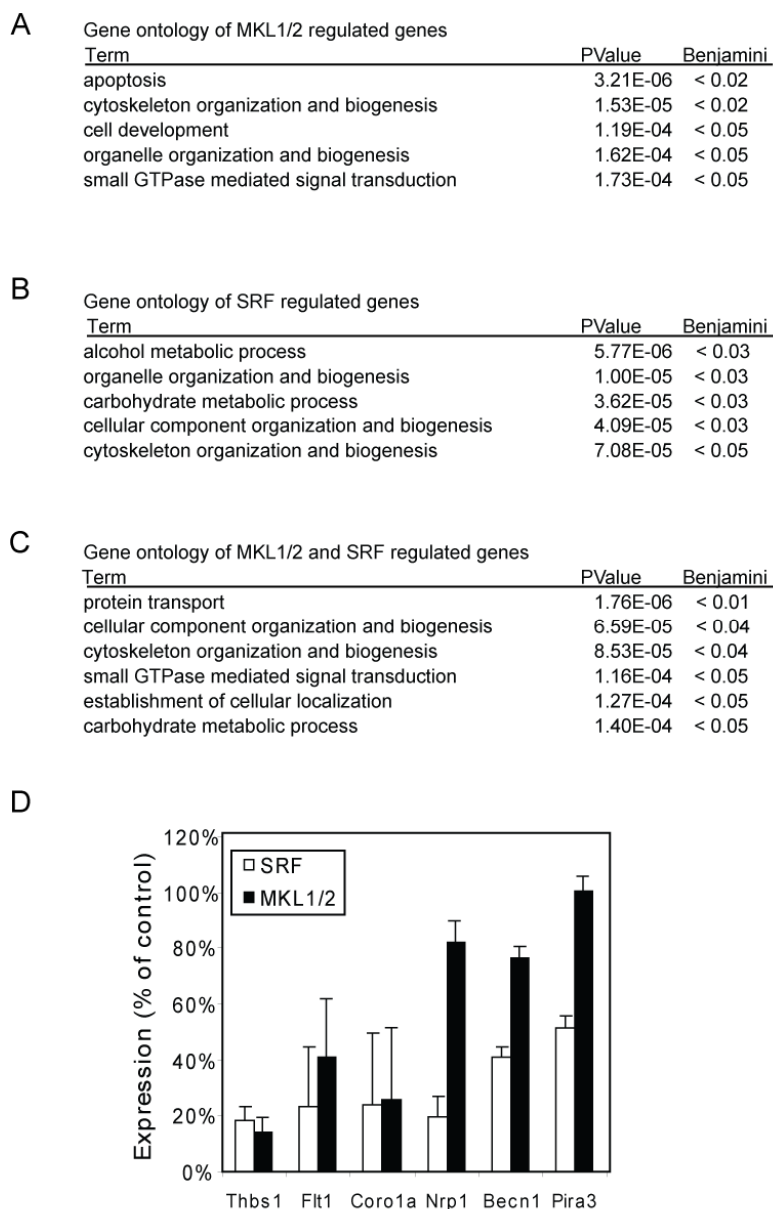


Figure 3 MKL and SRF regulate gene expression in macrophages. (A-C) Gene ontology (GO) functional annotation analysis of MKL (A), SRF (B), and both MKL and SRF (C) target genes identified by expression microarray analysis of non-specific control or specific siRNA transfected primary macrophage samples. Data shown are selected enriched terms using those genes that were significantly changed (FDR<0.01) in each condition. (D) qPCR confirmation of MKL and SRF target genes identified by the microarray analysis. Target gene expression after 48 hours of SRF (white bars) or MKL (black bars) siRNA treatment is represented as the average percent of control of three independent experiments.

Table 2 Full results of gene ontology analysis for genes regulated by MKL. Significant GO terms identified through the analysis of the genes that are regulated by MKL according to the expression microarray results.

Term	Input genes in term	Total input genes	PValue	Benjamini
GO:0006915~apoptosis	121	2264	3.10E-06	0.016
GO:0007242~intracellular signaling cascade	191	2264	4.40E-06	0.011
GO:0008152~metabolic process	969	2264	5.96E-06	0.010
GO:0012501~programmed cell death	121	2264	7.29E-06	0.009
GO:0006464~protein modification process	240	2264	1.20E-05	0.012
GO:0043412~biopolymer modification	248	2264	1.27E-05	0.011
GO:0007010~cytoskeleton organization and biogenesis	91	2264	1.53E-05	0.011
GO:0044238~primary metabolic process	877	2264	2.63E-05	0.017
GO:0016265~death	122	2264	2.75E-05	0.014
GO:0008219~cell death	122	2264	2.75E-05	0.014
GO:0009605~response to external stimulus	89	2264	3.32E-05	0.015
GO:0006796~phosphate metabolic process	138	2264	6.93E-05	0.029
GO:0006793~phosphorus metabolic process	138	2264	6.93E-05	0.029
GO:0043687~post-translational protein modification	208	2264	6.97E-05	0.025
GO:0048468~cell development	186	2264	1.19E-04	0.040
GO:0044237~cellular metabolic process	868	2264	1.58E-04	0.050
GO:0006996~organelle organization and biogenesis	171	2264	1.62E-04	0.048
GO:0007264~small GTPase mediated signal transduction	72	2264	1.73E-04	0.048

Table 3 Top genes downregulated by SRF knockdown. Microarray results showing the genes whose expressions are most decreased, compared to NS control, after treatment with SRF siRNAs in primary macrophages.

Accession #	Symbol	Description	fold decrease
NM_011580	Thbs1	thrombospondin 1	4.90
NM_001033042	AI595366	expressed sequence AI595366	3.48
NM_009876	Cdkn1c	cyclin-dependent kinase inhibitor 1C	3.38
NM_025425	Rpl3l	ribosomal protein L3-like	2.89
NM_016917	Slc40a1	solute carrier family 40 (iron-regulated transporter), member 1	2.69
NM_009898	Coro1a	coronin, actin binding protein 1A	2.59
NM_008348	Il10ra	interleukin 10 receptor, alpha	2.54
NM_022320	Gpr35	G protein-coupled receptor 35	2.53
NM_028784	F13a1	coagulation factor XIII, A1 subunit	2.39
NM_153592	Erlin2	ER lipid raft associated 2	2.33
NM_031167	Il1rn	interleukin 1 receptor antagonist, transcript variant 1	2.31
NM_027950	Osgin1	oxidative stress induced growth inhibitor 1	2.14
XM_001480835	ENSMUSG00000043795	predicted gene, ENSMUSG00000043795	2.13
NM_009984	Ctsl	cathepsin L	2.10
XM_001475483	LOC100046056	similar to Pre-B-cell leukemia transcription factor interacting protein 1	2.06
NM_013658	Sema4a	semaphorin 4A	2.06
NM_174995	Mgst2	microsomal glutathione S-transferase 2	2.06
NM_009695	Apoc2	apolipoprotein C-II	2.04
NM_013680	Syn1	synapsin I	2.04
NM_153560	Fam102a	family with sequence similarity 102, member A	2.03
NM_027836	Ms4a7	membrane-spanning 4-domains, subfamily A, member 7, variant 1	2.00
NM_023680	Tnfrsf22	tumor necrosis factor receptor superfamily, member 22	1.97
NM_007836	Gadd45a	growth arrest and DNA-damage-inducible 45 alpha	1.97
NM_020493	Srf	serum response factor	1.93
NM_026346	Fbxo32	F-box protein 32	1.92
NM_175649	Tnfrsf26	tumor necrosis factor receptor superfamily, member 26	1.90
NM_029884	Hgsnat	heparan-alpha-glucosaminide N-acetyltransferase	1.89
NM_178111	Trp53inp2	transformation related protein 53 inducible nuclear protein 2	1.88
NM_010228	Flt1	FMS-like tyrosine kinase 1	1.87
NM_008823	Cfp	complement factor properdin	1.86
NM_008397	Itga6	integrin alpha 6	1.85
NM_178420	Nlr1	NLR family member X1	1.83
NM_145634	Cd300lf	CD300 antigen like family member F	1.83
NM_008873	Plau	plasminogen activator, urokinase	1.81
NM_007642	Cd28	CD28 antigen	1.81
XM_001479169	LOC100048845	similar to CD28 antigen	1.80
NM_178045	Rassf4	Ras association domain family member 4	1.79
NM_011819	Gdf15	growth differentiation factor 15	1.77
NM_027667	Arhgap19	Rho GTPase activating protein 19	1.76
NM_008607	Mmp13	matrix metalloproteinase 13	1.76
NM_172161	Irak2	interleukin-1 receptor-associated kinase 2	1.75
NM_009405	Tnni2	troponin I, skeletal, fast 2	1.75
NM_011436	Sorl1	sortilin-related receptor, LDLR class A repeats-containing	1.75
NM_010902	Nfe2l2	nuclear factor, erythroid derived 2, like 2	1.73
NM_016961	Mapk9	mitogen-activated protein kinase 9, transcript variant 2	1.73
NM_010634	Fabp5	fatty acid binding protein 5, epidermal	1.72
NM_172285	Plcg2	phospholipase C, gamma 2	1.71
NM_134023	Tbc1d10a	TBC1 domain family, member 10a	1.70
NM_009610	Actg2	actin, gamma 2, smooth muscle, enteric	1.70
NM_133954	AA960436	expressed sequence AA960436	1.68

Table 4 Full results of gene ontology analysis for genes regulated by SRF Significant GO terms identified through the analysis of the genes that are regulated by both SRF according to the expression microarray results.

Term	Input genes in term	Total input genes	PValue	Benjamini
GO:0006066~alcohol metabolic process	37	1093	5.77E-06	0.029
GO:0006996~organelle organization and biogenesis	98	1093	1.00E-05	0.025
GO:0005996~monosaccharide metabolic process	24	1093	1.15E-05	0.019
GO:0019318~hexose metabolic process	23	1093	2.84E-05	0.035
GO:0007242~intracellular signaling cascade	103	1093	3.10E-05	0.031
GO:0005975~carbohydrate metabolic process	46	1093	3.62E-05	0.030
GO:0016043~cellular component organization and biogenesis	188	1093	4.09E-05	0.029
GO:0007010~cytoskeleton organization and biogenesis	49	1093	7.08E-05	0.043
GO:0044262~cellular carbohydrate metabolic process	34	1093	7.68E-05	0.042

Table 5 Top genes commonly downregulated by MKL or SRF knockdown. Microarray results of the genes whose expressions are most decreased, compared to NS control, after treatment with either MKL1/2 or SRF siRNAs in primary macrophages.

Accession #	Symbol	Description	MKL fold down	SRF fold down
NM_011580	Thbs1	thrombospondin 1	4.60	4.90
NM_009876	Cdkn1c	cyclin-dependent kinase inhibitor 1C (P57)	2.35	3.38
NM_016917	Slc40a1	solute carrier family 40 (iron-regulated transporter), member 1	1.97	2.69
NM_009898	Coro1a	coronin, actin binding protein 1A	2.65	2.59
NM_008348	Il10ra	interleukin 10 receptor, alpha	1.56	2.54
NM_028784	F13a1	coagulation factor XIII, A1 subunit	1.83	2.39
NM_009984	Ctsl	cathepsin L	1.62	2.10
NM_013658	Sema4a	semaphorin 4a	1.48	2.06
NM_153560	C230093N12Rik	RIKEN cDNA C230093N12 gene	2.16	2.03
NM_027836	Ms4a7	membrane-spanning 4-domains, subfamily A, member 7	1.65	2.00
NM_020493	Srf	serum response factor	1.57	1.93
NM_026346	Fbxo32	F-box only protein 32	1.76	1.92
NM_133970	D8Erd354e	DNA segment, Chr 8, ERATO Doi 354, expressed	1.96	1.89
NM_010228	Flt1	FMS-like tyrosine kinase 1	3.34	1.87
NM_008823	Cfp	complement factor properdin	1.61	1.86
BC024571	Itga6	integrin alpha 6	1.49	1.85
NM_008873	Plau	plasminogen activator, urokinase	1.70	1.81
AK164477	Arhgap19	Rho GTPase activating protein 19	2.02	1.76
NM_008607	Mmp13	matrix metalloproteinase 13	2.49	1.76
NM_172161	Ilr1k2	interleukin-1 receptor-associated kinase 2	1.80	1.75
NM_009610	Actg2	actin, gamma 2, smooth muscle, enteric	1.69	1.70
NM_177715	Kctd12	potassium channel tetramerisation domain containing 12	1.41	1.67
NM_009846	Cd24a	CD24a antigen	1.82	1.60
NM_011216	Ptpro	protein tyrosine phosphatase, receptor type, O	1.41	1.60
NM_173442	Gcnt1	glucosaminyl (N-acetyl) transferase 1, core 2	1.63	1.57
NM_007520	Bach1	BTB and CNC homology 1	1.79	1.56
NM_021273	Ckb	creatine kinase, brain	1.60	1.56
NM_008338	Ifngr2	interferon gamma receptor 2	1.41	1.55
NM_007574	C1qc	complement component 1, q subcomponent, C chain	1.65	1.53
NM_008685	Nfe2l3	nuclear factor, erythroid derived 2	1.40	1.52
NM_152220	Stx3	syntaxin 3	1.77	1.49
NM_016895	Ak2	adenylate kinase 2	1.57	1.47
NM_007392	Acta2	actin, alpha 2, smooth muscle, aorta	1.90	1.47
NM_134050	Rab15	RAB15, member RAS oncogene family	1.46	1.47
NM_010211	Fhl1	four and a half LIM domains 1	1.49	1.46
NM_010937	Nras	neuroblastoma ras oncogene	1.44	1.44
NM_175145	Tmem127	transmembrane protein 127	1.92	1.43
AK010701	Dedd2	death effector domain-containing DNA binding protein 2	1.67	1.42
NM_009502	Vcl	vinculin	1.97	1.42
NM_008879	Lcp1	lymphocyte cytosolic protein 1	1.51	1.40
NM_009777	C1qb	complement component 1, q subcomponent, beta polypeptide	1.58	1.39
NM_011338	Ccl9	chemokine (C-C motif) ligand 9	2.25	1.37
NM_028802	Prai4	preimplantation protein 4	1.59	1.36
NM_021439	Chst11	carbohydrate sulfotransferase 11	1.46	1.36
NM_007805	Cyb5b1	cytochrome b-561	1.48	1.36
NM_009579	Slc30a1	solute carrier family 30 (zinc transporter), member 1	1.49	1.35
NM_023063	Lima1	LIM domain and actin binding 1	1.40	1.35
NM_026163	Pkp2	plakophilin 2	1.79	1.34
NM_146045	B4gal7	xylosylprotein beta1,4-galactosyltransferase, polypeptide 7	1.42	1.34
NM_026647	Zdhhc21	zinc finger, DHHC domain containing 21	1.43	1.34

Table 6 Full results of gene ontology analysis for genes regulated by both MKL and SRF Significant GO terms identified through the analysis of the genes that are regulated by both MKL and SRF according to the expression microarray results.

Term	Input genes in term	Total input genes	PValue	Benjamini
GO:0015031~protein transport	80	1216	1.76E-06	0.009
GO:0007242~intracellular signaling cascade	124	1216	2.93E-06	0.007
GO:0045184~establishment of protein localization	83	1216	3.25E-06	0.005
GO:0046907~intracellular transport	76	1216	1.78E-05	0.022
GO:0008104~protein localization	85	1216	2.64E-05	0.026
GO:0051179~localization	252	1216	3.22E-05	0.026
GO:0006810~transport	220	1216	3.54E-05	0.025
GO:0051234~establishment of localization	225	1216	4.39E-05	0.027
GO:0033036~macromolecule localization	87	1216	4.57E-05	0.025
GO:0016043~cellular component organization and biogenesis	219	1216	6.59E-05	0.032
GO:0007010~cytoskeleton organization and biogenesis	56	1216	8.53E-05	0.038
GO:0007264~small GTPase mediated signal transduction	48	1216	1.16E-04	0.047
GO:0051649~establishment of cellular localization	87	1216	1.27E-04	0.048
GO:0005975~carbohydrate metabolic process	50	1216	1.40E-04	0.049

Table 7 Top SRF ChIP-seq peaks in primary macrophages. The data in this table represents the most significant peaks identified in the SRF ChIP seq in primary macrophages. Data highlighted in blue, orange, and green represent those peaks where a CARG box, PU.1 motif, or both motifs, respectively, were found within 200 base pairs of the center of the peak.

Chromosome	Start	End	Distance to TSS	Nearest PromoterID	Gene Name	Tag count
chr13	109010701	109011101	-90	NM_145456	Zswim6	475.5
chr5	143171718	143172118	-54	NM_007393	Actb	421.29
chr7	121923955	121924355	-34	NM_021608	Dctn5	405.48
chr11	69478309	69478709	328	NM_011900	Mpdu1	381.76
chr2	91530749	91531149	-28	NM_172669	Ambra1	327.54
chr1	174260123	174260523	47117	NM_026234	Pigm	324.16
chr11	87565602	87566002	-10934	NM_172449	Bzrap1	310.6
chr18	34986607	34987007	-374	NM_007913	Egr1	309.47
chr14	68809845	68810245	-2480	NM_018781	Egr3	307.21
chr10	119611590	119611990	-96	NM_026617	Tmbim4	304.96
chr10	82383231	82383631	-31865	NM_021439	Chst11	302.7
chr13	30981359	30981759	-47	NM_025588	Exoc2	298.18
chr7	125235587	125235987	-37	NM_029842	Jmjd5	297.05
chr19	5488951	5489351	-815	NM_027877	Mus81	291.4
chr7	27199577	27199977	4064	NM_018820	Sertad1	283.5
chr8	87553981	87554381	-1224	NM_010499	Ier2	282.37
chr11	98498412	98498812	-32	NM_009439	Psmc3	281.24
chr5	139598659	139599059	-51	NM_197980	Cox19	263.17
chr11	84732437	84732837	-86	NM_001005223	Znhit3	262.04
chr11	103083515	103083915	-224	NM_016896	Map3k14	256.39
chr6	52772148	52772548	129022	NM_025816	Tax1bp1	256.39
chr6	87468544	87468944	-6803	NM_175476	Arhgap25	248.48
chr9	57652382	57652782	-20211	NM_019689	Arid3b	243.96
chr2	142754243	142754643	-67	NM_021335	Snrbp2	234.93
chr8	87866539	87866939	2114	NM_008416	Junb	231.54
chr4	144461544	144461944	-4465	NM_001128198	Vps13d	223.63
chr11	75006524	75006924	-46	NM_144491	Dph1	220.25
chr2	172061929	172062329	-79	NM_024199	Cstf1	219.12
chr16	95837667	95838067	26583	NM_011809	Ets2	216.86
chr6	83447948	83448348	24404	NM_013764	Dguok	215.73
chr19	53582179	53582579	-251	NM_001085390	Dusp5	214.6
chr2	166792941	166793341	-38921	NM_001033196	Znfx1	213.47
chr4	44093619	44094019	11447	NM_015828	Gne	206.69
chr7	64252870	64253270	18116	NM_026483	Mphosph10	198.79
chr15	76453166	76453566	-84	NM_025842	Vps28	197.66
chr13	114739315	114739715	-73	NM_130796	Snx18	196.53
chr15	78704156	78704556	-89	NM_145929	Gga1	194.27
chr6	99436391	99436791	195643	NM_025829	Eif4e3	193.14
chr3	67073369	67073769	-32	NM_025822	Rsrc1	192.01
chr5	114133885	114134285	35620	NM_011779	Coro1c	190.88
chr7	114067091	114067491	6	NM_011965	Psmc1	189.75
chr10	105245321	105245721	-57	NM_207522	BC067068	188.62
chr13	76564529	76564929	-7	NM_001081352	Ttc37	187.49
chr1	4847594	4847994	-101	NM_011541	Tcea1	186.36
chr12	52892841	52893241	-100	NM_177171	Heatr5a	185.23
chr8	122698908	122699308	27448	NM_028071	Cotl1	182.97
chr7	44865508	44865908	6	NM_016849	Irf3	179.58
chr5	143170801	143171201	863	NM_007393	Actb	179.58
chr2	32058426	32058826	-66	NM_145145	Pomt1	176.2
chr14	19680434	19680834	-37351	NM_009502	Vcl	175.07
chr6	129198477	129198877	42385	NM_001033122	Cd69	173.94
chr12	86375012	86375412	12208	NM_010234	Fos	172.81
chr8	42096653	42097053	36756	NM_001040699	Mttr7	170.55
chr5	145420476	145420876	-75	NM_178576	Cpsf4	166.03
chr10	66932627	66933027	-399	NM_010118	Egr2	163.77
chr4	151169354	151169754	-124	NM_009079	Rpl22	163.77
chr8	122697817	122698217	28539	NM_028071	Cotl1	162.64
chr7	18540174	18540574	-187	NM_007949	Ercc2	161.51
chrY	1510282	1510682	-55165	NM_009571	Zfy2	158.12

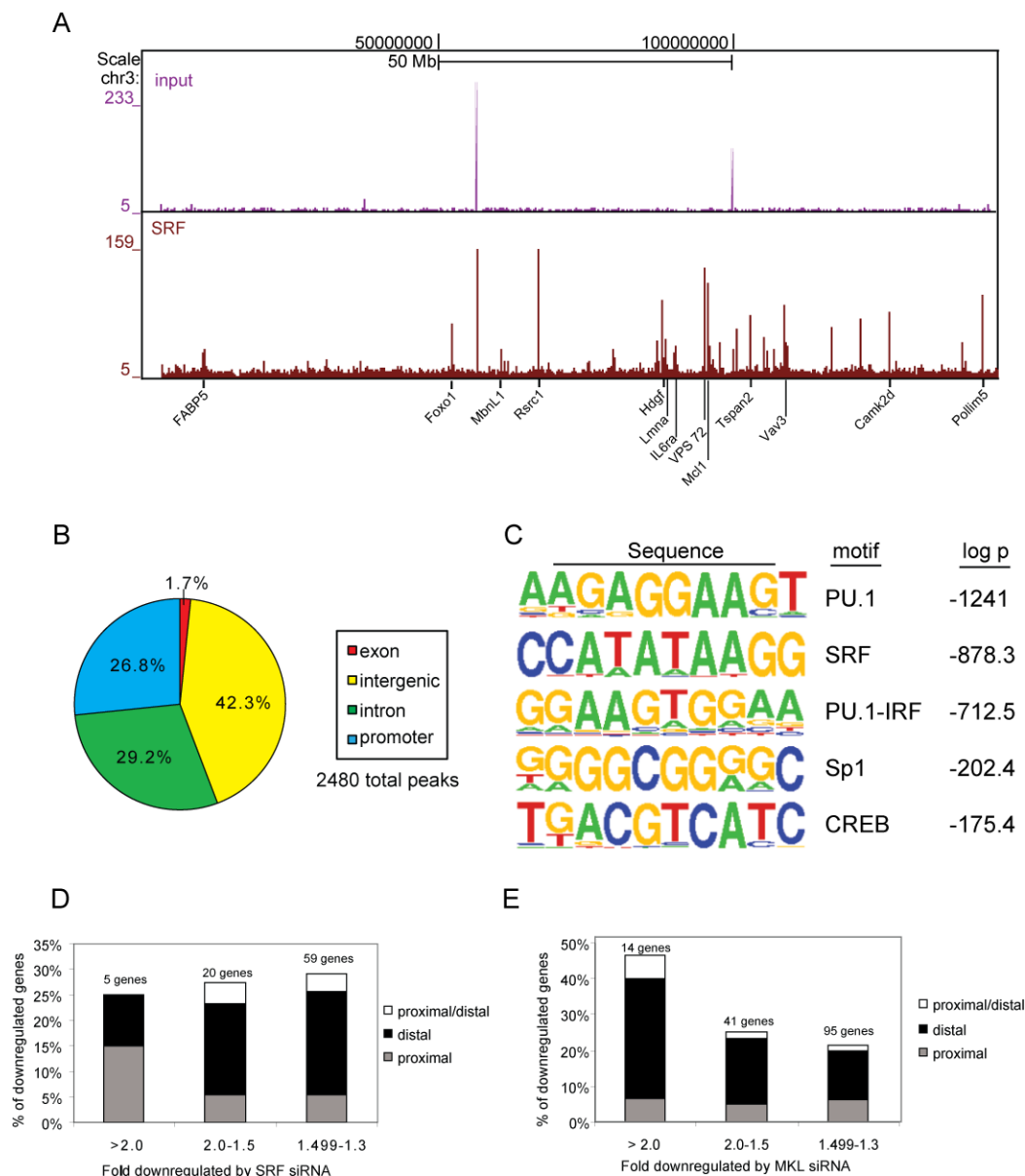


Figure 4 Genomic location annotation and motif analysis of SRF ChIP-seq peaks in primary macrophages. (A) UCSC genome browser image of ChIP-seq peak data for mouse chromosome 3 in primary macrophages. Both input (top lane) and SRF (bottom lane) peaks are shown. (B) Genomic location annotation analysis of significant (>12tags) SRF ChIP-seq peaks. (C) HOMER *de novo* motif analysis of all significant SRF peaks in primary macrophages. Peak sequences were defined as those sequences located +/- 200bp from the center of the SRF peak. (D and E) Graphs represent the percentages and locations (promoter or distal) of SRF or MKL dependent genes, respectively, identified by the microarray analysis that are associated with SRF ChIP-seq peaks.

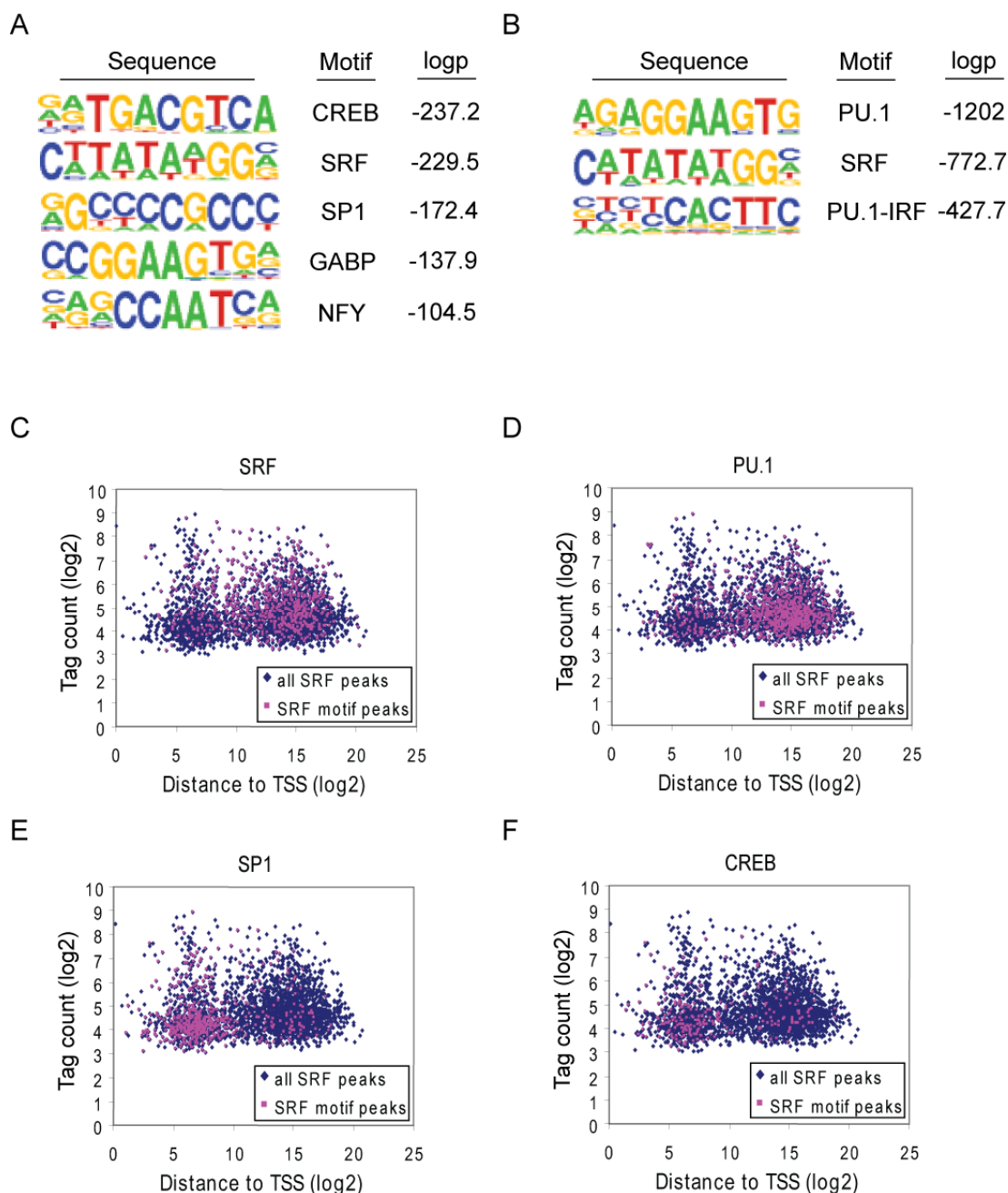


Figure 5 SRF associates with different transcription factors depending on genomic location. (A and B) HOMER de novo motif analysis of SRF ChIP-seq peak sequences, as described in Figure 4, located in promoter (-500bp to +500bp) (A) and distal (excluding promoter) (B) regions in primary macrophages. (C-F) Scatter plots representing all SRF peak locations relative to their corresponding tag counts (blue diamonds). Red squares denote those SRF peaks that have the indicated sequence motifs (SRF (C), PU.1 (D), SP1 (E), and CREB (F)) located within 200bp of the center of the peak.

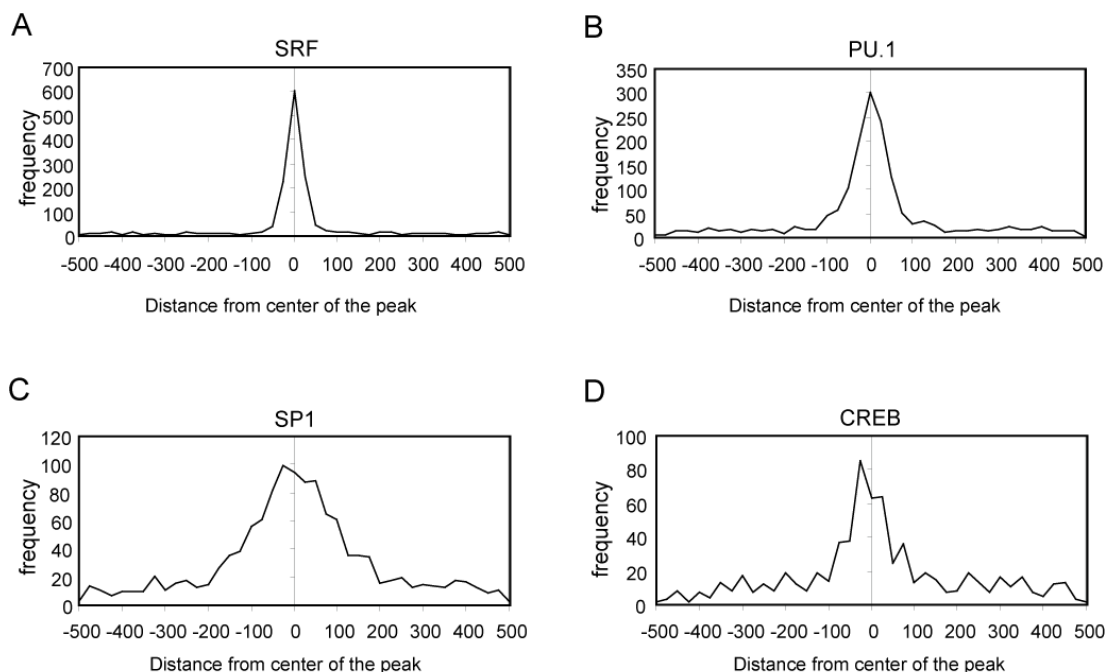


Figure 6 Motif distribution relative to the center of the SRF ChIP-seq peak. (A-D) Histogram analysis of the location of the indicated DNA sequence motifs relative to the center of the SRF ChIP-seq peaks in primary macrophages. The frequency for each graph was determined by counting the number of peaks that contain the indicated motif in 25bp windows starting at the center of the SRF peak and ending 500bp away from the center of the peak in each direction (+ and -).

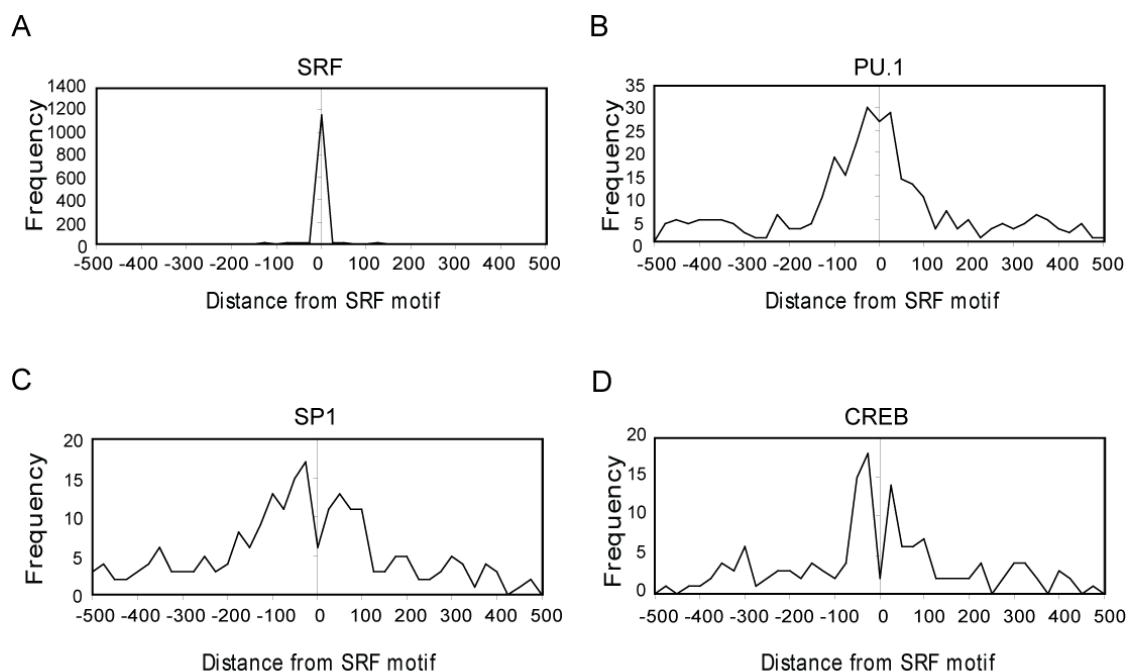


Figure 7 Motif distribution relative to the SRF motif. (A-D) Histogram analysis was performed as in Fig. 6 except that the SRF ChIP-seq peaks were first centered on the location of the SRF motif (0). Those peaks that were not identified as having SRF motifs were omitted from this analysis.

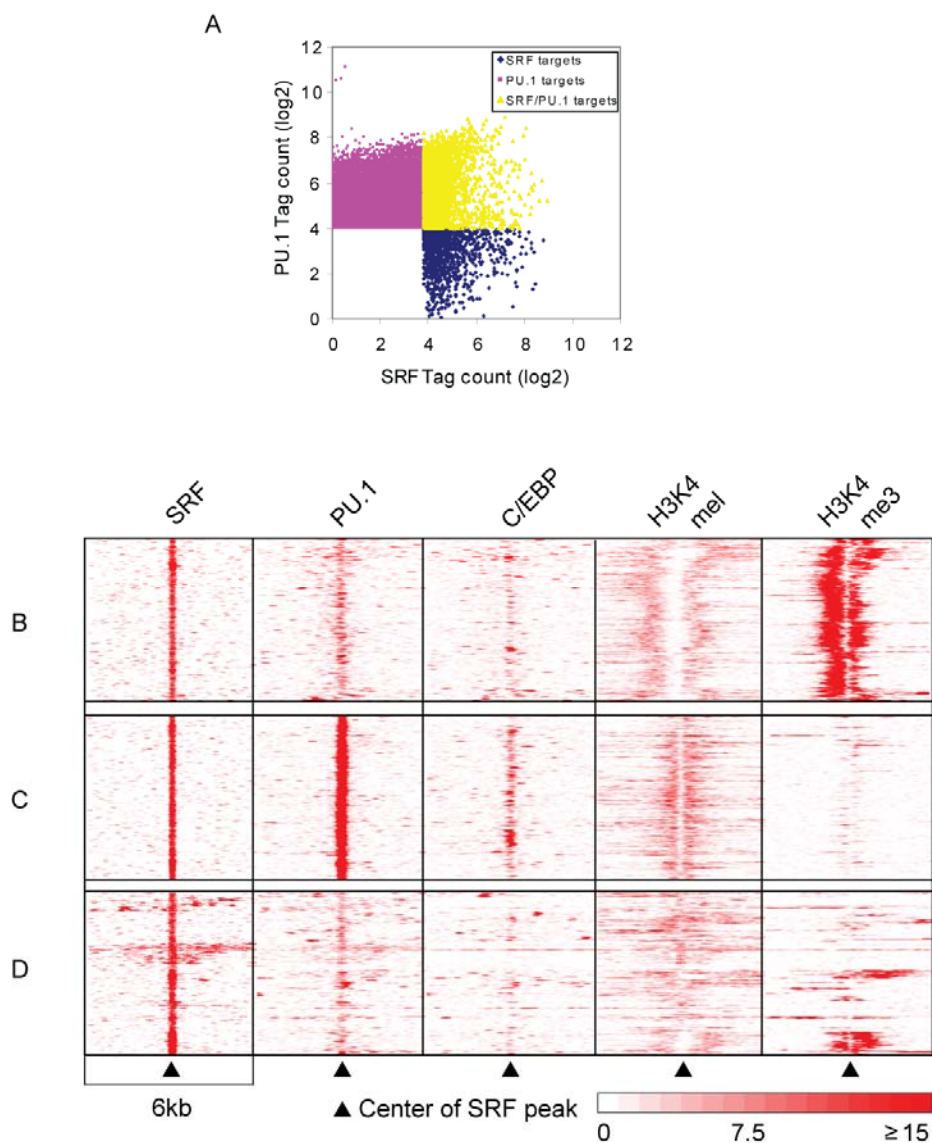


Figure 8 Distal SRF/PU.1 peaks are associated with the H3K4me1 enhancer mark. (A) Scatter plot comparison of the SRF and PU.1 ChIP-seq peaks in primary macrophages. Blue and pink dots represent those peaks that bind SRF (>12tags), but not PU.1 (<14tags) and bind PU.1 (>14tags), but not SRF (<12tags), respectively. Yellow dots represent those sites where the centers of the SRF and PU.1 peaks are located within 100bp of each other. (B-D) Heat map representation of the location of the indicated factors/histone marks within the 6kb region surrounding the center of the SRF peak. Each column represents all ChIP-seq data (no tag cutoff) in primary macrophages for the indicated factor/mark. Results were clustered and representative regions of the major binding patterns are shown. White represents areas of low numbers of sequence tags and red represents areas of high numbers of sequence tags. The full heat map can be found in Fig. 9.

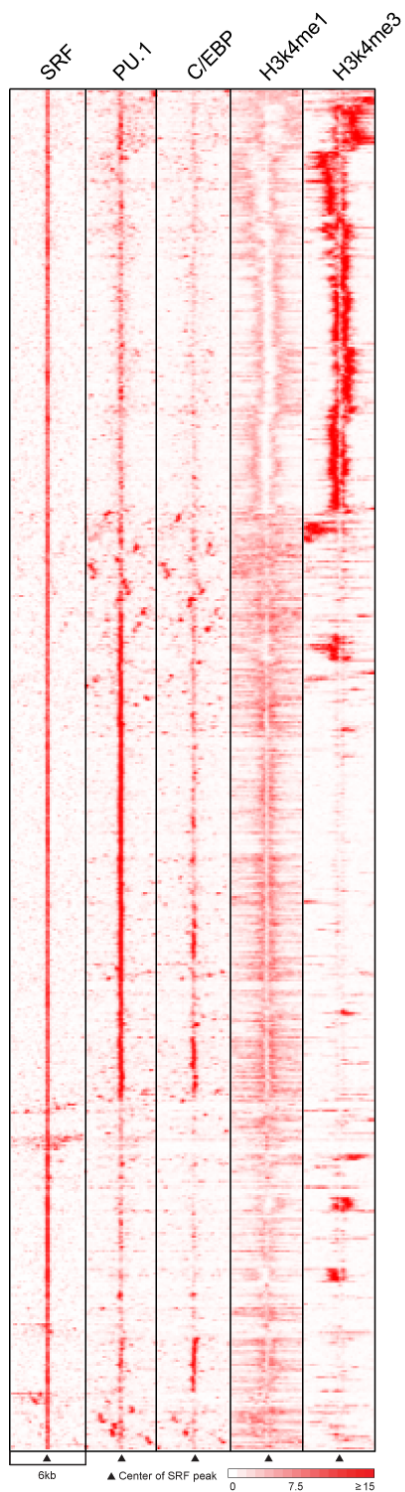


Figure 9 Heat map analysis of ChIP-seq data in primary macrophages. Full heat map of the ChIP-seq data in primary macrophages as described in Figure 8B-D.








Table 8 Top SRF ChIP-seq peaks in undifferentiated PUER cells. The data in this table represents the most significant peaks identified in the SRF ChIP seq undifferentiated PUER cells. Data highlighted in blue, orange, and green represent those peaks where a CA_rG box, PU.1 motif, or both motifs, respectively, were found within 200 base pairs of the center of the peak.

Chromosome	Start	End	Distance to TSS	Nearest PromoterID	Gene Name	Tag Count
chr13	109010706	109011106	-95	NM_145456	Zswim6	1339.25
chr5	143171724	143172124	-60	NM_007393	Actb	1338.1
chr19	53582180	53582580	-250	NM_001085390	Dusp5	991.48
chr8	87553982	87554382	-1225	NM_010499	Ier2	986.88
chr18	34986900	34987300	-81	NM_007913	Egr1	868.27
chr11	69478309	69478709	328	NM_011900	Mpdu1	854.45
chr11	87565600	87566000	-10936	NM_172449	Bzrap1	841.78
chr19	5488949	5489349	-813	NM_027877	Mus81	808.39
chr11	103083509	103083909	-218	NM_016896	Map3k14	804.93
chrX	162005820	162006220	-76199	NM_019397	Egfl6	772.69
chr15	77643239	77643639	25931	NM_022410	Myh9	726.63
chr8	87866536	87866936	2117	NM_008416	Junb	686.32
chr3	95743920	95744320	-5	NM_008562	Mcl1	669.05
chr7	27199570	27199970	4057	NM_018820	Sertad1	647.17
chr5	143170801	143171201	863	NM_007393	Actb	617.23
chr17	46019225	46019625	-74	NM_020493	Srf	568.87
chr18	5134064	5134464	87679	NM_153153	Svil	558.5
chr18	35013302	35013702	26321	NM_007913	Egr1	544.68
chr16	95837673	95838073	26589	NM_011809	Ets2	543.53
chr7	18289898	18290298	11706	NM_175530	Fbxo46	520.5
chr10	66932632	66933032	-394	NM_010118	Egr2	510.14
chr7	44865501	44865901	-1	NM_016849	Irf3	508.98
chr8	122697822	122698222	28534	NM_028071	Cotl1	488.26
chr10	119611597	119611997	-89	NM_026617	Tmbim4	487.11
chr9	65349066	65349466	28774	NM_153119	Plekho2	479.04
chr11	118869332	118869732	-12529	NM_013926	Cbx8	476.74
chr7	125235592	125235992	-32	NM_029842	Jrmd5	474.44
chr2	170304577	170304977	102342	NM_001110152	Pfdn4	473.29
chr14	68809847	68810247	-2478	NM_018781	Egr3	470.98
chr8	23957560	23957960	-15360	NM_001012667	AI316807	470.98
chr6	99436396	99436796	195638	NM_025829	Eif4e3	452.56
chr12	86375015	86375415	12211	NM_010234	Fos	446.8
chr4	118603578	118603978	-2899	NM_011400	Slc2a1	436.44
chr13	30981365	30981765	-53	NM_025588	Exoc2	432.98
chr8	12703570	12703970	-31693	NM_198031	Tubgcp3	431.83
chr11	120166113	120166513	-1739	NM_009609	Actg1	428.38
chr4	125549052	125549452	20	NM_025544	Mrps15	424.92
chr13	94404215	94404615	-833	NM_147176	Homer1	420.32
chr10	80104854	80105254	-105	NM_172457	Mobkl2a	418.01
chr2	142754249	142754649	-61	NM_021335	Snrpb2	412.25
chr5	139598660	139599060	-52	NM_197980	Cox19	407.65
chr7	121923945	121924345	-44	NM_021608	Dctn5	406.5
chr8	122698905	122699305	27451	NM_028071	Cotl1	397.28
chr11	75006515	75006915	-37	NM_144491	Dph1	392.68
chr1	145464861	145465261	-159	NM_145991	Cdc73	390.38
chr11	116036514	116036914	4622	NR_004853	Cdk3	386.92
chr10	127895802	127896202	772	NM_010860	Myl6	368.5
chr13	55891134	55891534	-50287	NM_011097	Pitx1	363.89
chr16	4795564	4795964	-5716	NM_029657	Mgrr1	346.62
chr4	122388646	122389046	-662	NM_007598	Cap1	339.71
chr5	143194138	143194538	-22474	NM_007393	Actb	337.4
chr19	34341651	34342051	-14812	NM_007987	Fas	337.4
chr17	24206470	24206870	8	NM_019988	Mist8	328.19
chr18	75573583	75573983	81053	NM_001042660	Smad7	325.89
chr4	135389069	135389469	-26629	NM_008321	Id3	325.89
chr9	50449854	50450254	30151	NM_178118	Dixdc1	324.74
chr2	91530755	91531155	-22	NM_172669	Ambra1	324.74
chr15	77627749	77628149	41421	NM_022410	Myh9	323.59
chr5	38849795	38850195	-151	NM_011715	Wdr1	321.28

Table 9 Top SRF ChIP-seq peaks in differentiated PUER cells. The data in this table represents the most significant peaks identified in the SRF ChIP seq in PUER cells treated with tamoxifen for 24 hours. Data highlighted in blue, orange, and green represent those peaks where a CA₂G box, PU.1 motif, or both motifs, respectively, were found within 200 base pairs of the center of the peak

Chromosome	Start	End	Distance to TSS	Nearest PromoterID	Gene Name	Tag Count
chr13	109010706	109011106	-95	NM_145456	Zswim6	1033.7
chr19	53582181	53582581	-249	NM_001085390	Dusp5	904.67
chr5	143171720	143172120	-56	NM_007393	Actb	865.53
chr8	87553985	87554385	-1228	NM_010499	Ier2	790.14
chr18	34986896	34987296	-85	NM_007913	Egr1	719.1
chr3	95743921	95744321	-4	NM_008562	Mcl1	700.25
chr11	87565598	87565998	-10938	NM_172449	Bzap1	621.96
chr11	103083510	103083910	-219	NM_016896	Map3k14	611.81
chr11	69478306	69478706	331	NM_011900	Mpdu1	610.36
chr11	120166108	120166508	-1734	NM_009609	Actg1	562.52
chr19	5488950	5489350	-814	NM_027877	Mus81	555.27
chr15	77643236	77643636	25934	NM_022410	Myh9	527.72
chr8	87866542	87866942	2111	NM_008416	Junb	524.83
chr5	143194144	143194544	-22480	NM_007393	Actb	517.58
chr7	27199569	27199969	4056	NM_018820	Sertad1	492.93
chr18	5134070	5134470	87685	NM_153153	Svil	490.03
chr18	35013304	35013704	26323	NM_007913	Egr1	471.18
chr17	46019221	46019621	-70	NM_020493	Srf	449.44
chr14	68809830	68810230	-2495	NM_018781	Egr3	449.44
chr8	122697823	122698223	28533	NM_028071	Cotl1	437.84
chr10	66932639	66933039	-387	NM_010118	Egr2	414.64
chr10	119611595	119611995	-91	NM_026617	Tmbim4	411.74
chr2	172061932	172062332	-76	NM_024199	Cstf1	408.84
chr7	44865495	44865895	-7	NM_016849	Irf3	397.24
chr10	80104855	80105255	-106	NM_172457	Mobkl2a	392.89
chr6	87468536	87468936	-6795	NM_175476	Arhgap25	371.15
chr2	143606390	143606790	-182	NM_019771	Dstn	369.7
chr19	46128962	46129362	-12	NM_001039352	Nolc1	365.35
chr7	18289900	18290300	11708	NM_175530	Fbxo46	363.9
chr7	121923945	121924345	-44	NM_021608	Dctn5	363.9
chr11	118869328	118869728	-12525	NM_013926	Cbx8	363.9
chr8	122670170	122670570	-5977	NM_028883	4632415K11Rik	363.9
chr12	86375015	86375415	12211	NM_010234	Fos	361
chr6	99436394	99436794	195640	NM_025829	Eif4e3	347.95
chr4	125549052	125549452	20	NM_025544	Mrps15	346.5
chr7	125235589	125235989	-35	NM_029842	Jmjd5	343.6
chr13	94404217	94404617	-831	NM_147176	Homer1	342.15
chr5	139598669	139599069	-61	NM_197980	Cox19	340.7
chr2	164506616	164507016	-7075	NM_028028	Zswim1	339.25
chr8	122698910	122699310	27446	NM_028071	Cotl1	337.8
chr2	142754249	142754649	-61	NM_021335	Snrpb2	330.55
chr15	78704175	78704575	-70	NM_145929	Gga1	327.65
chr8	108455017	108455417	-48	NM_013477	Atp6v0d1	326.2
chr16	4795566	4795966	-5714	NM_029657	Mgrn1	324.75
chr16	95837664	95838064	26580	NM_011809	Ets2	318.95
chr5	143170800	143171200	864	NM_007393	Actb	318.95
chr13	30981355	30981755	-43	NM_025588	Exoc2	310.26
chr4	118603577	118603977	-2900	NM_011400	Slc2a1	310.26
chr2	91530759	91531159	-18	NM_172669	Ambra1	310.26
chr9	57652381	57652781	-20210	NM_019689	Arid3b	307.36
chr8	12703572	12703972	-31695	NM_198031	Tubgcp3	297.21
chr5	142692948	142693348	-23257	NM_172725	C330006K01Rik	294.31
chr5	138061315	138061715	4438	NM_153510	Pilra	291.41
chr19	11678413	11678813	-6764	NM_026835	Ms4a6d	288.51
chr2	170304570	170304970	102335	NM_001110152	Pfdn4	287.06
chr9	65349067	65349467	28773	NM_153119	Plekho2	285.61
chr6	124885333	124885733	-1983	NM_026988	Ptms	282.71

A PUER 0h

Promoter			Distal		
Sequence	motif	log p	Sequence	motif	log p
	SP1	-1529		SRF	-1807
	NFY	-619.6		SP1	-467.6
	CREB	-464.8			
	Ets	-337.6			
	SRF	-314.4			

B PUER 24h








Promoter			Distal		
Sequence	motif	log p	Sequence	motif	log p
	SP1	-2162		SRF	-1433
	NFY	-970		PU.1	-1014
	CREB	-507.6		SP1	-673.7
	GABP	-485.7			

Figure 10 Motif analysis of SRF ChIP-seq peak sequences in undifferentiated and differentiated PUER cells. HOMER *de novo* motif analysis of SRF ChIP-seq peaks in undifferentiated (PUER 0) (A) and differentiated (PUER 24) (B) was performed as described in Figure 4C. The motif analysis was performed on significant peaks found in promoter (-500bp to +500bp) and distal (excluding promoter) regions as indicated.

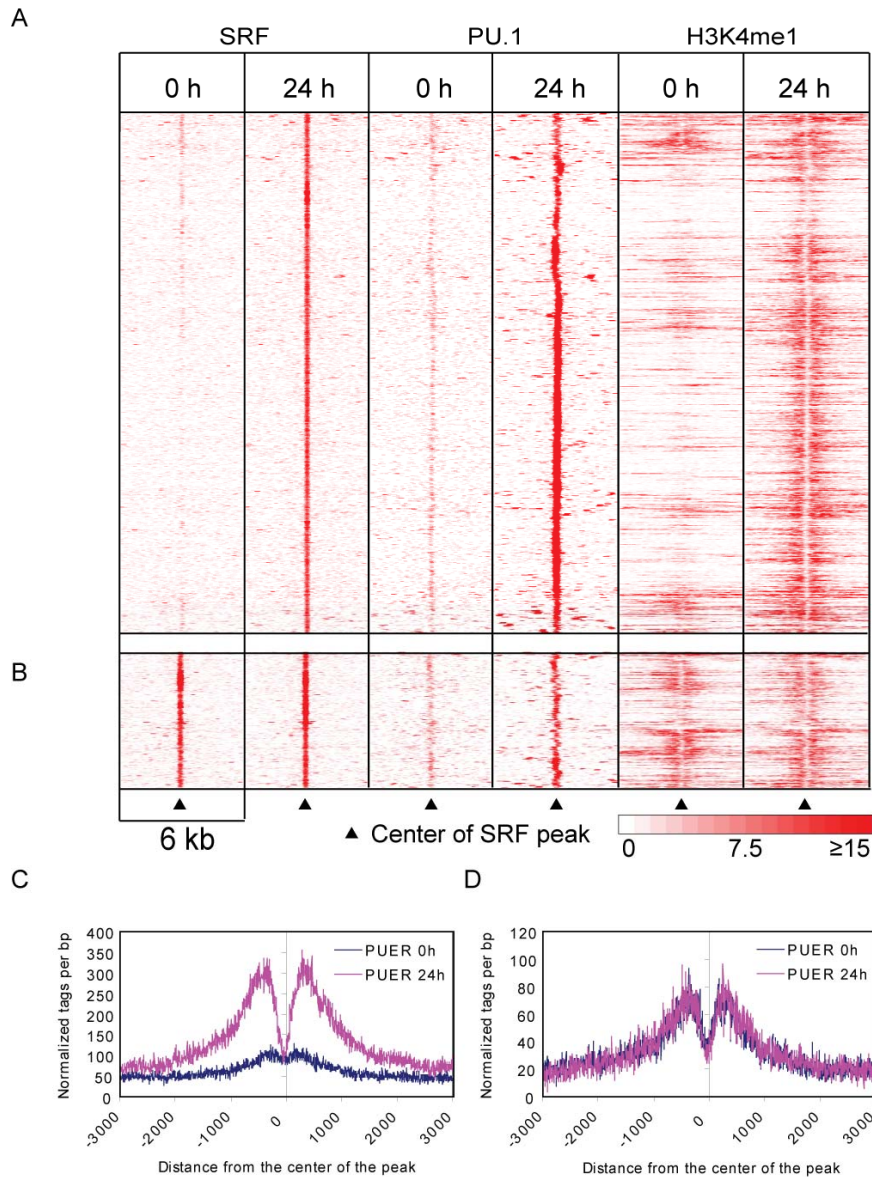


Figure 11 The H3K4me1 enhancer mark is acquired only at sites that gain both SRF and PU.1 during differentiation. (A and B) Heat map representation of SRF, PU.1, and H3K4me1 ChIP-seq distal peak (excluding area within ± 3 kb of TSS) data in undifferentiated (0h) and differentiated (24h tamoxifen treated) PUER cells. The data shown represent only those areas with SRF peaks at 24h that also gain PU.1. Each column represents a 6kb region centered on the location of SRF peaks after 24h of tamoxifen. Results were clustered and representative regions of the major binding patterns are shown. White represents areas of low numbers of sequence tags and red represents areas of high numbers of sequence tags. (C and D) Histogram analysis of the locations of the H3K4me1 marks relative to the center of the SRF peaks represented in (A) and (B), respectively.

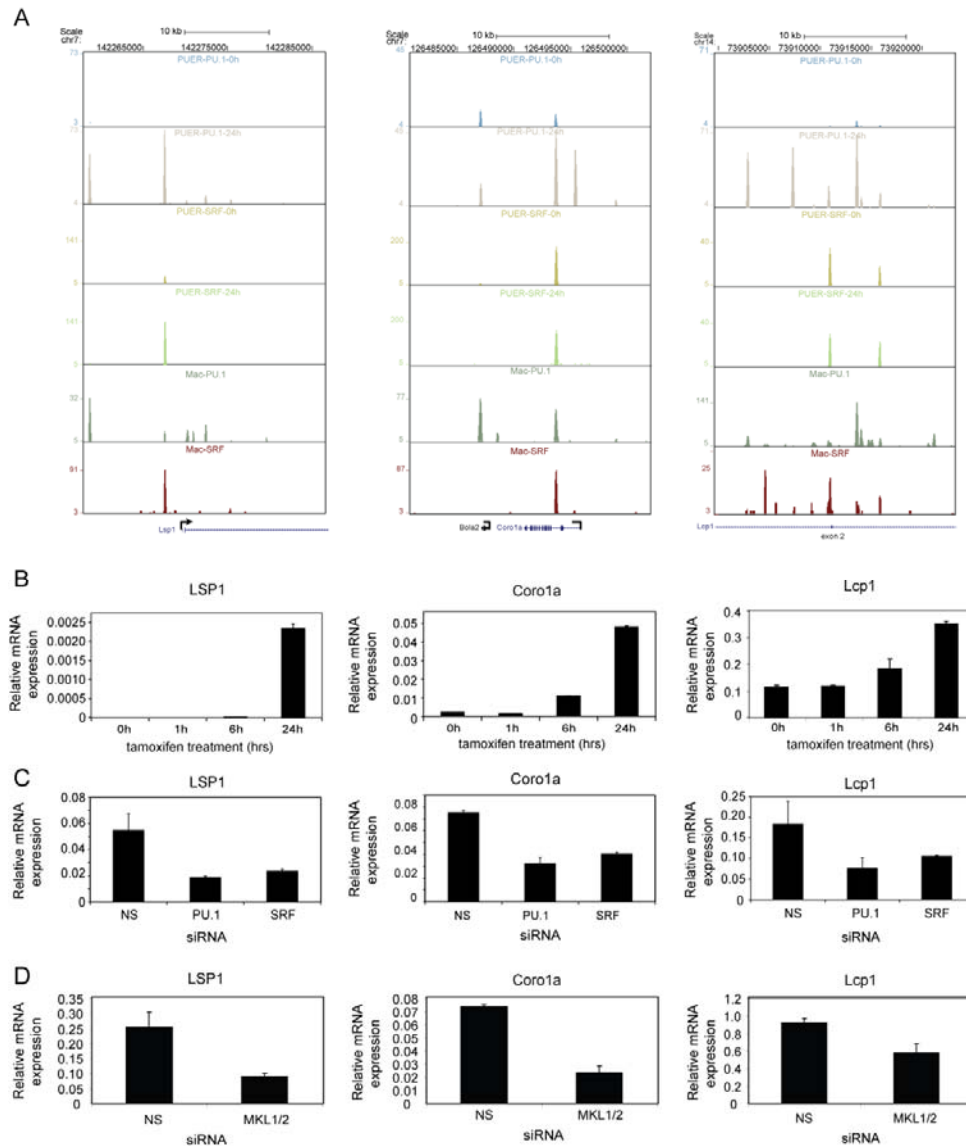


Figure 12 SRF, PU.1, and MKL regulate hematopoietic-specific cytoskeletal gene expression in macrophages. (A) UCSC genome browser images of PU.1 and SRF ChIP-seq peaks found near the *LSP1*, *Coro1a*, and *Lcp1* genes in undifferentiated (0h) and differentiated (24h tamoxifen treated) Puer cells and primary macrophages, as indicated. (B) qPCR analysis of the relative expression of *LSP1*, *Coro1a*, and *Lcp1* mRNAs (normalized to *GAPDH*) during Puer differentiation induced by tamoxifen. (C) qPCR analysis of the relative mRNA expression (normalized to *GAPDH*) of *LSP1*, *Coro1a*, and *Lcp1* mRNAs after treatment with non-specific (NS) control, PU.1, or SRF siRNAs for 30 hours in primary macrophages. (D) qPCR analysis of target genes as in (C) in primary macrophages treated with NS or MKL siRNAs for 48 hours. For panels (B-D), the data shown is representative of at least three independent experiments.

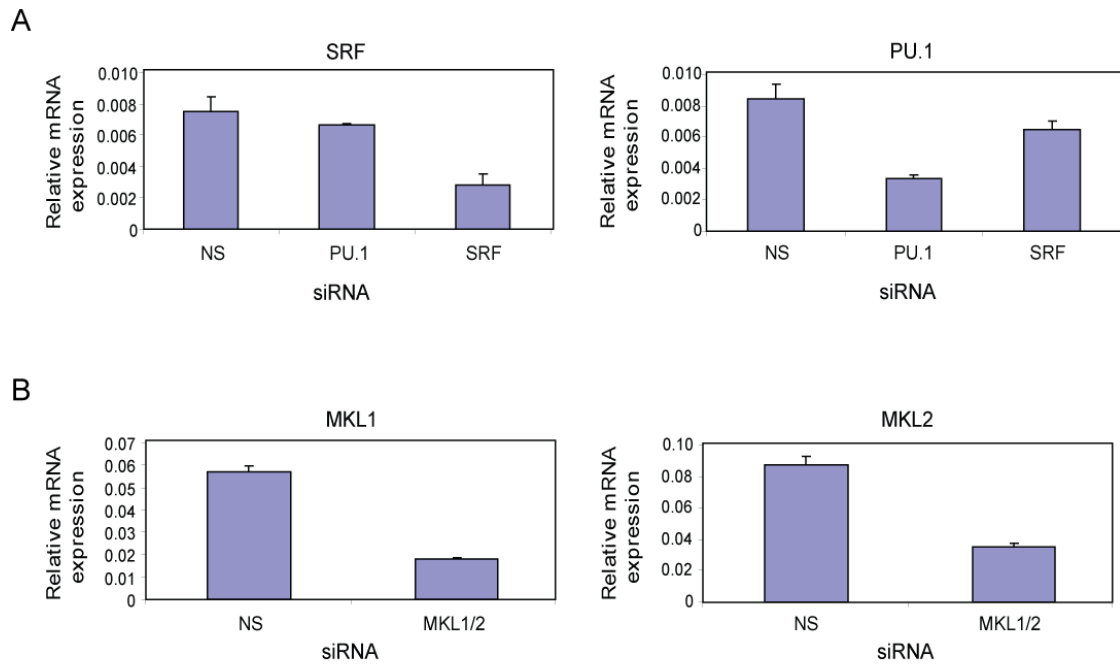


Figure 13 Confirmation of SRF, PU.1 and MKL1/2 knockdown in primary macrophages. qPCR analysis of *SRF* and *PU.1* (A) or *MKL1* and *MKL2* (B) mRNA expression in primary macrophages after treatment with the indicated siRNAs as described in Figure 12. The samples that are shown here are the same samples that are represented in Figure 11C and D.

Table 10 Transcription factors that are downregulated after SRF knockdown. The genes listed in this table are those transcription factors that are downregulated, based on the microarray analysis, in primary macrophages after treatment with SRF siRNA for 48 hours. Blue highlighting indicates those genes that also have significant SRF peaks associated with them according to the CHIP-seq data in primary macrophages.

NCBI Accession #	Gene Symbol	Fold downregulated
NM_010902	Nfe2l2	1.73
NM_007520	Bach1	1.56
NM_001077361	Fhl1	1.46
NM_010137	Epas1	1.44
NM_008667	Nab1	1.41
NM_008692	NFYc	1.36
NM_007953	Esrra	1.36
NM_001079513	ZBTB33	1.35
NM_181650	Prdm4	1.34
NM_133977	Trf	1.34
NM_007496	Zfx3	1.33
NM_183208	Zmiz1	1.33
NM_027901	Gtf3c2	1.32
NM_010828	Cited2	1.32

Chapter 3

Future Directions and Conclusion

Introduction

The work presented in Chapter 2 has successfully elucidated the binding pattern of SRF as well as the consequences of SRF or MKL knockdown on gene expression in macrophages, yet many questions still remain. A particularly interesting observation from the SRF ChIP-seq data is the relatively low percentage of SRF peaks that actually contain CArG boxes. In addition, the unexpected enrichment of SRF binding to distal genomic locations begs the question of exactly how SRF is able to regulate transcription from these sites. It is also unclear whether MKL or other cofactors, such as TCFs, are interacting with SRF at these distal sites and what their role may be in transcriptional regulation of cell-type versus non-cell-type specific gene expression. Finally, the functional consequences of SRF or MKL loss in macrophages have yet to be addressed. The goal of this section is to discuss the current hypotheses supported by the data presented in Chapter 2, as well as the experiments that could be used to answer the questions listed above that have arisen as a result of this study.

Understanding SRF association with CArG-less DNA

What is particularly surprising about the SRF ChIP-seq data presented in Chapter 2 is that not all of the binding sites that were identified by the analysis contained CArG motifs. One possibility for this result is that some of these sites have CArG-like sequences that deviate too much from the current definition of a CArG box (more than 1bp mismatch in the A-T core region),

which prevented their identification during the motif search. For example, visual inspection of some of the CArG-less peak sequences identified by the analysis had CArG-like sequences such as only one C or G flanking the AT core, a 7 nucleotide AT core instead of a 6 nucleotide AT core, or more than one C or G in the AT core while the 6bp spacing between the CC and GG was maintained. Several studies have shown that there are very stringent requirements for SRF binding to DNA because of the distance and spatial location of the nucleotides required for SRF homodimers to bind to DNA[50-53]. Additional studies suggest that the flanking sequences, in addition to the CArG box, are also very important for proper SRF-mediated expression of target genes[54]. Taken together, we hypothesize that there may be factor binding sites adjacent to these CArG-like sequences that could facilitate SRF binding to these non-optimal sites. To test this hypothesis, we propose that binding of SRF to these CArG-like sites would need to be independently determined using supershift assays with probes made from the peak sequences identified in the CHIP-seq assay, instead of random sequence, in order to maintain the genomic context that may be required for SRF binding to CArG-like sites. In addition to the direct binding assays, those CArG-like sites that are located in the promoter region could also be tested for function using wild-type and mutant target gene promoter luciferase reporter assays. If SRF binding is shown to occur at these CArG-like sites, then deletion or mutation analysis of the binding region itself should shed some light on the

requirements for SRF binding to non-consensus sites as well as the factors that may facilitate this binding.

Aside from those peak sequences that contain CArG-like binding sites, there is also a significant subset of peak sequences that have no recognizable CArG or CArG-like sequences in the SRF peak region. Interestingly, many of these peak sequences are those that contain motifs for known proximal promoter factors such as SP1. There are several possibilities to explain how ChIP of SRF leads to enrichment of these peak sequences: 1) SRF binds directly to a novel DNA sequence, 2) SRF associates with a complex composed of other DNA-binding transcription factors where DNA binding of SRF itself is not required, or 3) SRF is directly bound to DNA, but is also in complex with other DNA-binding transcription factors at a distant location leading to the purification of both directly bound and indirectly bound DNA during the ChIP process. Each of these potential mechanisms will be discussed below.

In an effort to rule out the possibility of SRF directly binding to novel DNA sequences or protein complexes, we performed supershift analyses using DNA probes comprised of sequences that were enriched in the ChIP-seq procedure, located in the proximal promoter region, but do not contain CArG boxes. For example, we performed supershift experiments using the peak sequence located at the proximal promoter (-34bp relative to the TSS) of the *Jmjd5* gene, which is also one of the most enriched sequences identified in

the SRF ChIP-seq analysis (Chapter 2, Table 7). Our results showed that SRF was not able to form any observable complexes on this sequence from *Jmjd5*, while the CArG-containing peak sequence from the β -actin promoter resulted in 100% of the complexes being supershifted by an antibody directed against SRF (data not shown). These experiments suggest that SRF is not directly binding to a novel DNA sequence or protein complex that would explain the relatively high enrichment of peak sequences that do not contain SRF binding motifs.

Because SRF does not appear to be directly binding to novel DNA sequences, we also investigated the possibility that SRF binds indirectly by forming complexes with other DNA binding factors in a CArG independent manner. *De novo* motif analysis of the CArG-less sequences identified in the primary macrophage SRF ChIP-seq experiment resulted in the enrichment of PU.1, CREB, and SP1 sites (data not shown), so we asked the question of whether or not any of these factors could be the key for CArG-less SRF association with DNA. SRF has already been shown to complex with SP1 and MyoD in muscle cells on the cardiac α -actin promoter, but this was found to be in a CArG and GC-box dependent manner[147]. Ternary complex formation of SRF and CREB on the Krox-20 promoter has also been observed in *Xenopus laevis*, but this complex also required intact CArG and CRE binding sites[148]. In contrast, PU.1 has been shown to be incapable of forming ternary complexes with SRF on the *egr1* promoter SREs and was also found to

directly bind SRF only in the context of a modified PU.1 protein that contained the Elk-1 B-box [48, 64]. In order to investigate the possibility of DNA independent binding, we performed co-immunoprecipitation experiments using macrophage cell extracts and antibodies directed against endogenous SRF, PU.1, SP1, and several SP1-related factors that can also bind GC-boxes. So far, we have failed to observe any interaction between SRF, PU.1, SP1, and KLF6, but additional experiments must be completed in order to confirm or rule out any interactions between SRF and CREB, SP3, or KLF13.

During further bioinformatic analysis of the SRF ChIP-seq data, we also observed that in many cases, the SP1, PU.1, and CREB sites identified in the CArG-less peak sequences were broadly distributed (\pm 100bp) relative to the center of the peaks (data not shown). In contrast, those peak sequences that do contain CArG boxes show the CArG motif primarily located within 25 base pairs of the center of the peak (Chapter 2, Fig. 6). We believe that the lack of close association of SP1, PU.1, and CREB motifs to the center of the SRF peaks (presumably the site of DNA binding) in addition to the negative interaction data suggests that complex formation between SRF with other DNA binding factors in the absence of direct SRF binding to DNA is not the reason for SRF association with CArG-less DNA sequences. Of course, we cannot rule out the possibility that a novel DNA binding or complex factor may be the reason for SRF association with CArG-less DNA. In order to address this question, we propose to purify the SRF complex using affinity purification

from macrophage cell nuclear extracts (RAW 264.7 cell line) followed by mass spectrophotometric analysis to identify any novel SRF binding partners.

An additional explanation for SRF association with CArG-less DNA stems from the dot plots shown in Chapter 2, Figure 5C-F that show a high correlation between enriched motifs and their location relative to the transcription start site. The fact that CArG motifs are enriched at distal regions relative to the TSS, in addition to the high association of SRF peaks with both PU.1 and the H3K4me1 enhancer mark led us to the hypothesis that SRF is functioning at enhancers through a DNA looping model[146]. In this model, DNA-bound factors at distal sites interact with protein complexes located at the proximal promoters of target genes forming loops in the DNA and leading to the enhancement of target gene expression. Such a model is attractive in this case because it would explain how SRF is associated with so many proximal promoter sequences that do not contain CArG boxes, as well as the reason for the large number of SRF peaks that are associated with the H3K4me1 enhancer mark.

So how does one prove such a complex model? The first step is to show that the distal and proximal promoter sequences are in fact spatially located near each other. One assay that can be used to show the close proximity of distant DNA elements is the chromosome conformation capture (3C) assay[149]. In this assay, cells are fixed with formaldehyde in the same fashion as for ChIP assay in order to fix the DNA/protein complexes in the

basal conformation. After isolation of the nuclei and incubation with detergent to solubilize the sample, the nuclei are incubated with an enzyme (usually a 6-base or 4-base cutter) to digest the chromatin. After inactivation of the enzyme, the digested chromatin is incubated with T4 DNA ligase, which ligates together the digested DNA fragments that are in close proximity to one another. Following ligation, the digested and ligated DNA fragments are purified and subjected to PCR using primers that have been designed such that one primer is in each fragment of interest and PCR product will only result if the desired fragments have been ligated to one another. Additional PCR primers are also designed in areas in between the two desired DNA fragments to serve as negative controls for the assay. A positive result is attained when the frequency of interaction (amount of PCR product) observed between the two DNA fragments of interest is significantly greater than the frequency of interaction seen for the intermediate DNA fragments used as negative controls after accounting for the efficiency of each PCR primer pair using a digested and ligated BAC clone encompassing the genomic region of interest.

Besides the basic 3C assay, related methods can also be used to identify long range interactions[150-151]. ChIP-3C incorporates the antibody affinity step used in the standard ChIP assay in conjunction with the 3C protocol in order to decrease the background associated with random ligation of the entire genome in the traditional 3C assay. Modifications of the 3C assay have also resulted in the development of the 4C and 5C assays which

use the amplification of ligated products and subsequent microarray hybridization or sequencing to identify novel interactions[152-153]. Additional methods, such as the fluorescence in situ hybridization (FISH) assay or a modified ChIP-assay using laser-fixed cells which only fixes protein to DNA, but not protein to protein may also be used to determine whether a specific DNA interaction results from the formation of protein complexes assembled at distant DNA sites[154-155].

The current disadvantage to all of the experimental methods outlined above is that they are extremely inefficient at identifying the many complex interactions that occur in the context of a cell. Because of the PCR-based results associated with each of the methods described, interactions can really only be identified for one target at a time and determining potential targets can prove to be extremely difficult because interactions with distal sequences may not be occurring with the nearest transcription start site, but rather with one or multiple sites located tens or hundreds of kilobases away, or even on different chromosomes. In order to fully understand the role of SRF at distal sites, an unbiased method for discovery of distal interactions must be employed. Recent work published by Lieberman-Aiden *et. al.* has described such an unbiased method (called Hi-C), which is a modification of the original 3C assay that employs labeling and enrichment of ligated fragments followed by high throughput sequencing similar to that used in the ChIP-seq experiments previously described[156]. Use of such a method will broadly expand our

understanding of the three-dimensional nature of transcriptional regulation, not only for SRF, but for any other DNA binding transcription factor.

Elucidating the mechanism of SRF regulation from distal sites

Even if we are able to identify long range interactions between direct SRF distal sequences and indirectly associated distal sequences through some form of chromatin capture assay, we still would not be able to comment on the actual regulatory function of the sequence. In order to test enhancer function of these distal SRF-bound sequences, we propose to perform luciferase reporter enhancer assays. This assay requires cloning the putative enhancer sequence at the 3' end of the coding region of a luciferase construct that is driven by a minimal promoter. If the sequence functions as an enhancer, then the relative expression of the reporter gene will be increased. It would be most interesting to use putative enhancers, such as those described in Chapter 2 (Figure 12), that have PU.1 binding sites in addition to the SRF binding sites, so that the contribution of each factor could be assessed in the enhancer assay using reporters containing mutations in the SRF and PU.1 binding sites of the enhancer fragment.

The enhancer assay described above is great if the result is positive, but interpreting a negative result could prove to be much more difficult. In addition to the obvious interpretation that the sequence is not an enhancer, we also must consider that because of the physical constraints of the partially chromatinized structure of a plasmid, it is possible that the putative enhancer

region may not be able to reach the proper orientation to interact with the proximal promoter and subsequently enhance gene expression. However, recent data using transgenic mice has suggested that almost all functional enhancers have the coactivator p300 present in the area[157]. As an alternative to the enhancer assay, I propose to perform ChIP-seq analysis for p300 in macrophages. Through comparison of the SRF and p300 ChIP-seq data sets, I expect to identify a subset of distal regions that are bound by both SRF and p300, which would presumably be functional enhancers under basal conditions.

Another interesting question regarding the discovery of SRF localization to distal regions is what known co-activators or novel factors may also be associated at these sites? At this time, I have not studied any of the TCF factors in this regard, but I have tried to perform ChIP assays for MKL1 in macrophages. So far, these attempts have not been successful using either the antibodies developed as a part of this study, or those that are commercially available, so other approaches will be required. In the absence of good antibodies, other members of the lab have had some success using an *in vivo* biotin tagging system, described by Beckett *et. al.*, which allows for protein precipitation using streptavidin coated beads for ChIP and subsequent sequencing analysis[158]. Future studies will focus on developing tagged versions of MKL1 and MKL2, as well as any TCF factors, as required, that can be used in this system to examine the genome-wide location of these factors.

In addition, any novel factors purified as components of the SRF complex, as described earlier, may also be subjects of genome-wide localization studies using either traditional or biotin-tagged ChIP-seq methods.

Control of cell-type and non-cell-type specific gene expression by SRF

Previous data in the lab has shown that the transcription of a subset of cell-type specific genes is mediated by the association of cell-type specific transcription factors at distal sites that are characterized by the presence of the H3K4 monomethyl mark[113]. What is particularly interesting about this study, is that when H3K4me1 ChIP-seq data sets from multiple cell types were analyzed, the enriched motifs that were associated with the distal monomethyl mark were for those factors that are required for lineage commitment of the particular cell type. For example, PU.1 motifs are enriched in the distal H3K4me1 sites in macrophages, while KLF4, Oct4, Sox2 and Esrrb motifs are enriched in the distal H3K4me1 sites in ES cells.

While other studies in the lab have focused on how cell type-specific transcription occurs from the point of view of the cell-type specific transcription factor, the current study provides insight to how a ubiquitous transcription factor is able to regulate both cell-type and non-cell-type specific gene expression. My proposed model is that during differentiation, expression of a cell-type specific transcription factor is induced (e.g. PU.1) which mediates the imprinting of a chromatin structure that enables the binding of factors and expression of genes that are required for the maintenance and function of the

designated cell type. Through this process, those genes that are required for responses in other cell types are sequestered, become inaccessible, and are not expressed. In support of this model, SRF binding to an enhancer region for the muscle specific factor, MyoD, has been observed in experiments using muscle cells, but MyoD is not expressed and I do not observe SRF binding to this enhancer region in macrophages. Cooper *et. al.* has also shown that SRF binding patterns differ between several human cell lines of different lineages, but this study is limited in scope because they used promoter microarrays to determine the SRF targets enriched in their ChIP assay and cannot account for the contribution of factor binding at distal sites[159].

In order to clarify the role of SRF in cell-type specific transcription, I propose to compare the genome-wide localization of SRF in multiple cell types using ChIP-seq. What I would expect, based on my model, is that the SRF localization would be the same for those SRF-dependent genes that are commonly expressed in all cell types, but that the SRF binding patterns observed at some promoters and distal sites would differ between cell types because they would correspond to those sites that regulate the expression of cell-type specific genes. I would also expect that the motifs that are co-enriched with CArG boxes at these distal sites would correlate with the cell-type specific transcription factors required for lineage commitment of that cell.

In addition to ChIP-seq experiments in different cell types, it would also be interesting to perform ChIP-seq in macrophages after inflammatory stimuli,

such as cytokines, lipopolysaccharide (LPS) or plpC. It is currently unclear whether SRF is always bound to all of its accessible sites, or whether it can be actively recruited to additional sites upon activation. Of particular interest is the subset of genes that are bound by SRF, but are not expressed under basal conditions according to the expression microarray data. Based on the proposed model described above, SRF and PU.1 will bind to the DNA at accessible sites, but the binding is not restricted to only those promoters or enhancers that are required for basal gene expression, but also include those regulatory regions of genes that may be induced in response to stimulation (e.g. LPS, virus, etc.). It would be particularly interesting to test this model by performing ChIP-seq analysis for SRF, PU.1, and p300 at different time points after stimulation to see the extent of the relocalization of these factors that occurs during the response. I hypothesize that the genome-wide localization of SRF and PU.1 will change very little at promoter or enhancer sites, but that p300 will redistribute to different promoter and distal sites in accordance with the gene expression required for the response.

Assaying the requirement of MKL1 and SRF for macrophage function through mouse models

This project originally started with the identification of MKL1 in macrophages through a yeast two-hybrid screen using the nuclear receptor, peroxisome proliferator activated receptor gamma (PPAR γ), as bait. Further experiments to observe endogenous interaction or alterations in known

transcriptional function caused by the direct or functional interaction of these two proteins failed, so the new focus of the project became to examine the role of MKL1 and SRF in macrophages and resulted in the work presented here.

Not long after the focus of my project became MKL1, two groups published mouse knockout models of the protein, which showed defects in mammary myoepithelial cell differentiation and resulted in premature involution of KO mothers and the inability to feed their pups[88-89]. Potential effects that MKL1 KO may have had on the immune system at that time had not been studied, so we requested, and were kindly granted, the opportunity to work with the mice developed in Dr. Stephen Morris' Laboratory.

Through a large number of studies that will not be described here (e.g. *in vitro* differentiation, phagocytosis and bacterial killing assays, atherosclerosis models, etc.) we concluded that there is little to no effect on macrophage differentiation and function as a result of the loss of MKL1. Based on gene expression profiling and the observation that many of the target genes identified by the siRNA experiments were not conserved in the MKL1 knockout, we believe that the lack of a macrophage phenotype in this model is the result of compensation for loss of MKL1 by MKL2. MKL2 is not upregulated with MKL1 loss, but is clearly expressed in macrophages and has been shown in other cell systems to be able to coactivate many of the same genes that are also targets of MKL1[72, 79]. Interestingly, some of the MKL1-dependent target genes that were identified in the siRNA microarray analysis

also showed decreases in mRNA expression in the MKL1 knockout, suggesting that MKL1 and MKL2 are largely redundant factors, but still hold some factor-specific functions that could be studied in greater detail. Due to the embryonic lethality of the MKL2 KO mouse and the compensation issues described here, any further studies to examine the role of MKL1/2 in macrophage function will require the development of a dual conditional knockout system.

Once we concluded that the MKL1 mice were not a viable model system to study macrophage function, we turned our attention to mouse models of SRF. Due to the catastrophic phenotype of systemic loss of SRF, *in vivo*, functional studies of SRF have depended on the conditional knockout of SRF in distinct cell types through the use of cell-type specific promoter driven CRE transgenes. Because SRF function in macrophages has not been assessed using a knockout model, we decided to generate a myeloid specific knockout using SRF fl/fl mice (a kind gift from J. Miano) crossed with mice harboring a CRE transgene under the control of the myeloid specific CD11b promoter (a kind gift of D. Cleveland)[101, 160-161]. Unfortunately after generation of these mice, we discovered that CRE expression in these mice had been silenced, thus preventing excision at the SRF locus.

As an alternative strategy, we have crossed the SRF fl/fl mice to mice carrying an Mx-CRE transgene. In contrast to the CD11b knockout strategy, SRF will not be knocked out during development, but will be induced following

activation of the Mx promoter through serial injection of plpC. This strategy is perfect in the sense that this transgene has been used extensively and is known to mediate almost total excision of floxed alleles in the macrophages, but may be problematic because Mx expression is not restricted to the myeloid lineage and will result in the deletion of SRF in other tissues, such as the liver, spleen, and total bone marrow, which could potentially be fatal based on other conditional knockout models of SRF[98].

Assuming that we are able to recover bone marrow from SRF deleted mice following plpC injection, we plan to initially characterize the cellular components of total bone marrow (e.g. myeloid/granulocyte precursors, B-cells, T-cells, erythrocytes, etc.) using fluorescence-activated cell sorting (FACS) of cells stained with lineage specific markers to highlight any specific lineage deficiencies that may result from SRF deletion. In addition, we will differentiate wild-type and knockout bone marrow *in vitro* using M-CSF, followed by FACS analysis for the mature macrophage marker, F4/80, to test whether mature macrophages can indeed be recovered. Along these same lines, we can also test the presence of mature resident or thioglycollate-elicited macrophages present in the peritoneal cavity following plpC injection. Characterization of the cell types recovered will be important because subsequent studies of macrophage function, such as phagocytosis, bacterial killing, chemotaxis and invasion/migration, will only be relevant if the wild-type and knockout macrophages recovered are in similar stages of development.

Otherwise, studies will have to be focused on the source of any developmental defect in the hematopoietic system that is observed.

Conclusion

The current study provides strong evidence for a role of the transcription factors, SRF and MKL, in the regulation of gene expression in macrophages. We have shown, in contrast to currently published work on SRF and MKL in other cell systems, that much of the SRF binding to the genome occurs at distal sites, as opposed to proximal promoter regions. Of particular interest is the observation that both SRF and MKL regulate the expression of hematopoietic-specific cytoskeletal genes and that SRF binding to these genes occurs, not at promoters, but at distal sites in association with the macrophage and B-cell-specific transcription factor, PU.1. Such association has important implications for both the transcriptional control of specialized macrophage functions, as well as the general mechanisms for cell-type versus non-cell-type specific gene regulation and should be the focus of further study.

References

1. Pollard, J.W., *Trophic macrophages in development and disease*. Nat Rev Immunol, 2009. **9**(4): p. 259-270.
2. Miano, J.M., *Serum response factor: toggling between disparate programs of gene expression*. J Mol Cell Cardiol, 2003. **35**(6): p. 577-93.
3. Miano, J.M., X. Long, and K. Fujiwara, *Serum response factor: master regulator of the actin cytoskeleton and contractile apparatus*. Am J Physiol Cell Physiol, 2007. **292**(1): p. C70-81.
4. Cen, B., A. Selvaraj, and R. Prywes, *Myocardin/MKL family of SRF coactivators: key regulators of immediate early and muscle specific gene expression*. J Cell Biochem, 2004. **93**(1): p. 74-82.
5. Disanza, A., et al., *Actin polymerization machinery: the finish line of signaling networks, the starting point of cellular movement*. Cell Mol Life Sci, 2005. **62**(9): p. 955-70.
6. Chen, H., B.W. Bernstein, and J.R. Bamburg, *Regulating actin-filament dynamics in vivo*. Trends Biochem Sci, 2000. **25**(1): p. 19-23.
7. Carlier, M.-F., *Control of actin dynamics*. Current Opinion in Cell Biology, 1998. **10**(1): p. 45-51.
8. Carlier, M.F. and D. Pantaloni, *Control of actin assembly dynamics in cell motility*. J Biol Chem, 2007. **282**(32): p. 23005-9.
9. Pollard, T.D., *Regulation of actin filament assembly by Arp2/3 complex and formins*. Annu Rev Biophys Biomol Struct, 2007. **36**: p. 451-77.
10. Ferrari, G., et al., *A coat protein on phagosomes involved in the intracellular survival of mycobacteria*. Cell, 1999. **97**(4): p. 435-47.

11. Jongstra-Bilen, J. and J. Jongstra, *Leukocyte-specific protein 1 (LSP1): a regulator of leukocyte emigration in inflammation*. Immunol Res, 2006. **35**(1-2): p. 65-74.
12. Ridley, A.J., *Regulation of macrophage adhesion and migration by Rho GTP-binding proteins*. J Microsc, 2008. **231**(3): p. 518-23.
13. Toksoz, D. and K.D. Merdek, *The Rho small GTPase: functions in health and disease*. Histol Histopathol, 2002. **17**(3): p. 915-27.
14. Allen, W.E., et al., *A role for Cdc42 in macrophage chemotaxis*. J Cell Biol, 1998. **141**(5): p. 1147-57.
15. Yang, L., L. Wang, and Y. Zheng, *Gene targeting of Cdc42 and Cdc42GAP affirms the critical involvement of Cdc42 in filopodia induction, directed migration, and proliferation in primary mouse embryonic fibroblasts*. Mol Biol Cell, 2006. **17**(11): p. 4675-85.
16. Wheeler, A.P., et al., *Rac1 and Rac2 regulate macrophage morphology but are not essential for migration*. J Cell Sci, 2006. **119**(Pt 13): p. 2749-57.
17. Ridley, A.J., *Rho proteins, PI 3-kinases, and monocyte/macrophage motility*. FEBS Letters, 2001. **498**(2-3): p. 168-171.
18. Niedergang, F. and P. Chavrier, *Signaling and membrane dynamics during phagocytosis: many roads lead to the phagos(R)ome*. Curr Opin Cell Biol, 2004. **16**(4): p. 422-8.
19. Niedergang, F. and P. Chavrier, *Regulation of phagocytosis by Rho GTPases*. Curr Top Microbiol Immunol, 2005. **291**: p. 43-60.
20. Caron, E. and A. Hall, *Identification of two distinct mechanisms of phagocytosis controlled by different Rho GTPases*. Science, 1998. **282**(5394): p. 1717-21.

21. Olazabal, I.M., et al., *Rho-kinase and myosin-II control phagocytic cup formation during CR, but not FcγR, phagocytosis*. *Curr Biol*, 2002. **12**(16): p. 1413-18.
22. Wang, Q.Q., et al., *Integrin beta 1 regulates phagosome maturation in macrophages through Rac expression*. *J Immunol*, 2008. **180**(4): p. 2419-28.
23. Erwig, L.P., et al., *Differential regulation of phagosome maturation in macrophages and dendritic cells mediated by Rho GTPases and ezrin-radixin-moesin (ERM) proteins*. *Proc Natl Acad Sci U S A*, 2006. **103**(34): p. 12825-30.
24. Shurin, G.V., et al., *Small rho GTPases regulate antigen presentation in dendritic cells*. *J Immunol*, 2005. **174**(6): p. 3394-400.
25. Lockman, K., et al., *Sphingosine 1-phosphate stimulates smooth muscle cell differentiation and proliferation by activating separate serum response factor co-factors*. *J Biol Chem*, 2004. **279**(41): p. 42422-30.
26. Mulder, J., et al., *Inhibition of RhoA-mediated SRF activation by p116Rip*. *FEBS Lett*, 2005. **579**(27): p. 6121-7.
27. Kuwahara, K., et al., *Muscle-specific signaling mechanism that links actin dynamics to serum response factor*. *Mol Cell Biol*, 2005. **25**(8): p. 3173-81.
28. Grosse, R., et al., *A role for VASP in RhoA-Diaphanous signalling to actin dynamics and SRF activity*. *EMBO J*, 2003. **22**(12): p. 3050-61.
29. Settleman, J., *A nuclear MAL-function links Rho to SRF*. *Mol Cell*, 2003. **11**(5): p. 1121-3.
30. Miralles, F., et al., *Actin dynamics control SRF activity by regulation of its coactivator MAL*. *Cell*, 2003. **113**(3): p. 329-42.

31. Liu, H.W., et al., *The RhoA/Rho kinase pathway regulates nuclear localization of serum response factor*. *Am J Respir Cell Mol Biol*, 2003. **29**(1): p. 39-47.
32. Murai, K. and R. Treisman, *Interaction of serum response factor (SRF) with the Elk-1 B box inhibits RhoA-actin signaling to SRF and potentiates transcriptional activation by Elk-1*. *Mol Cell Biol*, 2002. **22**(20): p. 7083-92.
33. Geneste, O., J.W. Copeland, and R. Treisman, *LIM kinase and Diaphanous cooperate to regulate serum response factor and actin dynamics*. *J Cell Biol*, 2002. **157**(5): p. 831-8.
34. Gineitis, D. and R. Treisman, *Differential usage of signal transduction pathways defines two types of serum response factor target gene*. *J Biol Chem*, 2001. **276**(27): p. 24531-9.
35. Mack, C.P., et al., *Smooth muscle differentiation marker gene expression is regulated by RhoA-mediated actin polymerization*. *J Biol Chem*, 2001. **276**(1): p. 341-7.
36. Spencer, J.A., M.L. Major, and R.P. Misra, *Basic fibroblast growth factor activates serum response factor gene expression by multiple distinct signaling mechanisms*. *Mol Cell Biol*, 1999. **19**(6): p. 3977-88.
37. Montaner, S., et al., *Activation of serum response factor by RhoA is mediated by the nuclear factor-kappaB and C/EBP transcription factors*. *J Biol Chem*, 1999. **274**(13): p. 8506-15.
38. Carnac, G., et al., *RhoA GTPase and serum response factor control selectively the expression of MyoD without affecting Myf5 in mouse myoblasts*. *Mol Biol Cell*, 1998. **9**(7): p. 1891-902.
39. Sahai, E., A.S. Alberts, and R. Treisman, *RhoA effector mutants reveal distinct effector pathways for cytoskeletal reorganization, SRF activation and transformation*. *EMBO J*, 1998. **17**(5): p. 1350-61.

40. Hill, C.S., J. Wynne, and R. Treisman, *The Rho family GTPases RhoA, Rac1, and CDC42Hs regulate transcriptional activation by SRF*. Cell, 1995. **81**(7): p. 1159-70.
41. Shore, P. and A.D. Sharrocks, *The MADS-box family of transcription factors*. Eur J Biochem, 1995. **229**(1): p. 1-13.
42. Treisman, R., *Identification and purification of a polypeptide that binds to the c-fos serum response element*. EMBO J, 1987. **6**(9): p. 2711-7.
43. Prywes, R. and R.G. Roeder, *Purification of the c-fos enhancer-binding protein*. Mol Cell Biol, 1987. **7**(10): p. 3482-9.
44. Treisman, R., *Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors*. Cell, 1986. **46**(4): p. 567-74.
45. Minty, A. and L. Kedes, *Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif*. Mol Cell Biol, 1986. **6**(6): p. 2125-36.
46. Phan-Dinh-Tuy, F., et al., *The 'CC.Ar.GG' box. A protein-binding site common to transcription-regulatory regions of the cardiac actin, c-fos and interleukin-2 receptor genes*. Eur J Biochem, 1988. **173**(3): p. 507-15.
47. Kitabayashi, I., et al., *Two cis-regulatory elements that mediate different signaling pathways for serum-dependent activation of the junB gene*. J Biol Chem, 1993. **268**(19): p. 14482-9.
48. Watson, D.K., et al., *FLI1 and EWS-FLI1 function as ternary complex factors and ELK1 and SAP1a function as ternary and quaternary complex factors on the Egr1 promoter serum response elements*. Oncogene, 1997. **14**(2): p. 213-21.

49. Norman, C. and R. Treisman, *Analysis of serum response element function in vitro*. Cold Spring Harb Symp Quant Biol, 1988. **53 Pt 2**: p. 719-26.
50. Pellegrini, L., S. Tan, and T.J. Richmond, *Structure of serum response factor core bound to DNA*. Nature, 1995. **376**(6540): p. 490-8.
51. Leung, S. and N.G. Miyamoto, *Point mutational analysis of the human c-fos serum response factor binding site*. Nucleic Acids Res, 1989. **17**(3): p. 1177-95.
52. Stepanek, J., et al., *C-->G base mutations in the CArG box of c-fos serum response element alter its bending flexibility. Consequences for core-SRF recognition*. FEBS J, 2007. **274**(9): p. 2333-48.
53. Huet, A., et al., *Mechanism of binding of serum response factor to serum response element*. FEBS J, 2005. **272**(12): p. 3105-19.
54. Chang, P.S., et al., *Muscle specificity encoded by specific serum response factor-binding sites*. J Biol Chem, 2001. **276**(20): p. 17206-12.
55. Sartorelli, V., K.A. Webster, and L. Kedes, *Muscle-specific expression of the cardiac alpha-actin gene requires MyoD1, CArG-box binding factor, and Sp1*. Genes Dev, 1990. **4**(10): p. 1811-22.
56. Nishida, W., et al., *A triad of serum response factor and the GATA and NK families governs the transcription of smooth and cardiac muscle genes*. J Biol Chem, 2002. **277**(9): p. 7308-17.
57. Nakamura, M., et al., *Transcriptional activation of beta-tropomyosin mediated by serum response factor and a novel Barx homologue, Barx1b, in smooth muscle cells*. J Biol Chem, 2001. **276**(21): p. 18313-20.
58. Sepulveda, J.L., et al., *Combinatorial expression of GATA4, Nkx2-5, and serum response factor directs early cardiac gene activity*. J Biol Chem, 2002. **277**(28): p. 25775-82.

59. Phiel, C.J., et al., *Differential binding of an SRF/NK-2/MEF2 transcription factor complex in normal versus neoplastic smooth muscle tissues*. J Biol Chem, 2001. **276**(37): p. 34637-50.
60. Buchwalter, G., C. Gross, and B. Wasylyk, *Ets ternary complex transcription factors*. Gene, 2004. **324**: p. 1-14.
61. Treisman, R., R. Marais, and J. Wynne, *Spatial flexibility in ternary complexes between SRF and its accessory proteins*. EMBO J, 1992. **11**(12): p. 4631-40.
62. Verger, A. and M. Duterque-Coquillaud, *When Ets transcription factors meet their partners*. Bioessays, 2002. **24**(4): p. 362-70.
63. Yang, S.H., et al., *The mechanism of phosphorylation-inducible activation of the ETS-domain transcription factor Elk-1*. EMBO J, 1999. **18**(20): p. 5666-74.
64. Shore, P. and A.D. Sharrocks, *The transcription factors Elk-1 and serum response factor interact by direct protein-protein contacts mediated by a short region of Elk-1*. Mol Cell Biol, 1994. **14**(5): p. 3283-91.
65. Cruzalegui, F.H., E. Cano, and R. Treisman, *ERK activation induces phosphorylation of Elk-1 at multiple S/T-P motifs to high stoichiometry*. Oncogene, 1999. **18**(56): p. 7948-57.
66. L'Honore, A., et al., *MyoD distal regulatory region contains an SRF binding CArG element required for MyoD expression in skeletal myoblasts and during muscle regeneration*. Mol Biol Cell, 2003. **14**(5): p. 2151-62.
67. Kim, H.J., J.H. Kim, and J.W. Lee, *Steroid receptor coactivator-1 interacts with serum response factor and coactivates serum response element-mediated transactivations*. J Biol Chem, 1998. **273**(44): p. 28564-7.

68. Jung, D.J., et al., *Novel transcription coactivator complex containing activating signal cointegrator 1*. Mol Cell Biol, 2002. **22**(14): p. 5203-11.
69. Hill, C.S., J. Wynne, and R. Treisman, *Serum-regulated transcription by serum response factor (SRF): a novel role for the DNA binding domain*. EMBO J, 1994. **13**(22): p. 5421-32.
70. Sotiropoulos, A., et al., *Signal-regulated activation of serum response factor is mediated by changes in actin dynamics*. Cell, 1999. **98**(2): p. 159-69.
71. Vickers, E.R., et al., *Ternary complex factor-serum response factor complex-regulated gene activity is required for cellular proliferation and inhibition of apoptotic cell death*. Mol Cell Biol, 2004. **24**(23): p. 10340-51.
72. Selvaraj, A. and R. Prywes, *Megakaryoblastic leukemia-1/2, a transcriptional co-activator of serum response factor, is required for skeletal myogenic differentiation*. J Biol Chem, 2003. **278**(43): p. 41977-87.
73. Du, K.L., et al., *Megakaryoblastic leukemia factor-1 transduces cytoskeletal signals and induces smooth muscle cell differentiation from undifferentiated embryonic stem cells*. J Biol Chem, 2004. **279**(17): p. 17578-86.
74. Wang, D., et al., *Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor*. Cell, 2001. **105**(7): p. 851-62.
75. Nagase, T., et al., *Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro*. DNA Res, 2000. **7**(1): p. 65-73.
76. Ma, Z., et al., *Fusion of two novel genes, RBM15 and MKL1, in the t(1;22)(p13;q13) of acute megakaryoblastic leukemia*. Nat Genet, 2001. **28**(3): p. 220-1.

77. Wang, D.Z. and E.N. Olson, *Control of smooth muscle development by the myocardin family of transcriptional coactivators*. *Curr Opin Genet Dev*, 2004. **14**(5): p. 558-66.
78. Sasazuki, T., et al., *Identification of a novel transcriptional activator, BSAC, by a functional cloning to inhibit tumor necrosis factor-induced cell death*. *J Biol Chem*, 2002. **277**(32): p. 28853-60.
79. Cen, B., et al., *Megakaryoblastic leukemia 1, a potent transcriptional coactivator for serum response factor (SRF), is required for serum induction of SRF target genes*. *Mol Cell Biol*, 2003. **23**(18): p. 6597-608.
80. Hinson, J.S., et al., *Smooth muscle cell-specific transcription is regulated by nuclear localization of the myocardin-related transcription factors*. *Am J Physiol Heart Circ Physiol*, 2007. **292**(2): p. H1170-80.
81. Posern, G., et al., *Mutant actins that stabilise F-actin use distinct mechanisms to activate the SRF coactivator MAL*. *EMBO J*, 2004. **23**(20): p. 3973-83.
82. Vartiainen, M.K., et al., *Nuclear actin regulates dynamic subcellular localization and activity of the SRF cofactor MAL*. *Science*, 2007. **316**(5832): p. 1749-52.
83. Wang, Z., et al., *Myocardin and ternary complex factors compete for SRF to control smooth muscle gene expression*. *Nature*, 2004. **428**(6979): p. 185-9.
84. Arsenian, S., et al., *Serum response factor is essential for mesoderm formation during mouse embryogenesis*. *EMBO J*, 1998. **17**(21): p. 6289-99.
85. Li, S., et al., *The serum response factor coactivator myocardin is required for vascular smooth muscle development*. *Proc Natl Acad Sci U S A*, 2003. **100**(16): p. 9366-70.
86. Li, J., et al., *Myocardin-related transcription factor B is required in cardiac neural crest for smooth muscle differentiation and*

- cardiovascular development*. Proc Natl Acad Sci U S A, 2005. **102**(25): p. 8916-21.
87. Oh, J., J.A. Richardson, and E.N. Olson, *Requirement of myocardin-related transcription factor-B for remodeling of branchial arch arteries and smooth muscle differentiation*. Proc Natl Acad Sci U S A, 2005. **102**(42): p. 15122-7.
 88. Li, S., et al., *Requirement of a myocardin-related transcription factor for development of mammary myoepithelial cells*. Mol Cell Biol, 2006. **26**(15): p. 5797-808.
 89. Sun, Y., et al., *Acute myeloid leukemia-associated Mkl1 (Mrtf-a) is a key regulator of mammary gland function*. Mol Cell Biol, 2006. **26**(15): p. 5809-26.
 90. Cheng, E.C., et al., *Role for MKL1 in megakaryocytic maturation*. Blood, 2009. **113**(12): p. 2826-34.
 91. Ayadi, A., et al., *Net-targeted mutant mice develop a vascular phenotype and up-regulate egr-1*. EMBO J, 2001. **20**(18): p. 5139-52.
 92. Costello, P.S., et al., *Ternary complex factor SAP-1 is required for Erk-mediated thymocyte positive selection*. Nat Immunol, 2004. **5**(3): p. 289-98.
 93. Cesari, F., et al., *Mice deficient for the ets transcription factor elk-1 show normal immune responses and mildly impaired neuronal gene activation*. Mol Cell Biol, 2004. **24**(1): p. 294-305.
 94. Wiebel, F.F., et al., *Generation of mice carrying conditional knockout alleles for the transcription factor SRF*. Genesis, 2002. **32**(2): p. 124-6.
 95. Niu, Z., et al., *Conditional mutagenesis of the murine serum response factor gene blocks cardiogenesis and the transcription of downstream gene targets*. J Biol Chem, 2005. **280**(37): p. 32531-8.

96. Niu, Z., et al., *Serum response factor orchestrates nascent sarcomerogenesis and silences the biomineralization gene program in the heart*. Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17824-9.
97. Koegel, H., et al., *Loss of serum response factor in keratinocytes results in hyperproliferative skin disease in mice*. J Clin Invest, 2009. **119**(4): p. 899-910.
98. Sun, K., et al., *Hepatocyte expression of serum response factor is essential for liver function, hepatocyte proliferation and survival, and postnatal body growth in mice*. Hepatology, 2009. **49**(5): p. 1645-54.
99. Li, S., et al., *Requirement for serum response factor for skeletal muscle growth and maturation revealed by tissue-specific gene deletion in mice*. Proc Natl Acad Sci U S A, 2005. **102**(4): p. 1082-7.
100. Latasa, M.U., et al., *Delayed liver regeneration in mice lacking liver serum response factor*. Am J Physiol Gastrointest Liver Physiol, 2007. **292**(4): p. G996-G1001.
101. Ramanan, N., et al., *SRF mediates activity-induced gene expression and synaptic plasticity but not neuronal viability*. Nat Neurosci, 2005. **8**(6): p. 759-67.
102. Alberti, S., et al., *Neuronal migration in the murine rostral migratory stream requires serum response factor*. Proc Natl Acad Sci U S A, 2005. **102**(17): p. 6148-53.
103. Fleige, A., et al., *Serum response factor contributes selectively to lymphocyte development*. J Biol Chem, 2007. **282**(33): p. 24320-8.
104. Wickramasinghe, S.R., et al., *Serum response factor mediates NGF-dependent target innervation by embryonic DRG sensory neurons*. Neuron, 2008. **58**(4): p. 532-45.
105. Schrott, G., et al., *Serum response factor is required for immediate-early gene activation yet is dispensable for proliferation of embryonic stem cells*. Mol Cell Biol, 2001. **21**(8): p. 2933-43.

106. Schrott, G., et al., *Serum response factor is crucial for actin cytoskeletal organization and focal adhesion assembly in embryonic stem cells*. J Cell Biol, 2002. **156**(4): p. 737-50.
107. Weinhold, B., et al., *Srf(-/-) ES cells display non-cell-autonomous impairment in mesodermal differentiation*. EMBO J, 2000. **19**(21): p. 5835-44.
108. Escalante, R., et al., *The MADS-box gene srfA is expressed in a complex pattern under the control of alternative promoters and is essential for different aspects of Dictyostelium development*. Dev Biol, 2001. **235**(2): p. 314-29.
109. Fraser, A.G., et al., *Functional genomic analysis of C. elegans chromosome I by systematic RNA interference*. Nature, 2000. **408**(6810): p. 325-30.
110. Guillemin, K., et al., *The pruned gene encodes the Drosophila serum response factor and regulates cytoplasmic outgrowth during terminal branching of the tracheal system*. Development, 1996. **122**(5): p. 1353-62.
111. Montagne, J., et al., *The Drosophila Serum Response Factor gene is required for the formation of intervein tissue of the wing and is allelic to blistered*. Development, 1996. **122**(9): p. 2589-97.
112. Han, Z., et al., *A myocardin-related transcription factor regulates activity of serum response factor in Drosophila*. Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12567-72.
113. Heinz, S., et al., *In preparation*. 2009.
114. Gupta, P., et al., *PU.1 and Partners: Regulation of Hematopoietic Stem Cell Fate in Normal and Malignant Hematopoiesis*. J Cell Mol Med, 2009.
115. Serbina, N.V., et al., *Monocyte-mediated defense against microbial pathogens*. Annu Rev Immunol, 2008. **26**: p. 421-52.

116. Gordon, S., *The macrophage: past, present and future*. Eur J Immunol, 2007. **37 Suppl 1**: p. S9-17.
117. Mosser, D.M. and J.P. Edwards, *Exploring the full spectrum of macrophage activation*. Nat Rev Immunol, 2008. **8**(12): p. 958-69.
118. Ridley, A.J., *Rho GTPases and cell migration*. J Cell Sci, 2001. **114**(Pt 15): p. 2713-22.
119. Vicente-Manzanares, M. and F. Sanchez-Madrid, *Role of the cytoskeleton during leukocyte responses*. Nat Rev Immunol, 2004. **4**(2): p. 110-22.
120. Chimini, G. and P. Chavrier, *Function of Rho family proteins in actin dynamics during phagocytosis and engulfment*. Nat Cell Biol, 2000. **2**(10): p. E191-6.
121. Sun, Q., et al., *Defining the mammalian CArGome*. Genome Res, 2006. **16**(2): p. 197-207.
122. Wang, D.Z., et al., *Potential of serum response factor activity by a family of myocardin-related transcription factors*. Proc Natl Acad Sci U S A, 2002. **99**(23): p. 14855-60.
123. Posern, G., A. Sotiropoulos, and R. Treisman, *Mutant actins demonstrate a role for unpolymerized actin in control of transcription by serum response factor*. Mol Biol Cell, 2002. **13**(12): p. 4167-78.
124. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery*. Genome Biol, 2003. **4**(5): p. P3.
125. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.

126. Robertson, G., et al., *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing*. Nat Methods, 2007. **4**(8): p. 651-7.
127. Benner, C., et. al., *In preparation*. 2009.
128. Kent, W.J., et al., *The human genome browser at UCSC*. Genome Res, 2002. **12**(6): p. 996-1006.
129. Marecki, S. and M.J. Fenton, *PU.1/Interferon Regulatory Factor interactions: mechanisms of transcriptional regulation*. Cell Biochem Biophys, 2000. **33**(2): p. 127-48.
130. Kanno, Y., et al., *Immune cell-specific amplification of interferon signaling by the IRF-4/8-PU.1 complex*. J Interferon Cytokine Res, 2005. **25**(12): p. 770-9.
131. Heintzman, N.D., et al., *Histone modifications at human enhancers reflect global cell-type-specific gene expression*. Nature, 2009. **459**(7243): p. 108-12.
132. Heintzman, N.D., et al., *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome*. Nat Genet, 2007. **39**(3): p. 311-8.
133. Bonifer, C., et al., *How transcription factors program chromatin--lessons from studies of the regulation of myeloid-specific genes*. Semin Immunol, 2008. **20**(4): p. 257-63.
134. Heintzman, N.D. and B. Ren, *The gateway to transcription: identifying, characterizing and understanding promoters in the eukaryotic genome*. Cell Mol Life Sci, 2007. **64**(4): p. 386-400.
135. Mellor, J., P. Dudek, and D. Clynes, *A glimpse into the epigenetic landscape of gene regulation*. Curr Opin Genet Dev, 2008. **18**(2): p. 116-22.

136. Walsh, J.C., et al., *Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates*. *Immunity*, 2002. **17**(5): p. 665-76.
137. Laslo, P., et al., *Multilineage transcriptional priming and determination of alternate hematopoietic cell fates*. *Cell*, 2006. **126**(4): p. 755-66.
138. Pieters, J., *Coronin 1 in innate immunity*. *Subcell Biochem*, 2008. **48**: p. 116-23.
139. Lin, C.S., et al., *Human plastin genes. Comparative gene structure, chromosome location, and differential expression in normal and neoplastic cells*. *J Biol Chem*, 1993. **268**(4): p. 2781-92.
140. Iwasaki, K., et al., *Rho/Rho-associated kinase signal regulates myogenic differentiation via myocardin-related transcription factor-A/Smad-dependent transcription of the Id3 gene*. *J Biol Chem*, 2008. **283**(30): p. 21230-41.
141. Morita, T., T. Mayanagi, and K. Sobue, *Dual roles of myocardin-related transcription factors in epithelial mesenchymal transition via slug induction and actin remodeling*. *J Cell Biol*, 2007. **179**(5): p. 1027-42.
142. Spencer, J.A. and R.P. Misra, *Expression of the serum response factor gene is regulated by serum response factor binding sites*. *J Biol Chem*, 1996. **271**(28): p. 16535-43.
143. Li, L., et al., *Evidence for serum response factor-mediated regulatory networks governing SM22alpha transcription in smooth, skeletal, and cardiac muscle cells*. *Dev Biol*, 1997. **187**(2): p. 311-21.
144. Kim, S., et al., *A serum response factor-dependent transcriptional regulatory program identifies distinct smooth muscle cell sublineages*. *Mol Cell Biol*, 1997. **17**(4): p. 2266-78.
145. Shen, X., et al., *The correlations of the function and positional distribution of the cis-elements CArG around the TSS in the genes of *Mus musculus**. *Genome*, 2009. **52**(3): p. 217-21.

146. Farnham, P.J., *Insights from genomic profiling of transcription factors*. Nat Rev Genet, 2009. **10**(9): p. 605-16.
147. Biesiada, E., et al., *Myogenic basic helix-loop-helix proteins and Sp1 interact as components of a multiprotein transcriptional complex required for activity of the human cardiac alpha-actin promoter*. Mol Cell Biol, 1999. **19**(4): p. 2577-84.
148. Watanabe, T., et al., *Functional role of a novel ternary complex comprising SRF and CREB in expression of Krox-20 in early embryos of Xenopus laevis*. Dev Biol, 2005. **277**(2): p. 508-21.
149. Dekker, J., et al., *Capturing chromosome conformation*. Science, 2002. **295**(5558): p. 1306-11.
150. Simonis, M., J. Kooren, and W. de Laat, *An evaluation of 3C-based methods to capture DNA interactions*. Nat Methods, 2007. **4**(11): p. 895-901.
151. Fullwood, M.J. and Y. Ruan, *ChIP-based methods for the identification of long-range chromatin interactions*. J Cell Biochem, 2009. **107**(1): p. 30-9.
152. Zhao, Z., et al., *Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions*. Nat Genet, 2006. **38**(11): p. 1341-7.
153. Dostie, J., et al., *Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements*. Genome Res, 2006. **16**(10): p. 1299-309.
154. Hu, Q., et al., *Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules*. Proc Natl Acad Sci U S A, 2008. **105**(49): p. 19199-204.

155. Nagaich, A.K. and G.L. Hager, *UV laser cross-linking: a real-time assay to study dynamic protein/DNA interactions during chromatin remodeling*. Sci STKE, 2004. **2004**(256): p. p113.
156. Lieberman-Aiden, E., et al., *Comprehensive mapping of long-range interactions reveals folding principles of the human genome*. Science, 2009. **326**(5950): p. 289-93.
157. Visel, A., et al., *ChIP-seq accurately predicts tissue-specific activity of enhancers*. Nature, 2009. **457**(7231): p. 854-8.
158. Beckett, D., E. Kovaleva, and P.J. Schatz, *A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation*. Protein Sci, 1999. **8**(4): p. 921-9.
159. Cooper, S.J., et al., *Serum response factor binding sites differ in three human cell types*. Genome Res, 2007. **17**(2): p. 136-44.
160. Boillee, S., et al., *Onset and progression in inherited ALS determined by motor neurons and microglia*. Science, 2006. **312**(5778): p. 1389-92.
161. Miano, J.M., et al., *Restricted inactivation of serum response factor to the cardiovascular system*. Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17132-7.