Regulation of Cardiac Gene Expression by Transcriptional and Epigenetic Mechanisms and Identification of a Novel Chromatin Remodeling Factor

DISSERTATION

zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften
(Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von
Dipl.-Biochem. Jenny Schlesinger
aus Berlin
September 2011

1. Gutachter: Prof. Dr. T. Schmülling,
Institut für Angewandte Genetik, Freie Universität Berlin,
Albrecht-Thaer-Weg 6, D-14195 Berlin

2. Gutachter: Prof. Dr. Silke R. Sperling,
Max-Planck-Institut für molekulare Genetik,
Ihnestr. 73, D-14195 Berlin

Disputation am 13.12.2011

Für meine Eltern

CONTENT

1 LIST OF MANUSCRIPTS ENCLOSED IN THIS THESIS

Manuscript 1

The Cardiac Transcription Network Modulated by Gata4, Mef2a, Nkx2.5, Srf, Histone Modifications, and MicroRNAs.

Schlesinger J*, Schueler M*, Grunert M*, Fischer JJ*, Zhang Q, Krueger T, Lange M, Tönjes M, Dunkel I, Sperling SR. PLoS Genet. 2011 Feb;7(2).

Doi:10.1371/journal.pgen.1001313

Manuscript 2

Dynamics of Srf, p300 and histone modifications during cardiac maturation in mouse.

Schueler M*, Zhang Q*, Schlesinger J, Tönjes M, Sperling SR.

In preparation

Manuscript 3

Evaluation of the LightCycler 1536 Instrument for high-throughput quantitative real-time PCR.

Schlesinger J, Tönjes M, Schueler M, Zhang Q, Dunkel I, Sperling SR. Methods. 2010 Apr;50(4):S19-22.

Doi: 10.1016/j.ymeth.2010.01.007

Manuscript 4

Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex.

Lange M, Kaynak B, Forster UB, Tönjes M, Fischer JJ, Grimm C, Schlesinger J, Just S, Dunkel I, Krueger T, Mebus S, Lehrach H, Lurz R, Gobom J, Rottbauer W, Abdelilah-Seyfried S, Sperling S. Genes Dev. 2008 Sep 1;22(17):2370-84.

Doi: 10.1101/gad.471408

2 INTRODUCTION

2.1 Transcriptional regulation in eukaryotes

In 1958, Francis Crick first formulated the theory that has nowadays become the central dogma of molecular biology. It defines the transfer of the genetic information encoded in the DNA sequence into a functional gene product.[1] Gene expression is the process of transcription of genes into RNAs. These are either, in case of mRNAs translated into proteins which are the fundamental functional unit of every cell, or directly utilized like tRNAs, rRNAs or microRNAs. The mechanisms that control the transcription of RNAs have been the subject of many genetic, biochemical and computational studies.[2-6] Thus, a full understanding of transcriptional regulation is today beyond our reach.

In eukaryotes, RNA polymerase II (Pol II), which is composed of twelve different subunits catalyzes the transcription of protein-coding genes. The entire Pol II transcription machinery comprises almost 60 different general transcription factors and co-activators.[7] General transcription factors (e.g. TFIIB, TFIID, TATA binding protein)[8] mediate promoter recognition and open double-stranded DNA structures during initiation. Positive and negative co-activators (e.g. Mediator)[9] transmit regulatory signals to the Pol II transcription machinery. Further factors enable the elongation and termination of transcription as well as RNA processing.[10]

In addition to these general transcription factors, higher eukaryotes express many sequence-specific transcription factors (TFs), which regulate gene transcription in respond to internal and external stimuli.[11-12] They are either ubiquitously or tissue-specifically expressed and necessary to recruit the Pol II transcription machinery to the transcriptional start site (TSS)[11] that is often referred as core or basal promoter. Many mammalian genes have multiple TSSs located in close proximity to each other generating additional diversity and complexity in the transcriptome.[13] Further, sequence-specific TFs can feature one or multiple DNA-binding domains and activate or repress transcription through binding to matching regulatory DNA sequences, which can be defined as *cis*-regulatory elements. These regions include the core promoter as well as proximal promoter regions directly up- and downstream of the TSS, within exons and introns, in 5' and 3' untranslated regions of genes, and even as far as 10kb away from the respective genes.[14-17] Depending on the regulatory function of TFs, *cis*-regulatory elements are often categorized into (activating) enhancers and (repressing) silencers.

While TFs direct the Pol II to their target genes, the accessibility of their regulatory elements is in turn dependant on the chromatin structure. The basic unit of chromatin is the

nucleosom,[18] a highly conserved nucleoprotein complex consisting of 146bp of double stranded DNA, which is wrapped in ~1.75 turns around an histone octamer composed of four core histone proteins (H2A, H2B, H3 and H4).[19] The individual nucleosomes are connected through short linker-DNA of different length that is stabilized by the fifth histone protein H1. Higher order chromatin structure achieved through further compaction of nucleosomes finally results in the condensation of around 2m of DNA packaged into the nucleus. Highly condensed 'heterochromatic' regions interfere with TF binding leading to predominantly transcriptionally inactive regions. On the other hand, less condensed 'euchromatic' regions are easily accessible for the transcriptional machinery and therefore associated with active gene transcription. The degree of condensation and changes in chromatin structure are dynamically controlled by epigenetic mechanisms including covalent modification of histone tails[20], chromatin remodeling complexes[21], and DNA methylation[22].

Finally, gene expression is controlled on a post-transcriptional level comprising RNA processing, and splicing, mRNA translation, mRNA transport, and mRNA and protein stability. The abundance of mRNA molecules within the cell is thereby of particular importance as it regulates the rate of protein synthesis. In the last two decades, it was found that microRNAs, short RNA molecules of around 22nt length, play a crucial role in the dynamic control of mRNA abundance in eukaryotic organism by specifically silencing their target genes.

Taken together, a multitude of highly complex mechanisms has been shown to dynamically regulate gene expression. Sequence-specific TFs control a correct spatial and temporal gene expression by altering the rate of transcription. Through their interactions with other TFs and transcriptional co-regulators, they form complex regulatory networks. Thus, the ability to bind to their targets is highly dependant on the chromatin structure, which is in turn regulated by epigenetic mechanisms adding further levels of regulation. Finally, the abundance of mRNA molecules within the cell is dynamically adjusted by post-transcriptional mechanisms such as microRNAs.

2.2 The mammalian heart

The mammalian adult heart is a four-chambered muscular organ that regulates the blood circulation, which is essential for nutrition and oxygen supply to the cells of the whole body. In human, deoxygenated blood from the body is transported through the superior and inferior caval vein into the right atrium, through the tricuspid valve to the right ventricle, and finally through the pulmonary valve into the pulmonary artery to the lungs (Figure 1). Oxygenated blood returns from the lungs through the pulmonary veins into the left atrium. It is passed through the mitral valve into the left ventricle before being pumped through the aortic valve to

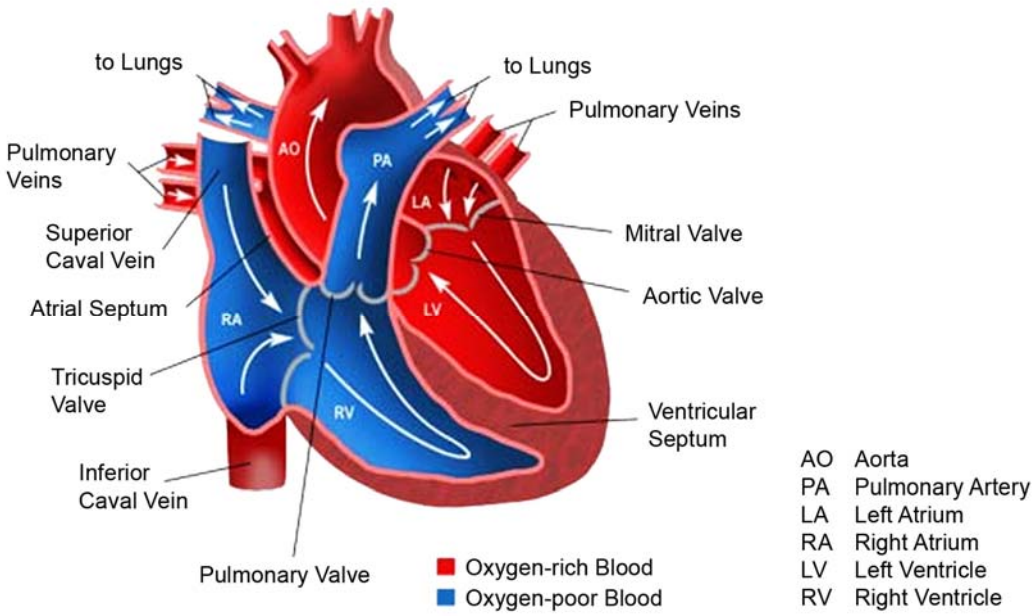


Figure 1: The structural features of the human adult heart.
Figure from http://www.medicalook.com/Heart_diseases/, modified.

the aorta into the arterial vascular circuit of the body. Specialized cells of the conduction system, which are organized into nodes, control the electrical impulse of the heartbeat.[23-24] The sinuatrial node, which is located at the junction between the right atrium and the superior caval vein, generates the initial electrical impulse, which is subsequently conducted throughout the atrial myocardium to the atrioventricular node. With a delay, the impulse is then rapidly transmitted along the His bundles and its bundle branches to the ventricular apex and finally throughout the ventricles by the Purkinje fibres.

In higher vertebrates, the heart is the first organ to form and function. All subsequent events in life depend on its repeated, rhythmic muscle contraction. Its fundamental functional units are cardiac myocytes. These are striated muscle cells which become specialized to form the ventricular and atrial myocytes as well as cells of the conduction system.[25] Further cell types including cardiac endothelial cells, cardiac fibroblasts and vascular smooth muscle cells contribute to the formation of the heart.[26-27] All these cell types arise from multipotent progenitors of the mesoderm where two pools of cardiac precursors exist defined as the first heart field (FHF) and the second heart field (SHF).[28] The precursors of the FHF contribute to the left ventricle while cells of the SHF contribute to the right ventricle, the outflow tract, sinus venosus and both left and right atria.[28]

During early mouse heart development at around embryonic day (E) 7.5, cardiac precursors from the FHF form a cardiac crescent-shaped structure in the anterior embryo with adjacent SHF cells medial and anterior to the FHF. The cells fuse along the ventral midline at around E8.0 to form a primitive linear heart tube, which ultimately starts to beat.[29] This tube is composed of an inner layer of endothelial cells shrouded by a myocardial cell

layer.[30] Cells from the SHF migrate to both ends of the tube where they also start to differentiate.[31] Around E9.0, the heart tube undergoes a complex rightward looping and myocardial cells expand and proliferate in a process called ventricular ballooning.[32] Distinct cardiac compartments become visible including the right and left ventricle as well as the outflow and inflow tract with atria and sinus venosus. The outflow tract will later become the aorta and pulmonary arteries, while the inflow tract will become the future atrioventricular canal. Endocardial cushions evolve in the atrioventricular canal depicting precursors of the tricuspid and mitral valve.[30] Endocardial cushions are also formed in the outflow region where they will give rise to the aortic and pulmonary valves.[30] Around E10 in mouse, the heart further matures and septation of the ventricles, atria and outflow tract results in a four-chambered structure.[29] The right atrium becomes directly connected to the right ventricle and the left ventricle to the outflow tract.[29] Both ventricles are septated by the interventricular septum, which arises from myocardial cells, and the atria are septated by the growth of two septa: the primary and the secondary septum.[28] Until birth, a small hole between both atrial septa remains open to allow the maternal circulation through the heart, but as soon as the lungs inflate pressure on the left side, both septa are pressed against each other and fuse. Around E16.5, the maturation of the heart is almost completed. Proliferation of myocardial cells is strongly reduced and further growth in volume of myocardial compartments occurs mainly by hypertrophic growth of already existing cardiomyocytes.[33] Finally, at birth, the majority of cardiomyocytes (97%) stops proliferation and exits the cell cycle.[33-34] Studies using rat hearts have shown that the postnatal growth of cardiomyocytes can be divided into three phases:[35-36] a hyperplastic phase (until day 4 after birth), a transitional phase with hyperplasia and hypertrophy concomitantly (between days 5 and 15) and a slow hypertrophic phase (from day 15 onwards).

The processes of heart formation involve the spatial and temporal orchestration of multiple molecular signaling pathways directing morphogenetic changes, which are precisely controlled by an evolutionary conserved network of transcription factors.[37] Mutations in these regulators of heart development during embryogenesis cause congenital heart disease (CHD). In human, CHD is the most common birth defect with an estimated incidence of around 1% of live births.[38] Most parts of the heart can be affected by defects ranging from minor to severe complex malformations such as cyanotic heart diseases, left-sided obstruction defects and septal defects.[28] Cyanotic heart disease are characterized by a blue skin colour resulting from the mixing of oxygenated and deoxygenated blood. Tetralogy of Fallot (TOF), the most common (6%) cyanotic heart defect, is the major cause for the so-called “blue baby syndrome”.[39] TOF has four clinical features: a ventricular septal defect (VSD), a biventricular connection of the aorta, right ventricular outflow tract obstructions, and right ventricle hyperthrophy.[40] It has long been considered that TOF occurs due to

environmental alterations,[41-42] but meanwhile several mutations in cardiac TFs (like e.g. Gata4 or Nkx2.5)[43-44] have been identified that can lead to TOF as well as other CHD. Thus, the majority of CHD could not be linked to mutations in TFs. In addition, CHD patients with the same TF mutations show variable expressivity and penetrance[45-46] indicating that a complex mixture of multiple genetic factors, environmental influences and epigenetic factors contribute to their incidence.[47]

2.3 Cardiac transcriptional regulation

The complex processes of heart development as described in the previous chapter are under the control of an evolutionarily conserved network of TFs that directs cardiac cell fates, myocyte differentiation and cardiac morphogenesis, thereby linking upstream signaling pathways with direct and indirect downstream target genes.[37,48] Gaining more insight into the function of these TFs will improve the understanding of these processes and will lead to a better understanding of the causes of CHD. Moreover, many studies have begun to unveil the important role of miRNAs and epigenetic factors such as histone modifications and chromatin remodeling factors for cardiogenesis. For that reason, it is necessary to connect the function of TFs with these molecular mechanisms. Likewise, the study of the interplay between TFs, epigenetic regulators and miRNAs will provide new insights into the transcriptional regulation of cardiac networks. A number of important TFs as well as studied epigenetic factors and miRNAs will be introduced in the following section.

2.3.1 DNA-binding transcription factors

During early cardiogenesis, several inductive signals, which derive from the underlying endoderm as well as adjacent ectoderm and extra-embryonic tissue, are committed to the cardiac mesoderm to regulate the contribution of progenitor cells to FHF and SHF cell lineages.[49] Among these inductive signals are activating bone morphogenetic proteins (BMPs) and basic fibroblast growth factors (FGFs) as well as inhibitory Wnt signaling proteins.[50] For example, the combination of Fgf8 and BMP signaling has been found to induce early cardiac differentiation.[51] In consequence of the inductive signals a core set of evolutionarily conserved TFs are activated including Gata, Mef2, Nkx2, Srf, Tbx and Hand.[37] These are considered as key regulators for the activation of muscle-specific genes and genes that control growth and patterning of the heart, often assisted by further cofactors.[37] Several studies have reported that they cross- and auto-regulate their own expression and the loss of function of any of these core TFs can ultimately lead to failure in cardiovascular development and cause CHD.

The zinc-finger TF Gata4, one of six members of the Gata family that bind the DNA sequence A/G GATA A/T, is expressed within the precardiac mesoderm from early stages (E7.0) and continues to be expressed within the adult heart.[50] Gata4 is known to regulate numerous cardiac genes encoding contractile elements like α- and β-myosin heavy-chain (α- and β-MHC).[52] Furthermore, it can physically interact with Nkx2.5, Mef2, Srf, Gata6, Tbx5 and Hand2.[52-57] Homozygous Gata4 null mice show severe morphogenic defects in heart tube formation and die between embryonic day E7.0 and E9.5.[58] The importance of Gata4 in heart development is further evidenced by the fact that in human GATA4 haploinsufficiency causes CHD including TOF.[43]

The myogenic MADS-box proteins of the myocyte enhancer factor-2 family (MEF2a, -b, -c, and –d) bind A/T rich regulatory promoter regions of several muscle relevant genes, thus being crucial for the regulation of skeletal and cardiac muscle cell differentiation. They are known to cooperate with other core cardiac TFs to regulate the expression in particular of contractile proteins via the interaction with MyoD family members. Their own expression, on the other hand, has been shown to be regulated by NKX2 proteins amongst others.[37] The four Mef2 family members can in part compensate each other's function.[59] Nevertheless, homozygous mutant mice lacking Mef2c die during early embryogenesis due to severe failure in heart tube looping and perturbations of cardiac muscle gene expression.[60] Mice lacking Mef2a, the Mef2 member predominantly expressed in post-natal cardiac muscle cells, die within the first week after birth in consequence of myofibrillar fragmentation, mitochondrial disorganization and impaired myocyte differentiation.[61] Moreover, a deletion of seven amino acids in human MEF2a causes coronary artery disease and myocardial infarction.[62]

Another MADS box TF important for heart and muscle development is the widely expressed serum response factor Srf. Srf binds to a DNA consensus sequence know as the CArG box [$CC(A/T)_6GG$], which is present in single or multiple copies of Srf target gene promoters.[63-65] Srf is suggested to be a master regulator of the actin cytoskeleton[66-67] as many of its target genes encodes for proteins involved in cytoskeleton, cell adhesion and contractility such as actin, myosin light-chain, and myosin heavy-chain.[67-69] Moreover, Srf is known to be involved in the control of muscle cell differentiation and cellular growth. In line with this, CArG boxes have been reported to be present primarily in promoters of muscle- and growth-factor-associated genes.[64-65,70-72] Since Srf is ubiquitously expressed and has relatively low intrinsic transcriptional activity, its activity to control heart and muscle gene expression is largely dependent on its interaction with positive and negative tissue-specific cofactors. For example, Srf together with Gata4 and Nkx2.5 can form complexes with each other to enhance transcription of cardiac differentiation genes.[73] A further potent co-activator of Srf is Myocardin, which is exclusively expressed in smooth muscle cells (SMCs) and

cardiomyocytes.[68] The crucial role of Srf for heart and muscle development has been demonstrated by several studies. Embryonic stem cells (ESCs) lacking Srf display defects in formation of cytoskeletal structures.[74] Neonatal cardiomyocytes lacking Srf show severe defects in the contractile apparatus.[75] Srf-deficient mice have strong gastrulation defects and die between E8.5 and E12.5.[76] Cre mice lacking Srf in 80% of cardiomyocytes display drastic heart defects and die at E11.5 while Cre mice lacking skeletal muscle Srf expression die during the perinatal period due to severe skeletal muscle hypoplasia.[77-78]

The homeobox gene Nkx2.5 is one of the earliest markers in vertebrate heart development that responds to inductive signals and initiates cardiogenesis. Nkx2.5 binds DNA through its homeodomain. Important interaction partners are Tbx5 and Gata4.[79] Mice lacking Nkx2.5 die due to abnormal morphogenesis of the heart tube and failure in left ventricular development with lethality at E9.5.[80] In human, more than 30 mutations have been identified within the *NKX2.5* gene and the majority of these account for the incidence of CHD like atrial septal defects (ASD), ventricular septal defects (VSD), left ventricle noncompaction, TOF, or disturbances in the atrioventricular conduction system.[44]

A further family of important regulators in heart development are the Tbx proteins that all feature a conserved 180 amino acid long region termed the T-box.[50] They are essential for early cardiac lineage determination, chamber specification and the development of the conduction system.[81] In line with this, mutations in human *TBX* genes cause CHD. One of the first evidences for the role of Tbx proteins in heart development is based on findings that mutations in *TBX5*, one of at least 20 family members, cause Holt-Oram syndrome, an autosomal dominant disorder characterized by a variety of structural and functional cardiac defects.[82]

The basic helix-loop-helix (bHLH) transcription factors Hand1 and Hand2 are early expressed in derivates of the FHF and SHF where they regulate in particular the growth of the ventricles.[83] Deletion of *Hand2* in mouse resulted in embryonic lethality at E10.5 due to abnormalities in formation of the right ventricle[84] and ESCs lacking *Hand1* are unable to contribute to the outer curvature of the heart that gives rise to the left ventricle.[85]

Additionally, there are many studies that demonstrate how the regulatory potential of several TFs is strongly dosage dependent. For example, heterozygous Tbx5 mutant mice have cardiac abnormalities comparable to patients with HOS, whereas Tbx5 homozygous mice have severe hypoplasia of posterior domains in the developing heart due to decreased expression of multiple genes.[86] Tbx20 deficiency in mice results in defects in heart formation only if the dosage falls below 50%.[87] Similar observations were obtained for Gata4. Knockout mice expressing 50% of Gata4 protein in the heart survive normally. In contrast, embryos expressing less than 30% die between E13.5 and E16.5 due to severe ventricular defects.[88]

Taken together, all these TFs have been shown to work cooperatively and their

combinatorial interactions regulate the expression of cardiac specific genes throughout the development to the point of adult life. However, the interaction of these TFs to coordinate the temporal and cell-specific regulation remains poorly understood.

2.3.2 Histone modifications and chromatin remodeling factors

The N-terminal tails of histones, which stick out of the globular histone octamer, are subject to a wide range of post-translational modifications (Figure 2). A large number of these covalent modifications are known and can be chemically grouped into acetylation, methylation, phosphorylation, ubiquitylation, sumoylation and ADP ribosylation.[90] These histone modifications undergo dynamic changes during embryonic development and in response to extracellular cues. They alter the chromatin structure either directly or indirectly and are present sequentially or in combination, leading to an activation or repression of transcription in the surrounding genomic regions.[91-92] The occurrence of distinct modification patterns has evoked the formulation of the 'histone code' hypothesis by Strahl and Allis in 2000.[92] It comprises the read-out of modification combinations, which is translated into an active euchromatic or a silent heterochromatic state. The chemical properties of histone modifications can generate direct changes in the chromatin structure via electrostatic repulsion or steric hindrance, respectively. On the other hand, histone modifications can be read by effector proteins, which bind to or are blocked by specific modifications and thereby mediate changes in chromatin structure and consequently contribute to gene regulation.

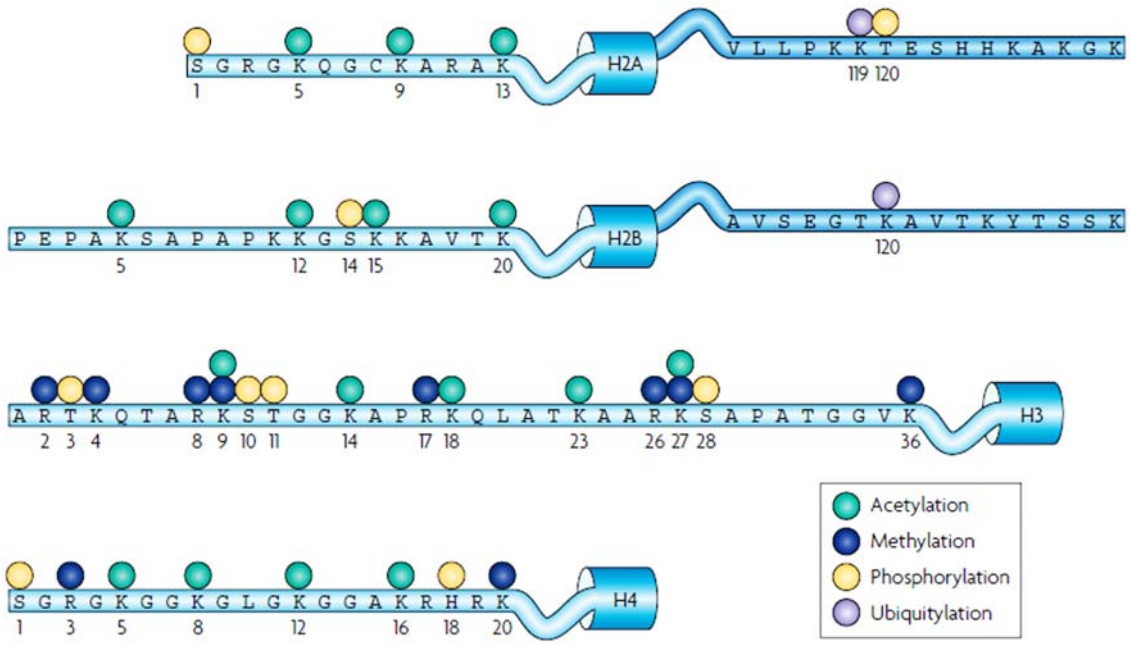


Figure 2: A selection of post-translational histone modifications. Amino-acid residues of histone H2A, H2B, H3 and H4, are subject to various post-translational modifications, especially those residues located at the N-terminus. These include methylation, acetylation, phosphorylation, and ubiquitylation. Figure from Spivakov and Fisher[89], modified.

So far, the best-characterized histone modifications are acetylation and methylation. Acetylation of lysine neutralizes the positive charge of these basic residues. As consequence, the interaction between histones and the negatively charged DNA is weakened directly creating a more open chromatin structure. A similar mechanism has been described for histone phosphorylation.[93] Consequently, hyperacetylation is in general associated with active chromatin, while regions with less acetylated histone tails point to repressive transcription. In 2007, the ENCODE (ENCyclopedia of DNA Elements) project analyzed the occurrence of several histone modifications including acetylation of histone H3 and H4 (H3ac and H4ac) within 1% of the human genome.[94] They observed that H3ac modifications are tightly associated with transcriptional start sites of actively expressed genes while H4ac have distributions that are more widespread. Two competitive enzyme families catalyze the dynamic balance between acetylated and deacetylated histones: histone acetyltransferases (HATs) and histone deacetylases (HDACs),[95] which in general do not bind directly to DNA, but are recruited by TFs and their cofactors.[96] Many transcriptional coactivators, such as Gcn5/PCAF, CBP/p300 and SRC-1, have been shown to possess intrinsic HAT activity, whereas transcriptional corepressor complexes, such as mSin3a, NCor/SmRT and NURD/Mi-2, contain subunits with HDAC activity.[95,97-98]

HATs and HDACs are key factors in development and disease. Most of them are early expressed during embryogenesis and several of these show relatively restricted expression patterns to specific cell types such as cardiomyocytes. For example, HDAC5 and HDAC9 are highly expressed in muscle, heart and brain,[99-100] and mice lacking both die due to abnormalities in growth and maturation of cardiomyocytes resulting in ventricular septum defects and thin-walled myocardium.[99] Homozygous mutant mice lacking HAT p300 die between E9.0 and E11.5 exhibiting defects in cell proliferation and heart development due to reduced expression of muscle structural proteins such as β-MHC and α-actinin.[101] In addition, class II HDACs (HDAC4, 5, 7 and 9), which mainly negatively regulate cardiac growth, have been postulated to repress cardiac hypertrophy, thus being a potent candidate for therapeutic treatment.[102-103] Conversely, the pharmacological suppression of HDACs was found to be a viable therapeutic strategy for cardiac hypertrophy in cell culture models.[104-105] Hence, the control of histone acetylation using HDAC inhibitors also provides promising clinical benefits as therapeutics for cardiac failure.[106]

Histone methylation is another well-studied histone modification, which mainly occur on lysine and arginine residues of histone tails.[90,107] Lysine residues can be mono-, di- or tri-methylated, whereas arginines are only mono- or di-methylated. The transfer of methyl-groups on the ε-nitrogen of lysine residues is mediated by either SET domain- or non-SET domain-containing lysine histone methyltransferases and their removal is catalyzed by histone demethylases.[107-108] As in the case of acetylation, the methylation of histones is

highly dynamic and provides docking sites for regulatory proteins. Depending on the individual position and the degree of methylation, gene transcription can be actively or repressively affected. For example, methylation of H3 on lysine 4 (K4), K17, and K36 is associated with active gene expression, whereas methylation of H3K9 and H3K29 is linked to transcriptionally inactive heterochromatin.[109-111] Furthermore, genome-wide studies using a combination of histone-specific chromatin immunoprecipitation (ChIP) with either DNA microarray analysis (ChIP-chip) or sequencing analysis (ChIP-seq) have shown that tri-methylation of H3K4 is enriched in promoter regions at 5' ends of transcriptionally active genes, whereas di-methylation of lysines is enriched in the middle and mono-methylation at the 3′ end of genes.[112-113]

The activity of several heart-specific TFs including Srf, Mef2, and Gata4 have been shown to be connected to the presence of histone modifications and to the activity of histone modifying enzymes. A well-described model for the activation of smooth muscle cell (SMC) genes was proposed by McDonald and Owens.[114] Here, the binding of Srf/Myocardin at CArG boxes of SMC gene promoters is achieved through a specific pattern of H4ac and H3K4me2. Subsequently, the Srf/Myocardin complex recruits other regulators like p300, leading to further histone acetylation including H3ac and an activation of SMC transcription. It is also known, that p300 acetylates lysine residues of Gata4 resulting in enhanced DNA binding affinity and an activation of Gata4-dependant transcription.[54] Furthermore, Srf activity is negatively regulated by HDAC4 in a Ca^{2+}-sensitive manner.[115] More evidence that Srf activity is intimately linked to HDAC activity is provided by the finding that Homeodomain only protein (Hop), which does not bind DNA, inhibits Srf-dependant transcription by recruiting HDAC2 to Srf target genes.[102] Mef2 is a further example for the dependence of heart-specific TFs on histone modifying enzymes. Mef2 proteins can act as transcriptional repressors through the interaction with class II HDACs. Its dissociation allows p300 to bind at the same interaction domain, thereby converting Mef2 from a transcriptional repressor to a transcriptional activator.[106,116-118]

Beside the modification of histone tails, ATP-dependant chromatin remodeling is an additional process regulating changes of chromatin structures, thereby controlling TF binding to regulatory elements. The ATPase subunit of ATP-dependant chromatin remodeling complexes utilizes the free energy of ATP to disrupt DNA contacts to nucleosomes, move nucleosomes along DNA, and remove or exchange nucleosomes[21]. About 30 genes encode the genetically non-redundant ATPase subunits of these complexes in mammals.[21] All chromatin remodeling ATPases found so far belong to the SNF2 family[119] and, in respect to their sequence and structure, the corresponding chromatin remodeling complexes can be broadly divided into four main subfamilies: SWI/SNF, ISWI, CHD and INO80,[21,120] which are characterized by a unique subunit composition.[121] In vertebrates, the Brahma-associated

factor (BAF) complex, which belongs to the SWI/SNF family, can function as both transcriptional activator and repressor. Moreover, the BAF complex can switch between these two modes of action in context of the same gene.[122] A schematic illustration of the subunit composition in the mouse BAF complex is depicted in Figure 3. It contains one of two alternative ATPases Brg1 or Brm. 11 further subunits vary depending on cell type and developmental stage.[120] The dynamic composition of these subunits enables the regulation of diverse biological processes. For example, mouse embryonic stem cells (mESCs) express a complex called esBAF, which is characterized by containing Brg1 but not Brm and Baf155 but not Baf170.[123-124] This complex regulates self-renewal and pluripotency of mESCs. During differentiation to neuronal progenitors, several subunits of the esBAF complex are exchanged. For instance, Baf60b is replaced by the incorporation of Baf60c.[123,125] Furthermore, this neuronal-progenitor-specific BAF complex (npBAF) replaces its subunits Baf45a and Baf53a during transition to the neuron-specific BAF complex (nBAF) with Baf45b and Baf53b.[125]

Brg1, one of the essential and ubiquitously expressed ATPase subunits of the BAF complex plays a critical role in regulating cardiac growth, differentiation and gene expression.[126] Heart restricted deletion of *Brg1* in mice and zebrafish reduces the proliferation of cardiac progenitors leading to variable defects in heart formation including disrupted chamber morphogenesis.[127] In addition to proliferation, Brg1 controls fetal cardiac differentiation by interacting with HDACs to repress α-myosin heavy chain (α-MHC) and activate β-myosin heavy chain (β-MHC).[126] In turn, β-MHC is the predominant isoform in

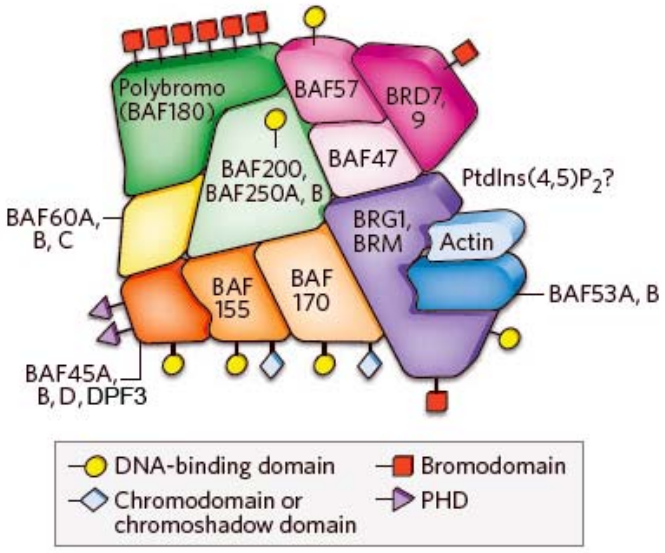


Figure 3: Schematic illustration of subunit compositions in the mouse BAF complex. The combinatorial assembly of 12 different subunits encoded by gene families (like BAF60A, B, and C, or DPF3) enables tissue-specific and developmental-stage-specific regulation of gene expression. For example, the core ATPase subunit of the mammalian BAF complex is either Brg1 or Brm. The different subunits can recruit the complex to their targets via histone modification "reader" domains. Figure from Ho and Crabtree[21], modified.

embryonic hearts while α-MHC is more expressed in the adult heart. As mentioned above, the combinatorial assembly of different subunits in the BAF complex enables tissue-type-specific and cell-type-specific regulation of gene expression. Two of these well-known tissue-specific subunits of the BAF complex are Baf60c and Dpf3. Baf60c (also known as Smarcd3) is specifically expressed in the precardiac mesoderm, somites and midbrain.[128] It is required for heart morphogenesis and differentiation of mesodermal cells into cardiomyocytes.[128] Studies have shown that Baf60c can physically interact with Gata4, Nkx2.5, and Tbx5 to recruit the BAF complex to cardiac specific target genes. Similar to Baf60c, the subunit Dpf3 (also known as Baf45c) was found to be incorporated into the BAF complex during transition from neuronal progenitor cells to post-mitotic neurons.[21,125] The detailed characterization of Dpf3 and its role in heart and muscle development are part of this thesis and are further addressed in chapter 6.

According to the current model of the function of ATP-dependent chromatin remodeling complexes, the core ATPase is essential for the nucleosomes' mobility.[129] The other subunits interact with a variety of specific TFs and recruit the complexes to their target sites.[130] Several recent studies have demonstrated that different subunits of chromatin remodeling complexes contain diverse conserved protein domains that can function as "readers" of histone modifications. For example, histone lysine acetylation marks are recognized by the bromodomain, an about 110 amino-acid long protein module, which is found in many chromatin-associated proteins.[131] Until the double PHD-finger containing protein DPF3 was identified,[132] the bromodomain was the only protein domain known to recognize histone acetylation. The bromodomain of BRG1 has been implicated in the recognition of acetylated lysines within H3 and H4, thereby recruiting SWI/SNF complexes to target promoters.[133] Nearly all known HAT-associated transcriptional co-activators contain bromodomains such as Gcn5p or its homolog PCAF (p300/CBP-associated factor).[131] Human TAFII250, a large subunit of the TFIID multiprotein complex that initiates assembly of the basal transcription machinery, contains a double bromodomain which targets multiple acetylated H4 peptides.[134] Histone lysine methylation on the other hand, is recognized by two general classes of structures, namely PHD (plant homeodomain) fingers and members of the Royal superfamily comprising chromo-, tudor-, and MBT (malignant brain tumor) domains.[112,135-136] Higher lysine methylation states like Kme2 or Kme3 are targeted by e.g. the chromodomain of HP1 (heterochromatin protein-1), the double chromodomain of CHD1 (chromo helicase DNA-binding protein) or the double tudor domain of JMJD2A (jumanji domain-containing protein 2A).[137-139] In addition, lower methylation states like Kme1 and Kme2 are recognized by the tandem tudor domain of 53BP1 (p53-binding domain) or the MBT repeats of L3MBTL1 (human lethal-(3) MBT repeat-like protein-1).[140-141] As for the Royal superfamily, PHD finger domains have also been shown to bind lysine methylation of histones. For example, the PHD

finger of BPTF, the largest subunit of the nucleosomal remodeling complex, was found to interact specifically with H3K4me3 to control HOX gene expression,[134] while the PHD finger-containing inhibitor of growth-2 (ING2) protein of the Sin3-HDAC complex interacts specifically with H3K4me3 to direct HAT and HDAC complexes to chromatin target sites.[142-143] Taken together, there are a number of conserved protein domains that specifically bind diverse histone modifications in a way that is dependent on both modification state and position within a histone sequence to control transcriptional gene expression.

2.3.3 MicroRNAs

In the last two decades of research, the important role of small endogenous non-coding microRNAs (miRNAs) has been uncovered. They were shown to control and fine-tune gene expression in almost every cellular process and provide an additional mechanism by which gene expression is regulated.[144-145] Mature miRNAs are single-stranded RNAs of ~22nt in length, which interact sequence-specifically with the 3' untranslated region (UTR) of target mRNAs to repress their expression at the post-transcriptional level.

Since the discovery of the first miRNA genes in 1993 and 2000, known as *lin-4* and *let-7* in *Caenorhabditis elegans* (C.elegans), the number of newly identified miRNA genes increases enormously each year.[146-147] In human, around 60% of all protein-coding genes are predicted to be regulated by miRNAs and genome-wide computational analyses revealed an estimated number of around 1000 miRNAs encoded in the human genome regulating on average hundreds of different targets each.[148-149] MiRNAs are encoded by their own sets of evolutionarily conserved genes and are either ubiquitously expressed or display strict cell-type or tissue-specific expression patterns, thereby controlling developmental timing, tissue differentiation and maintenance of tissue identity during embryogenesis and adult life.[150] Studies of miRNA distribution throughout the genome have shown that miRNA genes are often derived from polycistronic clusters and that they are transcribed as independent units from intergenic, intronic, or exonic regions.[151-152]

In most animals, miRNAs are transcribed by Pol II as long pri-miRNA precursors that form hairpin-shaped structures and receive a 5'cap and a poly-(A) tail at the ends similar to that of mRNAs (Figure 4).[153] The double-strand specific endonuclease RNase III (ds-RNaseIII) Drosha and its cofactor, the double-strand RNA binding domain protein DGCR8 (DiGeorge syndrome critical region 8), cleave the pri-miRNA into approximately 70nt long hairpins known as pre-miRNAs.[154] Subsequently, exportin-5 mediates the transport of pre-miRNAs into the cytoplasm where they are further processed by ds-RNaseIII Dicer into ~22nt long miRNA-miRNA* duplexes.[155-156] Both strands of the duplex can potentially act as a functional miRNA, but only one is finally incorporated into the RNA-induced silencing

complex (RISC). After assembly of the RISC complex, which contains TAR RNA binding protein (TRBP) and further dsRNA-binding proteins of the Ago family, the mature miRNA sequence binds with its 5' end, the so-called 'seed' region, to the complementary mRNA target sequence present in the 3'UTR. Consequently, targeted mRNAs can be accumulated, blocked for translation, marked for degradation or deadenylation, or stored in so-called P-bodies, which contain enzymes for RNA turnover.[157]

Many studies have started to unveil the central role for miRNAs as governors of gene expression during cardiac development, function and disease. MiR-1 and miR-133 are the first identified and most studied miRNAs that are implicated in cardiac and skeletal muscle development. In vertebrates, both miR-1 and miR-133 are generated from two distinct bicistronic transcripts that yield three related miRNA clusters: miR-1-1/133a-2, miR-1-2/133a-1, and miR-206/133b.[151,159] The former two clusters are specifically expressed in skeletal and cardiac muscle cells, whereas the latter one is skeletal muscle specific. Their expression is controlled by Srf, Mef2 and MyoD in feedforward- and negative feedback loops, respectively.[159] For instance, miR-1 represses the repressor HDAC4, leading to an activation of Mef2, which in turn activates expression of miR-1 and other target genes. In contrast, miR-133 expression is regulated by Srf and in turn represses the expression of Srf. The cardiac function of miR-1 and miR-133 was investigated by several studies using transgenic mouse

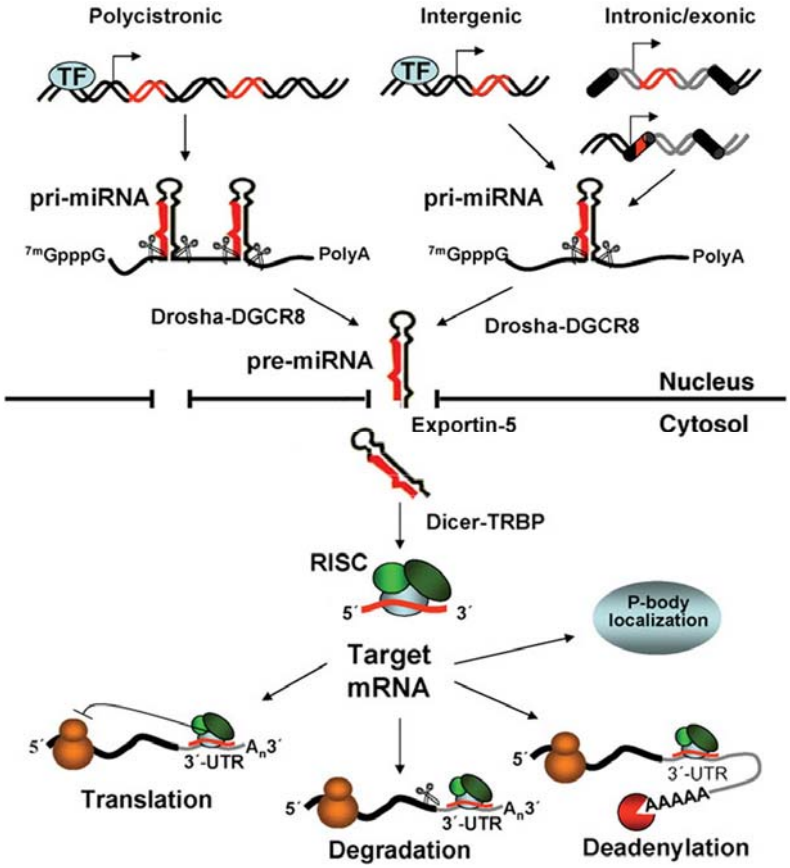


Figure 4: MiRNA genomic organization, biogenesis and function. Figure from Fazi and Nervi[158], modified.

models. Thereby, miR-1 was e.g. found to be responsible for the balance between differentiation and proliferation during cardiogenesis.[160] In line with this, cardiac-specific overexpression of miR-1 results in an inhibition of proliferation leading to embryonic lethality at E13.5 due to thin-walled ventricles and heart failure.[160] MiR-1-2 knockout mice show late embryonic lethality with incomplete penetrance due to thick-walled chambers and septal-defects.[161] In contrast, miR-133 was found to function in the opposite way. MiR-133 inhibits differentiation and promotes proliferation of cardiac muscle cells.[162] The double knockout of miR-133a-1 and miR-133a-2 results in late embryonic and neonatal lethality due to ventricular septum defects and chamber dilatation, whereas no obvious cardiac abnormalities were observed in mice lacking only one of these miRNAs.[163] This reveals the essential and redundant roles for miR-133a-1 and miR-133a-2 in the heart. In addition, an increasing number of studies shows dysregulated expression profiles of miRNAs in human heart failure and animal models of heart disease, pointing to their crucial role as fine-tuners for cardiac gene expression.[164-167] As an example, Thum and colleagues[167] investigated miRNA profiles in left ventricular tissue from patients with heart failure in comparison with tissue from healthy adult and fetal human hearts. They observed that the majority of up-regulated (e.g. miR-1, 21, 125, 208) and down-regulated (e.g. miR-16, 107, 126) miRNAs in failing hearts versus non-failing hearts were regulated concordantly in fetal heart tissue suggesting that alterations in miRNAs expression contribute largely to an activation of the 'fetal gene' program which is a well-known signatures of failing myocardium.[164,167]

MiRNAs have been thought to act primarily at the post-transcriptional level. However, more and more miRNAs were identified which influence gene expression at the transcriptional level. For example, miR-10a directly targets homologous DNA region in the promoter region of Hoxd4 gene and repress its expression whereby H3K27me3 is suggested to be involved.[168] A link between miRNAs and chromatin remodeling was further demonstrated by a study showing that miRNAs can induce heterochromatic features in promoters with sequence complementarity.[169]

Taken together, a full understanding of how miRNAs regulate the complex processes in cardiogenesis and cardiac function is still limited, as only a small number of miRNAs expressed in the heart have been functionally analyzed so far.

2.4 Experimental technologies

In order to study cardiac development and function an adequate system is required that resembles as best as possible the geno- and phenotypical characteristics of cardiomyocytes *in vivo*. Model organisms such as the zebrafish *Danio rerio*, the frog *Xenopus leavis* or the fruit fly *Drosphila melanogaster* were frequently used to study heart formation of lower

vertebrates and invertebrates. Moreover, the mouse is extensively used as it can be genetically manipulated while serving as a good model for the mammalian heart. The heart is a mixture of different cell types including cardiomyocytes, endothelial cells, cardiac fibroblasts and vascular smooth muscle cells[26-27] making it technically challenging to distinguish them and moreover to obtain enough homogenous cell populations. Isolated embryonic and neonatal rat primary cardiomyocytes are often used as model system for *in vitro* studies.[170] However, their use is limited as they lack many adult cardiomyocyte characteristics. In addition, they stop proliferation as well as non-cardiomyocytes become overgrown after a few days in culture making it problematic to obtain enough cells for the respective experiments. *In vitro* models using murine pluripotent embryonic stem or embryonic carcinoma cells have been also successfully employed to study structural and functional properties in cardiomyogenesis.[171-174] In particular, murine pluripotent P19 embryonic carcinoma cells, which are derived from a teratocarcinoma, can differentiate into a heterogeneous cell population under certain growth conditions resulting in a high percentage of cardiomyocytes.[171,175-176] Nevertheless, to obtain a cell line with pure or highly enriched population of cardiomyocytes from differentiating embryonic stem cells is still challenging. Thus, for a number of applications it is more helpful to use a cell culture model that can be maintained at a constant physiological state.

HL-1 cells were the first cardiac muscle cell line that could be continuously passaged in culture while maintaining a differentiated cardiac-specific phenotype. In 1998, Claycomb and colleagues developed this cell line from the AT-1 murine atrial cardiomyocyte tumor lineage.[177] The cells possess the ability to contract and retain differentiated cardiac morphological, biochemical, and electrophysiological properties. The gene expression pattern is similar to that of adult atrial myocytes as cardiac and muscle specific genes are expressed including e.g. α-cardiac myosin heavy chain, α-cardiac actin, connexin 43, desmin, sarcomeric myosin, and atrial natriuretic factor. Furthermore, the characterization using microscopic, genetic, immunohistochemical, electrophysiological, and pharmacological techniques has demonstrated how similar HL-1 cells are to primary cardiomyocytes.[170,177] Thus, they are a suitable model system to study pathological conditions such as hypoxia, hyperglycemia-hyperinsulinemia, and ischemia-reperfusion.[170,178] Moreover, to elucidate normal cardiac function, they can be used to investigate signalling pathways, apoptosis and cell cycle mechanisms, as well as electrical, metabolic and transcriptional regulation. H9C2 cells are another clonal cell line derived from rat embryonic ventricular tissue. These cells possess properties of adult cardiac and skeletal muscle cells and are frequently used to study mechanisms of heart development and function.[179-181]

In order to study the function of DNA binding proteins such as histone marks or transcription factors, it is necessary to identify the region where binding occurs. One

traditional technique that enables the analysis of protein-DNA or protein-RNA interactions is the electrophoretic mobility shift assay (EMSA). Here specific short oligonucleotides were synthesized and incubated with the protein of interest resulting in case of a positive binding event to a mobility shift in gel electrophoreses. Both the protein and the DNA sequence have be known beforehand making it a method that has little predictive power on a genome-wide scale. Therefore, the invention of chromatin immunoprecipitation (ChIP) technique combined with microarray analysis or sequencing has become the method of choice. Moreover, it is suitable to couple ChIP experiments with gene expression studies to obtain information about the functional consequence of DNA binding that can be activating, repressive or non-functional. The expression status of genes that are marked by a certain histone modification or harbour binding sites for a particular transcription factor can be analyzed with qPCR or the 'out-of-date' Northern Blotting. Both techniques are based on the detection of gene expression via sequence-specific primers or probes, respectively, and in general only one gene per reaction can be measured which have to be known beforehand. In the past, the application of microarrays was more widespread enabling the screening of expression levels of thousands of genes in one single sample. Nowadays, next-generation sequencing has replaced microarrays as it is possible to analyse the expression of all genes in parallel without any pre-design of the analysed sequence. In addition, overexpression or RNAi knockdown experiments coupled with subsequent gene expression analysis are helpful techniques to obtain more information about the individual function of DNA binding regulatory factors.

2.4.1 Chromatin immunoprecipitation

Chromatin immunoprecipitation combined with microarray detection (ChIP-chip) or next generation sequencing (ChIP-seq) has become a powerful technique for genome-wide identification of protein-DNA interactions such as chromatin histone marks or transcription factor binding sites. The major steps involved are schematically illustrated in Figure 5.

In a typical ChIP experiment the starting material of cultured cells or tissue samples are fixed with formaldehyde to cross-link the non-covalent protein-DNA interactions as well as protein-protein interactions allowing the identification of direct and indirectly bound DNA sequences. In the next step, the cross-linked DNA is sheared by sonication into short fragments of usually 100 to 500bp in size. A small aliquot of the fragmented DNA is directly purified as reference 'Input' sample for later normalization. The specific DNA fragments are then enriched and isolated by immunoprecipitation with an antibody against the protein of interest. If such a 'ChIP-grade' antibody is not available, the protein can also be overexpressed with a tag (e.g. Flag[182] or biotin[183]) that allows the subsequent isolation via an antibody against the

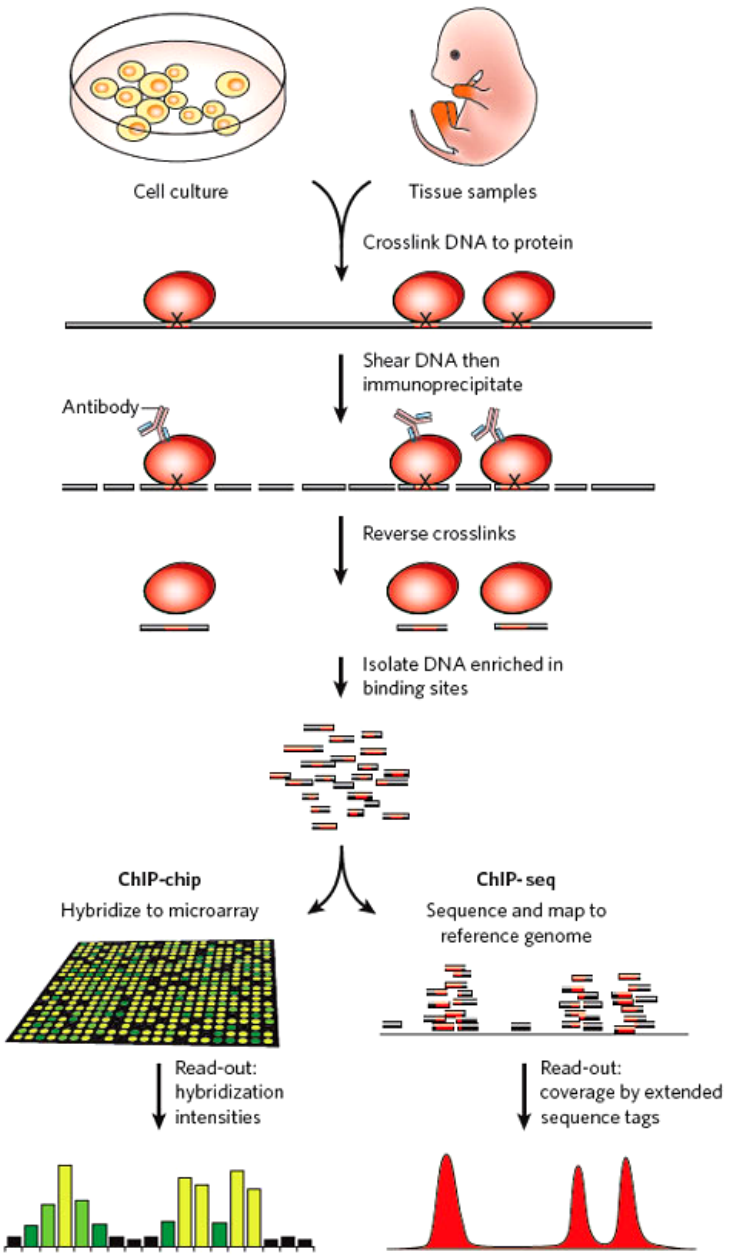


Figure 5: **Schematic overview of the individual steps involved in ChIP-chip and ChIP-seq.** Figure from Visel *et al.*[184]

tag. However, the overexpression of a protein with an additional tag can lead to undesired alterations of the physiological conditions resulting in an abnormal binding behavior. In the last step, the enriched DNA fragments are reverse cross-linked from the proteins and purified via DNA precipitation.The enrichment relative to the non-immunoprecipitated Input DNA can then be quantified for individual binding positions using qPCR. To get a genome-wide readout of the specific binding sites, either ChIP-chip or ChIP-seq are the methods of choice. For ChIP-chip, an additional amplification step is sometimes necessary, as the amount of DNA is not sufficient. Afterwards, the enriched and amplified DNA pool is hybridized to a microarray. Therefore the ChIP as well as the reference Input sample are labelled with two

different fluorescence dyes such as Cy5 or Alexa647 (for the ChIP sample) and Cy3 or Alexa555 (for the Input). Subsequently, both samples are combined and hybridized to a single or multiple microarrays. The fluorescence intensities of the two dyes were measured and set in relation to each other to determine the protein-DNA interaction sites in the ChIP experiment. In general, the arrays represent the genome of interest in form of overlapping fragments (tiling arrays). Full-genome tiling arrays for higher organisms are expensive due to the large genome size. Alternatively, arrays can be custom designed for regions of interest such as promoters of selected genes. It is known that ChIP-chip data from higher organisms can be noisy and tend to have low resolution[185-186] occasionally dependent on the fragment size and the spacing of the probes on the array. As consequence, ChIP-chip has become more and more replaced by ChIP-seq as an alternative method. Here, next-generation sequencing is used as final step to determine the sequence of the immunoprecipitated DNA fragments. The sequence reads are then computationally mapped to the reference genome and further comprehensive statistical approaches are implemented to identify the genomic locations. According to recent studies, ChIP-seq is more sensitive compared to ChIP-chip without the need to restrict the analysis to selected regions.[184,186-188]

2.4.2 qPCR

Quantitative real-time PCR (qPCR) is a widely used, sensitive and accurate method for absolute and relative quantification of gene expression. In addition, it is routinely applied for the quantification of genomic DNA as in ChIP experiments making it a very suitable technique in research. The detection chemistry of all real-time PCR procedures are based on monitoring amplification products either through binding of double-stranded DNA (e.g. SYBR Green I)[189] or through target-specific hybridization to single-stranded DNA (e.g. hydrolysis probe assays).[190-191] SYBR Green I is a non-specific fluorescent dye that intercalates with any double-stranded DNA that is amplified via specific primer pairs. Hydrolyses probe assays utilize in addition to sequence-specific primers short oligonucleotides (e.g TaqMan probes) that are labelled with a fluorescent reporter that can be detected after hybridization of the probe with its complementary DNA target after every PCR cycle. Thus, the use of reporter probes significantly increases the specificity and enables quantification even in the presence of non-specific DNA amplification products or primer dimers. In both methods, the intensity of the emitted fluorescence is directly proportional to the amount of amplified DNA in the PCR reaction allowing the quantification of the initial DNA concentration of the sample. The PCR reaction is influence by a number of factors including the efficiency of the PCR reaction, which has to be validated beforehand to guarantee that the quantity of DNA (theoretically) doubles every cycle. Moreover, the determination of the relative amount of DNA is more

conventional. Therefore, internal controls like “housekeeping genes” with stable expression values or reference Input DNA samples are used for normalization.

In context of this thesis, the qPCR method was applied to validate the large amount of sequencing and microarray data. So far, the throughput was limited to 384 reactions per run by the limitations of the instruments applicable for routine laboratories. Commonly used instruments are for example the LightCycler 480 from Roche or the ABI PRISM 7900HT system from Applied Biosystems. At present, the recently launched LightCycler 1536 Instrument from Roche enables the analysis 1536 qPCR reactions per run. In manuscript 3 of this thesis, it is described how the new system was evaluated for its applicability. Subsequently, we applied the new system for the comprehensive analysis of ChIP experiments performed in mouse hearts that is described in manuscript 2.

2.5 Purpose and aims

Transcription factors are long time thought to be the main driving force directing normal cardiac development and function. However, how transcription factors regulate the overall transcription network and how they act in a combinatorial manner are still unanswered questions. The function of transcription factors is directly dependant on epigenetic factors and it has become apparent that also miRNAs play a powerful role in regulating cardiac gene expression. Despite many studies which address these different levels regulating cardiac development and function independently, we still lack data showing their combined interactions.

To gain insight into the transcriptional regulation of cardiac mRNA profiles we used a systems biology approach comprising technical methods such as ChIP-chip, ChIP-seq, RNAi mediated knockdown, microarray expression profiling as well as miRNA-seq. We aimed to elucidate transcriptional dependencies analyzing genome-wide DNA binding behavior of the four key transcription factors Gata4, Mef2a, Nkx2.5 and Srf and studied the impact of the co-occurring histone modifications H3ac, H4ac, H3K4me2 and H3K4me3 and miRNAs on target gene expression in the cardiomyocyte cell line HL-1. In addition, observed interdependencies between Srf and H3ac were further investigated in conjunction with H3K4me2 and the acetyltransferase p300 *in vivo* using mouse hearts from a time-series around birth.

qPCR is routinely applied in this thesis as method to verify the large amount of sequencing and microarray data as well as to analyze comprehensive ChIP experiments performed in mouse hearts. Consequently, we evaluated the novel LightCycler® 1536 Real-Time PCR System in context of a beta-side testing of the Roche Company for its applicability in medium and high-throughput experiments.

Finally, the chromatin remodeling factor DPF3 was characterized in more detail as it initially emerged as potentially cardiac relevant due to its up-regulation in patients with Tetralogy of Fallot. Different *in vivo* and *in vitro* biochemical methods were used to investigate the spatial and temporal expression of DPF3, to identify its interaction partners, and moreover to elucidate its molecular function.

3 MANUSCRIPT 1

The cardiac transcription network modulated by Gata4, Mef2a, Nkx2.5 and Srf, histone modifications and microRNAs

Jenny Schlesinger[#], Markus Schueler[#], Marcel Grunert[#], Jenny J. Fischer[#], Qin Zhang, Tammo Krueger, Martin Lange, Martje Tönjes, Ilona Dunkel, Silke R. Sperling. PLoS Genet. 2011 Feb;7(2)

[#]These authors contributed equally to this work.

doi:10.1371/journal.pgen.1001313

The original article is online available at

http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1001313

3.1 Synopsis

In this study we investigated the interplay of transcription factor binding, co-occurring histone modifications and miRNAs in regulating cardiac transcription networks in murine cardiomyocytes. The aim of our study was to understand how these molecular levels are involved in regulating cardiac transcription profiles and how they are connected to each other. Therefore, we focused on the four TFs Gata4, Mef2a, Nkx2.5, and Srf that are essential for heart development and function.[58,61,80,192]

In order to determine their direct target genes, we performed chromatin immunoprecipitation followed by array analysis (ChIP-chip) in the contracting cardiomyocyte cell line HL-1. We identified several hundreds of TF binding sites for each factor. The majority of binding sites was located within close proximity to the transcriptional start site. Consequently, we assigned to each binding site the corresponding target gene. We investigated which Gene Ontology (GO) terms were significantly overrepresented among these and found that most of the targets are involved in heart development and function. For example, the GO term “heart looping” was significantly overrepresented among Nkx2.5 targets, a process which is affected in Nkx2.5 knockout mice during embryonic development.[80]

As the expression of genes is mostly regulated by multiple TFs, we analyzed how often Gata4, Mef2a, Nkx2.5, and Srf bind to the same target genes. We observed frequent co-binding with more than 20% of genes bound by more than one TF indicating a complex and cooperative regulation. A penal of transcription factor binding sites was confirmed using quantitative real-time PCR. Furthermore, we validated a novel Srf binding site in the regulatory region of mouse miR-125b-1 as well as an Nkx2.5 binding element in the core promoter region of human *DPF3* using luciferase reporter gene assays.

To elucidate the functional consequence of reducing the quantity of the TFs, we performed siRNA-mediated knockdown in HL-1 cells for each factor and measured resultant mRNA profiles with genome-wide expression arrays. In conjunction with the primarily activating function of the TFs, the majority of deregulated transcripts were down-regulated in siRNA treated samples compared to control samples treated with unspecific siRNA. According to the observed frequent co-binding behavior, we found that the different factors share a similar proportion of differentially expressed genes in the respective knockdown. We confirmed a comprehensive set of differentially expressed genes by quantitative real-time PCR. In addition, we determined the overlap between direct target genes identified by ChIP and differentially expressed genes of the respective siRNA mediated knockdown. Both datasets shared GO terms reflecting heart and muscle development and function. Nevertheless, only a small fraction of directly bound genes (~10%) was concomitant

differentially expressed, and vice versa. This observation is in line with other studies[193-195] and alludes to the combinatorial nature of gene regulation and to potential buffering effects between Gata4, Mef2a, Nkx2.5, and Srf. Further bioinformatical analyses could substantiate this evidence as genes which were bound by multiple transcription factors, were significantly less likely differentially expressed. Likewise, TFs, which have a high number of common binding sites, share only a small number of co-regulated genes in siRNA mediated knockdown. There are several further explanations for the observed small overlap such as non-functional binding,[196] dosage dependant regulation by TFs[86,88] or redundant paralogs which can compensate the loss of one factor.[193] Moreover, TF binding depends on binding affinity and accessibility of binding sites, which is regulated amongst others by epigenetic mechanisms such as histone modifications.

In order to investigate to which extent histone modifications influence gene expression, we ananlysed the co-occurrence of four activating histone modifications (H3ac, H4ac, H3K4me2 and H3K4me3) and compared these with wild type expression levels of direct target genes. We found that around 80% of TF binding sites carried one or more of these histone modifications. Further, our data revealed a significant impact for H3ac on target gene expression. The expression levels of direct Nkx2.5 and Mef2a targets were significantly higher than the reference group, independent of whether H3ac was present or not. Whereas, target genes directly bound by Gata4 and Srf were only significantly higher expressed when they were additionally marked by H3ac. Finally, using ChIP followed by qPCR after Srf knockdown, we observed a significant decrease of H3ac at promoter regions, which are normally marked by H3ac and Srf in control cells. The correlation of H3ac and Srf target gene expression was further validated with genome-wide ChIP-seq experiments in HL-1 cells as well as with ChIP-qPCR using mouse hearts of three developmental stages around birth.

Finally, we investigated the impact of miRNAs on gene expression of target genes. Therefore, we focused on Srf, which is known to regulate cardiac relevant miRNAs.[159] Using our Srf ChIP-seq data, we could identify 22 miRNAs with a Srf binding site within a region of 10kb, including well-known cardiac and muscle relevant miRNAs like miR-208, miR-1 miR-125b, miR-133, miR-143 and miR145. Next, we carried out siRNA-mediated knockdown of Srf with subsequent quantification of miRNA expression levels using miRNA-seq technique. We identified 42 differentially expressed miRNAs and the majority of these were down-regulated, providing evidence that Srf mainly activates miRNAs. Consequently, we performed target prediction of all deregulated miRNAs to figure out how these might contribute to the deregulation of Srf target genes in the Srf knockdown. We identified 192 miRNA targets out of 429 differentially expressed Srf target genes, and a high fraction of these were found to be upregulated in Srf knockdown. Thus, the differential expression of

miRNAs potentially explain up to 45% of all differentially expressed genes and emphasize the essential impact of miRNAs as co- regulators of transcriptional profiles.

Finally, we constructed a transcription network centered on Srf where we integrated information from an extensive literature search and our results for Srf and H3ac obtained with ChIP-chip, ChIP-seq, miRNA-seq and siRNA-knockdown. This detailed Srf centered network presents the different nodes of regulating cardiac gene expression and can build a basis for further gene focused studies.

Taken together, our study provides insights into the complex transcriptional regulation of cardiac networks. We show the combinatorial influence of four transcription factors regulating cardiac mRNA profiles, and provide evidence that histone modifications and miRNAs modulate their functional consequence. We conclude that the different levels regulating the cardiac transcriptome should be viewed in context to each other.

3.2 Experimental contribution

For this work I performed and analyzed major parts of wet lab experiments including chromatin immunoprecipitation, RNAi transfection, RNA isolation, qPCR measurements, immunohistology and protein analyses. Furthermore, I participated in the discussion and conception of this study and wrote parts of the manuscript.

Contributions of co-authors:
Conception: SR. Sperling, JJ. Fischer
Performed the experiments: JJ. Fischer, Q. Zhang, M. Lange, M. Tönjes, I. Dunkel
Bioinformatic analysis: M. Schueler, M. Grunert, T. Krüger
Wrote the paper: M. Schueler, M. Grunert, SR. Sperling

4 MANUSCRIPT 2

Dynamics of Srf, p300 and histone modifications during cardiac maturation in mouse

Schueler M[#], Zhang Q[#], **Schlesinger J**, Tönjes M, Sperling SR

In preparation.

[#] Authors contributed equally.

4.1 Synopsis

In this study we aimed to gain a deeper understanding of the Srf-driven regulation of transcription in murine cardiomyocytes. Using chromatin immunoprecipitation (ChIP) followed by medium-throughput real-time qPCR, we investigated the dynamic binding behavior of Srf and the histone acetyltransferase p300 in correlation to their occurrence with H3ac and H3K4me2 at selected regulatory regions in a time-series of mouse hearts of three developmental stages.

Transcription factors are known to be major regulators of gene transcription. In addition, epigenetic modifications of chromatin and DNA have been recognized as important factors in controlling their activity. The serum response factor Srf is a very important transcriptional regulator for myogenic differentiation and cardiac development and function.[75] It binds a DNA consensus sequence known as CArG box which is predominantly found at promoters of muscle- and growth-factor associated genes.[64,197] Several studies have further demonstrated that activating histone modifications like H4ac and H3K4me2 are enriched at GArG boxes in particular at smooth muscle cell (SMC) gene promoters.[114,198-200] According to a model proposed by McDonald and Owens,[114] Srf together with its cofactor Myocardin is capable to bind and recruit other regulators like the histone acetyltransferase p300 what leads to further histone acetylation including H3ac and an activation of SMC gene transcription. Although the important role of Srf and the above mentioned activating histone modifications in SMC gene activation is well described, we lack information about the dynamics of their inter-regulatory dependencies in cardiomyocytes *in vivo*.

In our proceeding analysis using the contracting atrial cardiomyocytes cell line HL-1, we correlated gene expression profiles with the occurrence of histone modifications.[201] Using a bioinformatic analysis of variance (ANOVA) we unveiled a significant influence of H3ac on Srf target gene expression. These findings differ from studies in SMCs where H4ac was found to be the causative link between Srf and transcriptional activation.

Thus, to study the impact of H3ac and its dynamics in an *in vivo* situation, we carried out ChIP experiments for Srf, p300, H3ac and H3K4me2 with mouse hearts of three developmental stages shortly before and after birth: one prenatal stage (E18.5) and two postnatal stages (P0.5 and P4.5). For 191 potentially regulatory promoter regions of heart and muscle relevant genes, we designed corresponding TaqMan assays and measured the enrichment of bound DNA in triplicates for each individual time point and factor plus Input control using the LightCycler® 1536 Real-Time PCR System from Roche.

In order to test the reliability of our data, the mean enrichments were compared to the previous ChIP-chip/seq results gained in HL-1 cells. We found that regions bound by antibodies against H3ac, H3K4me2 and Srf in HL-1 cells were also significantly more

enriched in mouse hearts. Next, to determine alterations between the individual stages, we compared the average enrichment over all measured regions for each factor in each stage. We observed a general trend in the enrichment of bound DNA for each factor during the time-series that correlated with additionally measured protein level. As these overall trends in protein level were not the focus of our study but could lead to superficially high correlations between the time points, we carried out a second normalization step introducing a linear shift to remove these trends.

Finally, fold changes were calculated between mean enrichments of each two subsequent stages (E18.5/P0.5 and P0.5/P4.5) and pairwise correlation coefficients were computed. We found high to modest correlation for all studied factors. The highest correlations were obtained for H3K4me2/H3ac, H3ac/p300 and H3K4me2/p300. This is in line with other studies reporting that methylation of H3 recruits histone acetylation complexes including p300 and induces histone acetylation.[94,202-204] In addition, we observed significant correlation between Srf/H3ac and Srf/H3K4me2 what supports the model proposed by McDonald and Owens.[114] However, we found the lowest correlation for Srf and p300. The fact that this correlation was even lower than between Srf and H3ac is suggesting an additional mechanism for Srf-triggered acetylation, which is independent from p300. Moreover, using qualitative and quantitative linear modeling techniques changes in enrichment of H3ac were found to be significantly dependent according to changes in Srf and p300 binding both alone and in conjunction, which further substantiated the hypothesis of an additional Srf-H3ac coupled mechanism.

In a last step, we were interested whether observed changes in enrichment also have a functional influence on the expression level of respective genes. Hence, we measured expression levels of 44 genes associated with one or multiple analyzed regulatory regions and calculated fold changes for the individual stages. As a result, we found significant dependency between changes in gene expression and changes in H3ac enrichment as well as Srf binding, which is in line with the activating function of these two regulatory factors. Moreover, manual inspection revealed that many gene expression levels were found to be dependent on only a subgroup of the studied histone modifications and regulatory factors, respectively, indicating high variability of combinatorial regulation.

In summary, the study revealed a strong correlation between the occurrence of H3ac and H3K4me2 marks as well as Srf and p300 binding at potent regulatory regions of heart and muscle relevant genes pointing to a common regulatory mechanism which is triggered by Srf and has H3ac as outcome which is to a certain degree independent from p300. In addition, dynamic changes of the studied factors in the course of the three developmental stages were shown to have regulatory impact on expression levels of nearby genes.

4.2 Experimental contribution

For this study I performed about 50% of wet lab experiments. I designed and carried out the TaqMan qPCR measurement of ChIP samples together with Martje Tönjes. I analysed expression levels on mRNA and protein level together with Qin Zhang. In addition, I participated in the discussion and conception of the bioinformatic analysis together with Markus Schueler and wrote parts of the manuscript.

Contributions of co-authors:
Conception: SR. Sperling, M. Schueler
Performed the experiments: Q. Zhang, M. Tönjes
Bioinformatic analysis: M. Schueler
Wrote the paper: M. Schueler, SR.Sperling

Dynamics of Srf, p300 and histone modifications during cardiac maturation in mouse

Markus Schueler*, Qin Zhang*, Jenny Schlesinger, Martje Tönjes, Silke R. Sperling
*These authors contributed equally to this work.
In preparation.

Abstract

The adaptation of the cellular network to functional changes, timing and patterning of gene expression is regulated by binding of transcription factors to gene regulatory elements, which in turn depends on co-occurring histone modifications. These two layers influence each other, enabling a further level of regulatory fine-tuning.

We analyzed the interdependencies between histone 3 acetylation, histone 3 lysine 4 dimethylation, the transcription factor Srf and the histone acetyltransferase p300 in an *in vivo* model using chromatin immunoprecipitation in a time-series during cardiac maturation in mouse.

We found a strong correlation between the presence of the two histone modifications and binding of Srf and p300. Using linear modeling techniques we could show that each factor contributes individually as well as conjointly to histone 3 acetylation and gene expression, probably aided by accompanying histone 3 lysine 4 dimethylation. We further demonstrate that changes in gene expression during cardiac maturation are attended by changes of the analyzed regulators while revealing a high variability of combinatorial regulation. Finally, we propose a model of Srf-driven gene expression in cardiomyocytes.

Introduction

Transcription factors can regulate the expression of genes in a tissue-specific and quantitative manner and are thus major regulators of embryonic developmental processes. The ability of transcription factors to bind to DNA is highly dependent on the accessibility of their binding sites. The majority of genomic DNA in eukaryotes is packaged into chromatin by association with histone proteins. Covalent modifications of histone tails and ATP-dependent chromatin remodeling facilitate access for DNA-binding transcription factors.[1] Thus, to understand networks directing gene expression the interplay between transcription factors, co-regulatory elements and epigenetic factors has to be considered.

Serum Response Factor (Srf) is an ubiquitously expressed MADS (MCM1, Agamous, Deficiens, SRF) box transcription factor[2] and binds to a DNA consensus sequence known as the CArG box [$CC(A/T)_6GG$]. Despite the ubiquitous expression of SRF, the CArG box is predominantly found in promoters of muscle- and growth-factor-associated genes.[3-4]

Functional analyses showed that embryonic stem cells lacking Srf display defects in spreading, adhesion and migration. These defects correlate with defective formation of cytoskeletal structures.[5] In Srf-null neonatal cardiomyocytes, severe defects in the contractile apparatus are observed.[6] Srf-deficient mice embryos (*Srf*$^{-/-}$) show a gastrulation defect, fail to develop the mesoderm and subsequently die between E8.5 and E12.5.[7] Cre mice with knockout of Srf in 80% of the cardiomyocytes display severe heart defects and die at E11.5[8] and Cre mice lacking skeletal muscle Srf expression die during the perinatal period from severe skeletal muscle hypoplasia.[9]

Studies analysing co-occurring histone modifications showed that histone H3 and H4 acetylation (H3ac and H4ac) are enriched at CArG boxes in cultured smooth muscle cells (SMCs).[3,10-12] Additionally, SMC genes show higher levels of H3 Lys4 dimethylation (H3K4me2) relative to non-SMC lineages.[3] A model of step-wise molecular events directing the activation of SMC gene transcription has been depicted by McDonalds and Owens[13] (Fig. 1 A). It comprises the combinatorial enrichment of both H3K4me2 and H4ac at regulatory CArG boxes of SMC genes. H4ac is thought to lead to an open chromatin state facilitating genomic access while H3K4me2 provides a docking site to the Srf/Myocardin complex.[13] Myocardin, one of the most potent co-activators of Srf, belongs to the SAP domain family of nuclear proteins and is exclusively expressed in SMCs and cardiomyocytes.[14] Binding of the Srf/Myocardin complex to the CArG box promotes the recruitment of other transcription factors via the powerful C-term transcriptional activation domain of Myocardin. Thus, the recruitment of the acetyltransferase p300 establishes additional histone acetylation marks which in turn activate transcription.[13] p300 is a known histone acetyltransferase and transcriptional co-activator, which was found to reside in enhancer and promoter regions.[15-16]

The model proposed has been established based on studies in SMCs, however, a systematic validation in an *in vivo* model comprising the dynamics of the interregulatory dependencies is pending. Further, the relevance of factor interaction for gene transcription in cardiomyocytes is unknown. Recently, we showed that the transcriptional activity of Srf in the mouse cardiomyocyte cell line HL-1 is highly depending on the co-occurrence of H3ac. In contrast to SMCs H4ac marks do not enhance Srf-driven gene expression in cardiomyoyctes.[17] To gain further insights, we used a time-series approach of mouse hearts and monitored the combinatorial regulation of gene expression driven by Srf, H3ac, H3K4me2, and p300. The dynamic changes of transcription factor binding and depletion or enrichment of histone modifications were studied during the cardiac maturation process occurring at the prenatal and postnatal developmental phase (E18.5, P0.5 and P4.5). Based on our previous genome-wide study[17] we selected 191 cardiac enhancer and promoter regions for this focused approach and used chromatin immunoprecipitation (ChIP) followed by high-throughput qPCR to quantify respective dynamics. In addition, we measured the

corresponding changes of genes expression showing the functional consequences of the observed alterations of factor binding and histone modifications. Based on our findings we propose a model of Srf-driven gene expression in cardiomyocytes (Fig. 1 B).

Results and discussion

Our previous analysis performed in a cardiomyocyte cell line (HL-1) showed that the activating potential of Srf is significantly depending on accompanying H3ac.[17] In contrast to the Srf-driven gene expression model proposed for SMC, we do not observe a dependency for H4ac or H3K4me2 (Fig. 2). To further elaborate on the dynamic relationship between Srf and histone modifications, we switched towards an *in vivo* time-series model of cardiac maturation in mouse and investigated the relationship between binding of Srf, the acetyltransferase p300 and H3ac as well as H3K4me2. We included H3K4me2 in our further studies even so we did not observe an enhancement of Srf-driven gene expression in our cell line analyses. H3K4me2 had been shown to function as a stabilizing element for Srf binding and therefore might act in concert with H3ac. This could potentially be underlined by our previous observation that co-localization of H3ac and H3K4me2 is associated with higher gene expression than isolated H3K4me2 marks.[18]

We used mouse hearts of three developmental stages (Fig. 3 A), one prenatal stage (E18.5) and two postnatal stages (P0.5, P4.5). From the fetal to the postnatal stage, the heart adapts to the body circulation and cardiomyocytes mature. During this process the heart increases in size, the cells elongate, myofibrils align and cell-cell contacts become bipolar.[19]

As an initial step, we determined total protein levels of Srf, p300, H3ac, H3K4me2, total H3 and GAPDH (normalization control) at all stages using Western Blot followed by densitometric measurements of the resulting bands (Figure 3 B and C). Due to alternative splicing events of Srf, four different isoforms are expressed in the heart,[20-21] showing only slight changes during the time-series. The level of p300 is strongly reduced at P4.5 (83% compared to E18.5). Both, the levels of H3ac and H3K4me2 show a high reduction directly after birth (80-95% compared to E18.5) whereas the total amount of H3 remains nearly constant.

We selected enhancer and promoter regions with potential Srf-driven regulatory influence on gene expression based on our previous genome-wide ChIP-chip/seq data of histone modifications and Srf binding (see Fischer *et al.*[18] and Schlesinger *et al.*[17]). A priori 422 genomic regions that shared enrichment for at least two of H3ac, H3K4me2 and Srf binding were selected located in a 10kb window of genes involved in cardiac development. We used a multistep filtering approach to focus on genes with either heart or muscle annotation as a GO term, an entry of cardiac disease in OMIM or relevance in a manually curated literature

search. Genomic regions of a maximum of 500bp in length were defined (Supplemental Fig. S1) and Taq Man assays for the following ChIP analysis were established for every window. Windows without unique primer pairs were discarded from the analysis. This resulted in a total of 191 genomic regions in the proximity of heart and muscle developmental genes, which were consequently analyzed.

ChIP followed by real-time qPCR was performed for each factor plus input DNA for each region with samples of mouse hearts of the time points E18.5, P0.5 and P4.5 in triplicates. This resulted in a total of 7578 measurements, which fulfilled our quality standards. Data normalization was performed computing ΔCP values using the measured input DNA. As a proof of principle, the resulting ChIP enrichments from stage P0.5 were compared to our previous ChIP-chip/seq results in HL-1 cells (Supplemental Fig. 2), which confirmed the reliability of the system. As we were interested in the dynamic changes of factor binding and histone modification over the time course, we checked for a potential overall bias, which might be introduced by the observed changes of the protein content during the time course. Consequently, we compared the average ChIP enrichment over all regions for each factor in each stage and found differences between the three stages (Fig. 3 D and Supplemental Fig. S3 A) recapitulating the findings from the overall protein analysis (Fig. 3 B and C). We eliminated this bias, as it would result in a general trend between individual stages and superficially high correlations between the individual factors. Therefore, we introduced a fixed linear shift for each single region and each single factor resulting in the same average enrichment in each stage for each factor (Supplemental Fig. S3 B). Finally, we calculated for each factor fold changes of ChIP enrichments between consecutive stages and grouped all changes into the categories significantly up, significantly down or unchanged according to a t-test using a significance level of 95%.

As a first analysis, scatter plots were used to visualize the level of correlation between the fold changes of individual factors (Fig. 4). We found high to modest correlation for all the measured factors, with Pearson correlation coefficients ranging from 0.753 to 0.364, which implicates that all factors are related to each other. Statistical significance of found correlation coefficients was evaluated using random experiments. As a null model, triplicate measurements for one factor and time point were randomly assigned to all regions without replacement and fold changes were again computed between successive timepoints. This process was repeated 20,000 times and an empirical p-value was derived by counting the number of tries were the random correlation coefficient exceeded the true coefficient. Applying a significance level of 99.9% we found significant correlation for all of the analyzed pairings (Supplemental Fig. S4).

The highest correlation coefficient of 0.753 was found between the changes in H3ac and H3K4me2 which is in accordance to the high correlation of absolute H3ac and H3K4me2 levels found by us[17] and others.[22] A recent study performed by Wang *et al*.[23] suggested that H3K4 methylation might directly facilitate histone acetylation events. The second highest correlation coefficient (0.593) was found between changes in H3ac enrichment and p300 followed by p300 and H3K4me2 (0.547). While the correlation between changes in p300 and acetylation level was highly expected given the function of p300, the correlation between p300 and methylation level is of high interest. A possible explanations for this correlation supposed by Pray-Grant *et al*.[24] and Wysocka *et al*.[25] is an initial opening of chromatin through recruitment of ATP-dependent chromatin remodeling complexes induced by histone methylation which in turn allows p300 to bind. Another possibility might be the direct recognition of methylated sites by histone acetylation complexes (including p300) as suggested by Martin *et al*.[26] In addition, significant correlation was found between Srf and H3ac (0.489) as well as Srf and H3K4me2 (0.481) supporting the model by McDonald and Owens[13] which suggest stabilization of Srf binding to the CArG box DNA motif via Myocardin or some Myocardin related factors that directly bind methylated histones. Further, this finding is in line with our previous observation in HL-1 cells such that siRNA mediated knockdown of Srf leads to a significant reduction of H3 acetylation at related genomic sites.[17] Consistent with the model in SMCs, significant correlation was found between changes in Srf and p300. However, given that the observed correlation (0.364) between these two factors is much weaker than between any of the other measured factors and especially between Srf and H3ac might suggest additional mechanisms of Srf-triggered acetylation, which are independent from p300 binding.

While our first analysis of correlation between changes in the individual factors was based on the combination of fold changes between all three stages, we also considered each fold change individually. In general, fold changes as well as correlation coefficients were higher for changes between E18.5 and P0.5 than between P0.5 and P4.5 while, except for Srf and p300, still remaining significant. A possible explanation is that the physiological and cellular change reflected in the molecular setting is greater from the pre- towards the postnatal phase than later on. This might cause a higher influence of experimental noise on the correlation.

To analyze the interdependency in regulation of the measured factors, we used ANOVA linear models to predict changes in H3ac modifications levels from changes in binding levels of p300 and Srf. We excluded the data for H3K4me changes as these were shown to be highly correlated with the dependent variable H3ac and could therefore obscure the models. After estimating coefficients for each of the predictive groups using a least-square approach, it was tested whether any of the coefficients was significantly different from zero implying a

significant influence on changes in H3ac level. As before, randomized experiments were used to derive empirical p-values for the individual models, now using the *percentage of variance explained* or R^2 as a measurement for the goodness of fit for each model.[27] The empirical p-value is therefore determined by the number of random models which yielded a higher R^2 when compared to the original model. Two single-factor ANOVAs were calculated using only categorized changes in either p300 or Srf to predict changes in H3ac (Fig. 5 A and B). Both models indicated a significant dependency of H3ac on the individual factor with an empirical p-value < $5x10^{-5}$. The estimated coefficients for up and down regulated regions were significantly higher and lower than zero. The distribution of the residuals and the normal Q-Q plots indicated validity of the linear model assumptions for both single-factor ANOVAs (Supplemental Fig. S5).

After proving the dependency of H3ac changes on each Srf and p300 individually, we were interested in the interaction between the two and if the observed changes were mainly dependent on the interplay between the two factors or on both factors independently. Therefore, a two-factor ANOVA was performed using categorized changes in both Srf as well as p300 as predictors for changes in H3ac (Fig. 5 C). Again, the model indicated significant dependency with an empirical p-value < $5x10^{-5}$. Like in the single-factor ANOVA, the estimated coefficients for the influence of up and down regulation of Srf and p300, which can be found in Tab. 1, were significantly different from zero, with the exception of up-regulated Srf regions. As a result, changes in both p300 and Srf had a significant influence on H3ac. Reduction of Srf binding on top of reduced p300 levels leads to further diminishment of H3ac, which substantiates the assumption that Srf triggers H3ac additionally by a p300 independent mechanism. Like for the single-factor ANOVAs, the distribution of the residuals and the normal Q-Q plots indicated validity of the linear model assumptions (Supplemental Fig. S6).

To gain insights of H3ac levels in regions with steady state levels of Srf and p300, we used density plots of mean levels for H3ac, p300 and Srf over all regions. We were interested if these stable Srf and p300 regions had a general preference in their absolute H3ac levels, such that higher p300 levels are associated with higher H3ac levels. Fig. 5 D shows the density plot of mean levels for H3ac over all regions. It indicates a bimodal distribution with two peaks that correlate with regions of high and low acetylation levels. The same two peaks were found in a two-dimensional density plot for absolute H3ac and p300 levels (Fig. 5 E). We did not observe a bimodal distribution for Srf (data not shown). However, in both plots the two peaks are not perfectly separated, indicating many regions with intermediate H3ac and p300 levels.

In line with this quantitative assessment of H3ac in steady state regions of Srf and p300, we were interested if the dynamics of H3ac levels were also quantitatively dependent on the dynamics of Srf and p300 levels. In other words, if stronger changes in any of these two factors lead to stronger changes in H3ac and which of the two factors is more influential. Therefore, a linear regression model was computed predicting the strength of H3ac fold changes from Srf and p300 fold changes directly. Like for the previous models, an empirical p-value was derived from random experiments. Again, the residuals and the Q-Q plots revealed validity of the linear model assumptions (Supplemental Fig. S7 A and B).
The linear regression model revealed that changes in H3ac level were not only qualitatively but also quantitatively dependent on changes in Srf as well as p300 level ($p < 5x10^{-5}$). In line with the results from the qualitative model, both factors showed a significant influence and to a roughly equal extend according to the high similarity in their estimated coefficients (Tab. 2). This finding further questions the proposed dependency between the two factors in the analyzed time-series, indicating p300-independent mechanisms of Srf-triggered H3ac. Finally, we computed the Pearson correlation coefficient between predicted and observed changes reaching an overall model accuracy of 0.67 (Supplemental Fig. S7 C). While far from perfect, this correlation demonstrates that changes in the two factors Srf and p300 already drive an appreciable fraction of the measured changes in histone 3 acetylation levels.

So far we used changes in H3ac level as the predicted read-out of changes in Srf and p300 level. In a last step, we were interested whether the observed changes also had a functional consequence on gene expression. Therefore, gene expression levels for 44 genes associated to one or multiple analyzed regulatory regions were measured in the same three stages using quantitative real-time PCR. Like for the ChIP measurements, a normalization step was introduced to remove potential trends in the mean expression over all measured genes between the three individual stages and fold changes were calculated to measure the effective change between the individual stages. To check whether changes in the individual factors lead to changes in gene expression, genes were grouped according to significant up or down regulated regions for each factor combining the two time points. Using a t-test between the two groups, we observed significant dependencies for changes of gene expression to changes of H3ac and Srf levels ($p = 0.017$ and $p = 0.03$ respectively, Fig. 6 A and B) and only an insignificant trend for p300 levels (data not shown).
In addition, we analyzed the dependency of gene expression changes on all measured factors and histone modifications. While the genes in general showed the proposed dependency on binding or modification changes, several were regulated by only subgroups of these (representative examples are shown in Fig. 6 C-H). This finding likely indicates a

high variability of combinatorial regulation between the investigated regulators. Though, it has to be kept in mind that we measured only three consecutive timepoints and might therefore have missed intermediate regulatory events.

Summary and conclusion

We investigated the interdependency between Srf and H3ac in conjunction with H3K4me2 and the acetyltransferase p300 *in vivo* using a time-series of maturating murine cardiomyocytes around birth. With respect to the limited amount of cardiac material, regulatory regions were selected that showed binding of at least two of the three factors Srf, H3ac and H3K4me3 in a preceding HL-1 study. These regions were subsequently profiled for their enrichment with the analyzed factors and histone modifications. Investigating changes between the successive stages, binding of the individual factors was in general found to be very dynamic with many significant changes between the individual stages. Most importantly, these changes showed high levels of correlation pointing to strong interdependencies. While a high correlation between changes of the histone acetyltransferase p300 and acetylation level was expected, the link between the other factors is not straightforward. Interestingly, the highest correlation was found between changes in H3ac and H3K4me2 levels. This supports findings from Wang *et al.*[23] in human CD4+T cells which revealed that histones in promoters that showed initial H3K4 methylation were acetylated after treatment with Trichostatin A, an histone deacetyltransferase inhibitor, while those promoters that did not had initial H3K4 methylation remained largely unacetylated. The coupling between histone tail methylation and acetylation might either be achieved through chromatin remodeler binding both modifications such as Dpf3b (Baf45c)[28] or BPTF,[29-30] the recruitment of ATP-dependent chromatin remodeling complexes which open the chromatin to allow other histone modification enzymes to bind[24-25] or by direct H3K4me2 recognition from histone acetylation complexes.[26]

Further, it has been suggested that H3K4 methylation increases the efficiency of p300 function and consecutive histone acetylation.[31] In line with this we found significant correlation between changes in H3K4me2 and p300 levels and finally coupling to H3ac, which might likely be established through tissue specific transcription factors. Following a model suggested by McDonald and Owens,[13] gene activation by Srf is performed in a step-wise procedure. In order to bind its CArG-box motif, initial histone modifications like H3K4me2 and H4ac must be present in smooth muscle cells. H4ac thereby leads to an open chromatin state facilitating genomic access while H3K4me2 provides a docking site to the Srf/Myocardin complex. This complex has been suggested to recruit further transcription factors and transcriptional co-factors including p300 that fully activate the gene expression program. Here, we focused on cardiomyocytes maturation using an *in vivo* time-series model

and suggest that H3ac instead of H4ac plays a crucial role for Srf-driven gene expression. Both quantitative and qualitative modeling of changes in H3ac showed a significant dependency on Srf and p300 levels either alone or in conjunction. A key question was whether the implemented models would substantiate the stepwise regulatory model in which H3ac levels are dependent on Srf binding through recruitment of p300 or if the presence of both factors independently influences H3ac levels. We found a minor correlation between Srf and p300 levels, however, both qualitative as well as quantitative linear modeling revealed a significant influence of Srf as well as p300 on H3ac levels, pointing to independent regulation. We therefore strongly believe that Srf-associated H3ac is not only achieved via the recruitment of p300 but is established by the interaction with additional histone acetylation factors. Furthermore, we found for cardiac relevant genes that gene expression levels and the presence of H3ac and Srf binding are highly correlated which does not apply for p300. Notably, on the single-gene level a panel of regulatory dependencies can be observed with dynamic subsets of regulators that co-operatively work to regulate correct temporal transcription of target genes.

Materials and methods

Cardiac samples

Mouse hearts at the indicated stages of CD1 strain were dissected in cold PBS from the rest of the body. For subsequent RNA isolation heart samples were directly snap frozen in liquid nitrogen and stored at -80°C. For ChIP experiments heart samples were minced and fixed with 1% formaldehyde solution.

Chromatin immunoprecipitation followed by qPCR

Chromatin immunoprecipitation (ChIP) experiments were performed using 30mg mouse heart tissue each. Tissue samples were minced and cross-linked for 15min at room temperature with fixation buffer (1% formaldehyde, 10mM NaCl, 100µM EDTA, 50µM EGTA, 5mM HEPES, pH 7.5). Cross-linking was terminated by addition of glycine to a final concentration of 125mM. Tissue samples were homogenized for disaggregation. Subsequently, cells were incubated for 15min at 4°C with lysis buffer (50mM HEPES, pH 7.5, 140mM NaCl, 1mM EDTA, 1% Triton, 0.1% Na-deoxycholate, 0.1% SDS, 1mM PMSF, 1x Complete Protease Inhibitor Cocktail (Roche)), collected by centrifugation and homogenized using a Douncer. Nuclei were collected by centrifugation for 15min at 20.000g and 4°C and resuspended in sonification buffer (10mM Tris-HCl, pH 8.0, 1mM EDTA and 0.5mM EGTA, 1mM PMSF, 1x Complete Protease Inhibitor Cocktail). The chromatin was fragmented by sonification to an average size of 500bp. For immunoprecipitation buffer conditions were adjusted to the following concentrations: 140mM NaCl, 1% Triton, 0.1% SDS, 0.1% Na-

deoxycholate). A fraction of material was saved as “Input”. Chromatin was pre-cleared by rotation with Dynabeads Protein A/G (Invitrogen) for 1h at 4°C. Antibodies were added as given in Supplemental Tab. 1 and incubated on a rotating wheel over night at 4°C. Protein A/G beads were added and rotation continued for 1h. The beads were washed twice with RIPA buffer A (10mM Tris- HCl, pH8.0, 140mM NaCl, 0.025% NaN3, 1% Triton, 0.1% SDS, 1% Na-deoxycholate), once with RIPA buffer B (10mM Tris-HCl, pH8.0, 500mM NaCl, 0.025% NaN3, 1% Triton, 0.1% SDS, 1% Nadeoxycholate), once with LiCl detergent solution (10mM Tris-HCl, pH8.0, 0.5% Na-deoxycholate, 1mM EDTA, 250mM LiCl, 0.5% Nonidet P-40) and once with TBS (20mM Tris-HCl, pH7.6, 150mM NaCl). Immunocomplexes were disrupted by eluting 10min at 65°C with 1% SDS/TE buffer (10mM Tris-HCl, pH7.6, 1mM EDTA, 1% SDS) and eluting a second time for 15min with 0.67% SDS/TE buffer (10mM Tris-HCl, pH7.6, 1mM EDTA, 0.67% SDS). Eluates were combined and reverse cross-linked by heating at 65°C over night. Subsequently, DNA was treated with RNase A and Proteinase K and purified by phenol-chloroform extraction.

ChIP samples of Srf, p300, H3ac and H3K4me2 for 191 regions (available upon request) were analyzed with TaqMan qPCR in context of a beta-side testing of the LightCycler® 1536 System from Roche Applied Science. TaqMan qPCR assays were designed using the Universal ProbeLibrary (UPL) and the ProbeFinder software which is freely accessible at www.roche-appliedscience.com. Primer sequences are available upon request. For the 1536-well plate format the Innovadyne™ Nanodrop™ Express pipetting robot (IDEX Health & Science LLC, Rohnert Park, CA USA) and the Velocity11 sealing machine (Agilent, Santa Clara, CA USA) were used. ChIP analysis was carried out in triplicates using the following program: Enzyme activation: 95°C for 1min; amplification (45cycles): 95°C for 1sec (ramp: 4.8°C/sec), 60°C for 30sec (ramp: 2.5°C/sec); cooling: 40°C for 30sec (ramp: 2.5°C/sec). The detection format was set to ‘Mono Color Hydrolysis/UPL Probes’ and the pipetting control to ‘Master Control’. The qPCR reaction contained 0.8µl ChIP sample and 1.2µl mastermix containing 1x RealTime ready DNA Probes Master, 300nM primer and 400nM probe.

Preparation of protein extracts and Western Blot analysis

Protein extraction and Western Blot was performed according to standard protocols. All antibodies with their respective dilution are given in Supplemental Tab. 1. Signals on the blots were subjected to densitometry measurements using NIH’s ImageJ software.

mRNA expression analysis

Total RNA of heart tissues was isolated using TRIzol reagent (Invitrogen) followed by DNase digest (Promega) and ethanol precipitation according to standard protocols. Reverse

transcription reactions were carried out via AMV-RT (Promega) with random hexamers (Amersham Pharmacia Biotech). To verify transcript expression levels of mouse hearts, quantitative real-time PCR measurements were performed using Absolute qPCR SYBR Green ROX Mix (Thermo Scientific) and the ABI PRISM 7900HT Sequence Detection System according to manufacturer's protocol. Gene expression was calculated using the ΔCT method with normalization to the housekeeping gene Hprt. Genes and corresponding primer sequences are available upon request.

Bioinformatic analysis

All computations, linear models, statistical tests and plots were carried out using R.[32] The two-dimensional density plot was computed using the *kde2d* function from R's MASS package.[33] Gene annotations were based on Ensembl, version 55.

To assess the influence of the individual factors on gene expression for genes with multiple assigned ChIP regions, the changes for these regions were combined using the following algorithm for each factor and time point comparison individually: At first, changes were categorized into up, down and unchanged as described in the manuscript. In a second step, genes that were associated to both up and down regulated regions were discarded from the analysis. Finally, genes which were associated to at least one region with a significant change were categorized accordingly and genes associated to only unchanged regions were discarded.

References

1. Simone, C. SWI/SNF: the crossroads where extracellular signaling pathways meet chromatin. *J Cell Physiol* **207**, 309-14 (2006).
2. Miano, J.M. Serum response factor: toggling between disparate programs of gene expression. *J Mol Cell Cardiol* **35**, 577-93 (2003).
3. McDonald, O.G., Wamhoff, B.R., Hoofnagle, M.H. & Owens, G.K. Control of SRF binding to CArG box chromatin regulates smooth muscle gene expression in vivo. *J Clin Invest* **116**, 36-48 (2006).
4. Sun, Q. *et al.* Defining the mammalian CArGome. *Genome Res* **16**, 197-207 (2006).
5. Schratt, G. *et al.* Serum response factor is crucial for actin cytoskeletal organization and focal adhesion assembly in embryonic stem cells. *J Cell Biol* **156**, 737-50 (2002).
6. Balza, R.O., Jr. & Misra, R.P. Role of the serum response factor in regulating contractile apparatus gene expression and sarcomeric integrity in cardiomyocytes. *J Biol Chem* **281**, 6498-510 (2006).
7. Arsenian, S., Weinhold, B., Oelgeschlager, M., Ruther, U. & Nordheim, A. Serum response factor is essential for mesoderm formation during mouse embryogenesis. *Embo J* **17**, 6289-99 (1998).
8. Miano, J.M. *et al.* Restricted inactivation of serum response factor to the cardiovascular system. *Proc Natl Acad Sci U S A* **101**, 17132-7 (2004).
9. 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* **102**, 1082-7 (2005).
10. Cao, D. *et al.* Modulation of smooth muscle gene expression by association of histone acetyltransferases and deacetylases with myocardin. *Mol Cell Biol* **25**, 364-76 (2005).
11. Manabe, I. & Owens, G.K. Recruitment of serum response factor and hyperacetylation of histones at smooth muscle-specific regulatory regions during differentiation of a novel P19-derived in vitro smooth muscle differentiation system. *Circ Res* **88**, 1127-34 (2001).
12. Qiu, P. & Li, L. Histone acetylation and recruitment of serum responsive factor and CREB-binding protein onto SM22 promoter during SM22 gene expression. *Circ Res* **90**, 858-65 (2002).
13. McDonald, O.G. & Owens, G.K. Programming smooth muscle plasticity with chromatin dynamics. *Circ Res* **100**, 1428-41 (2007).
14. Wang, D. *et al.* Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* **105**, 851-62 (2001).
15. Heintzman, N.D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311-8 (2007).
16. Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854-8 (2009).
17. Schlesinger, J. *et al.* The Cardiac Transcription Network Modulated by Gata4, Mef2a, Nkx2.5, Srf, Histone Modifications, and MicroRNAs. *PLoS Genet* **7**, e1001313 (2010).
18. Fischer, J.J. *et al.* Combinatorial effects of four histone modifications in transcription and differentiation. *Genomics* **91**, 41-51 (2008).
19. Hirschy, A., Schatzmann, F., Ehler, E. & Perriard, J.C. Establishment of cardiac cytoarchitecture in the developing mouse heart. *Dev Biol* **289**, 430-41 (2006).
20. Davis, F.J. *et al.* Increased expression of alternatively spliced dominant-negative isoform of SRF in human failing hearts. *Am J Physiol Heart Circ Physiol* **282**, H1521-33 (2002).
21. Kemp, P.R. & Metcalfe, J.C. Four isoforms of serum response factor that increase or inhibit smooth-muscle-specific promoter activity. *Biochem J* **345 Pt 3**, 445-51 (2000).
22. Koch, C.M. *et al.* The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res* **17**, 691-707 (2007).

23. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019-31 (2009).
24. Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd & Grant, P.A. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433**, 434-8 (2005).
25. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859-72 (2005).
26. Martin, D.G., Grimes, D.E., Baetz, K. & Howe, L. Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. *Mol Cell Biol* **26**, 3018-28 (2006).
27. Faraway, J. *Practical Regression and Anova using R*, (2002).
28. Lange, M. *et al.* Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev* **22**, 2370-84 (2008).
29. Li, H. *et al.* Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91-5 (2006).
30. Ruthenburg, A.J., Li, H., Patel, D.J. & Allis, C.D. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**, 983-94 (2007).
31. Wang, H. *et al.* Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Mol Cell* **8**, 1207-17 (2001).
32. *R: A language and environment for statistical computing*, (2005).
33. Venables, W.N., Ripley, B. & Venables, W.N. *Modern applied statistics with S*, (Springer, 2002).

Figures and tables

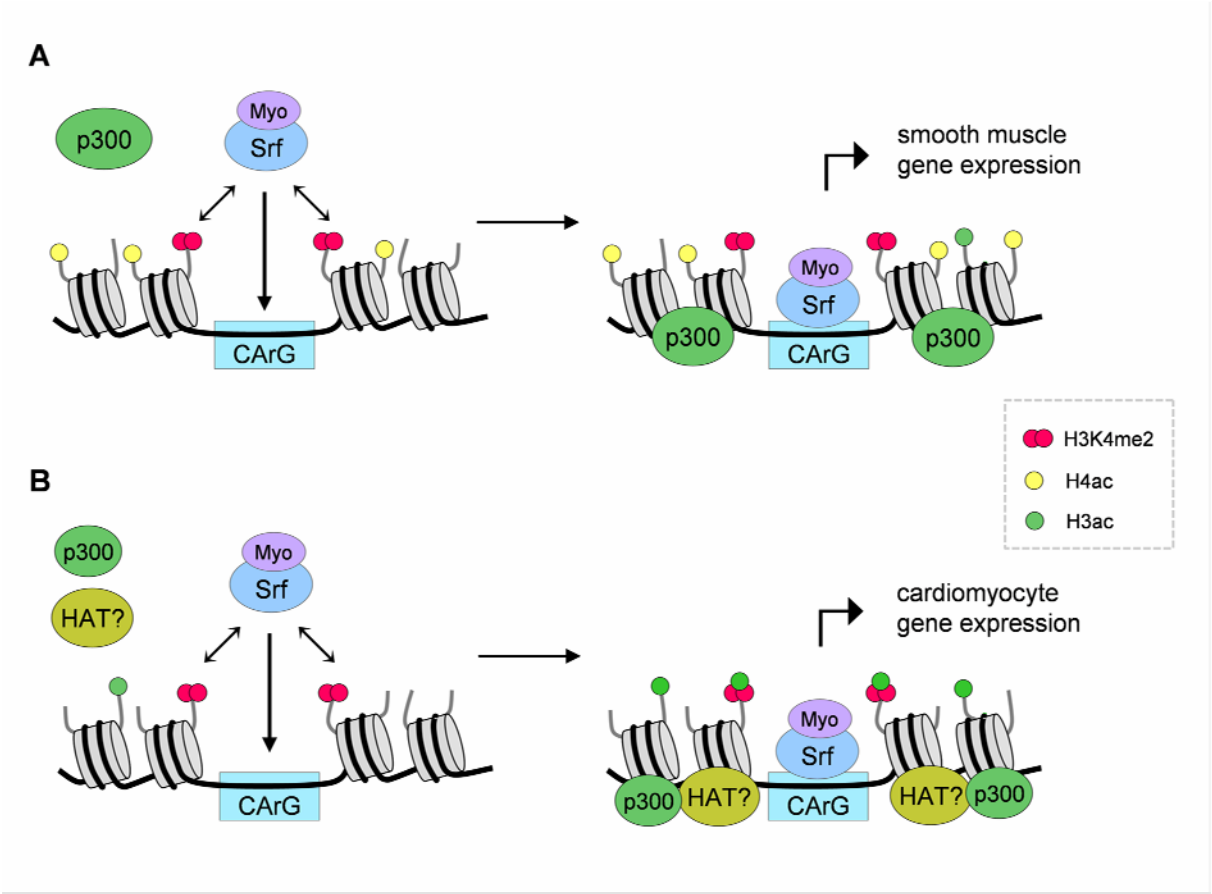
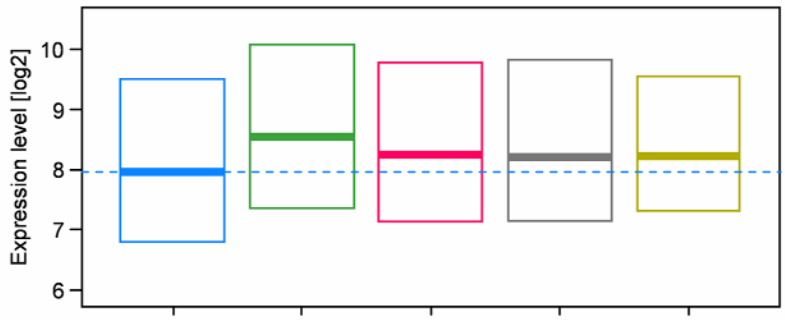


Fig. 1: (A) Model for smooth muscle cell (SMC) gene activation adapted from McDonald and Owens.[13] H4ac and H3K4me2 facilitate the binding of the Srf/Myocardin complex to regulatory CArG boxes. The complex then recruits p300, leading to further histone acetylation and active gene expression. (B) Model for Srf-driven gene expression in cardiomyocytes. In contrast to SMC, H3ac takes over the role of H4ac in facilitating gene expression. Furthermore, histone acetylation by p300 is accompanied by additional so far unexplored histone acetyltransferases (HAT).



	Srf only	H3ac	H3K4me2	H3K4me3	H4ac
Coefficient	/	1.035	-0.106	-0.397	-0.223
p-value	/	4.4×10^{-6}	0.633	0.092	0.237
Transcripts	611	267	273	270	208

Fig. 2: The influence of histone modifications at Srf binding sites on expression levels. Boxplots for the expression of transcripts with Srf binding site and specific histone modifications compared to genes with Srf binding site but no histone modification. A single transcript can belong to multiple groups if it has multiple histone modifications. Estimates for the influence of individual histone modifications and their associated p-values are derived by an ANOVA model. Figure taken from Schlesinger *et al*.,[17] modified.

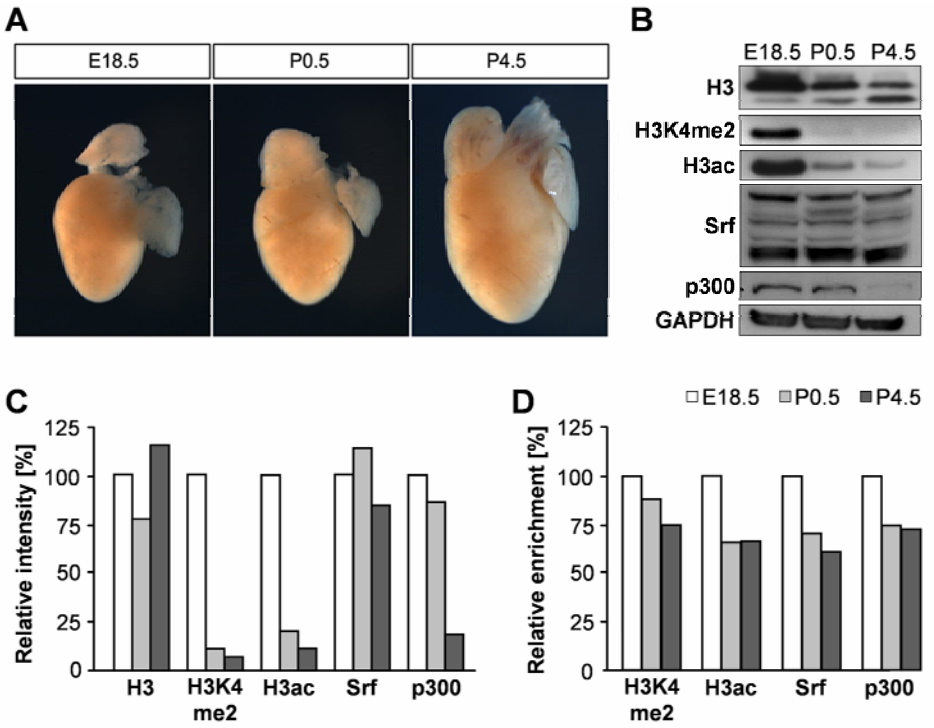


Fig. 3: (A) Three stages of mouse hearts during cardiomyocyte maturation around birth. (B) Western Blot analysis of protein expression in mouse hearts for Srf, p300, H3ac, H3K4me2 and H3. GAPDH served as loading control. Srf is expressed in four different isoforms. The H3 level is divided in phosphorylated H3 (upper band) and non-phosphorylated H3 (lower band). (C) Barplot of relative protein expressions of Srf, p300 and H3 and relative levels of histone modifications H3ac and H3K4me2. Total protein amounts were gathered using densitometry measurements. All intensities are normalized to GAPDH and shown relative to E18.5. (D) Barplot of relative average ChIP level over all regions for every factor in every stage. Measurements after ΔCP normalization showing a distinct trend between the individual stages. Consequently, an additional linear shift was introduced to remove the trend.

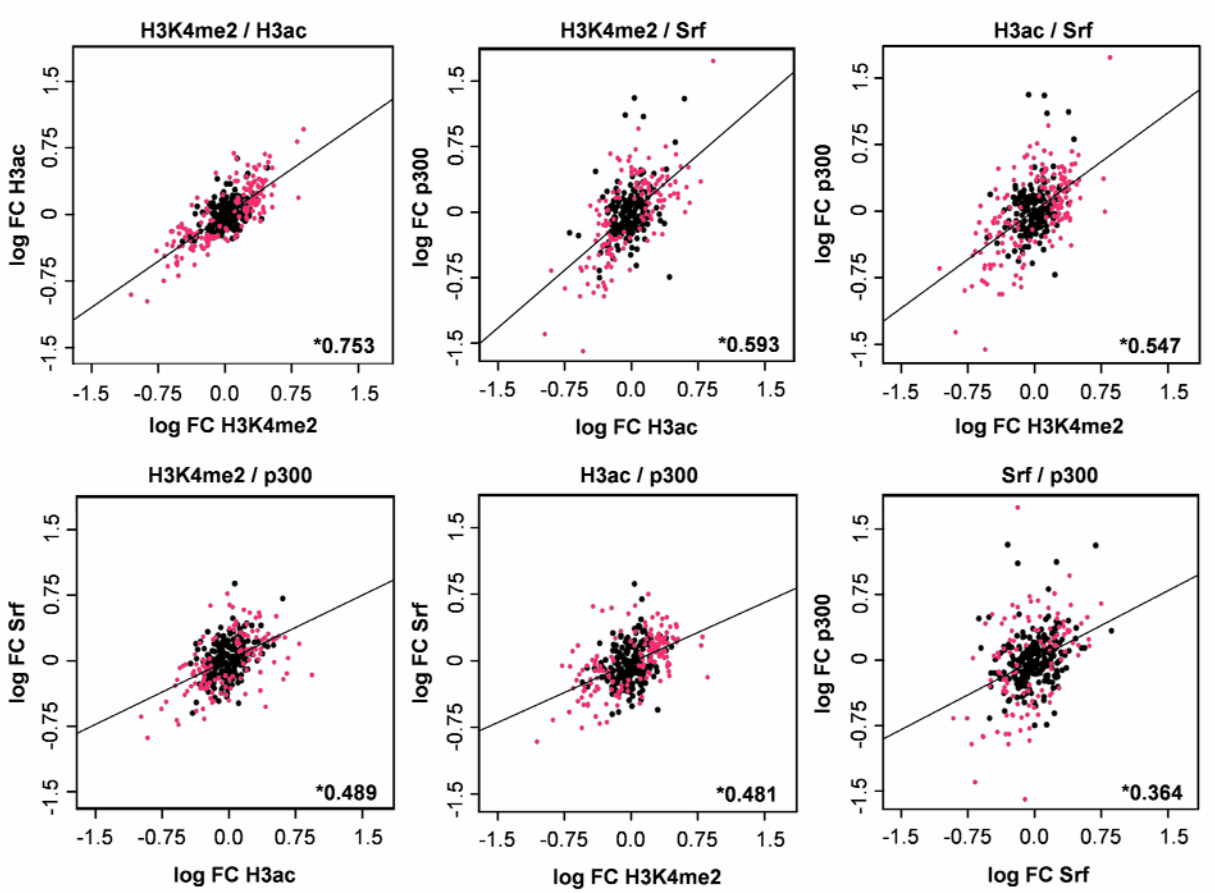


Fig. 4: Scatter plots of fold changes between the measured factors (combined time points). Red dots represent measurements with significant change for at least one of the factors. Lines represent the best linear fit for all measurements. Pearson correlation coefficients for all measurements are indicated in the lower right corner. Significance using an empirical p-value < 0.001 is indicated by *.

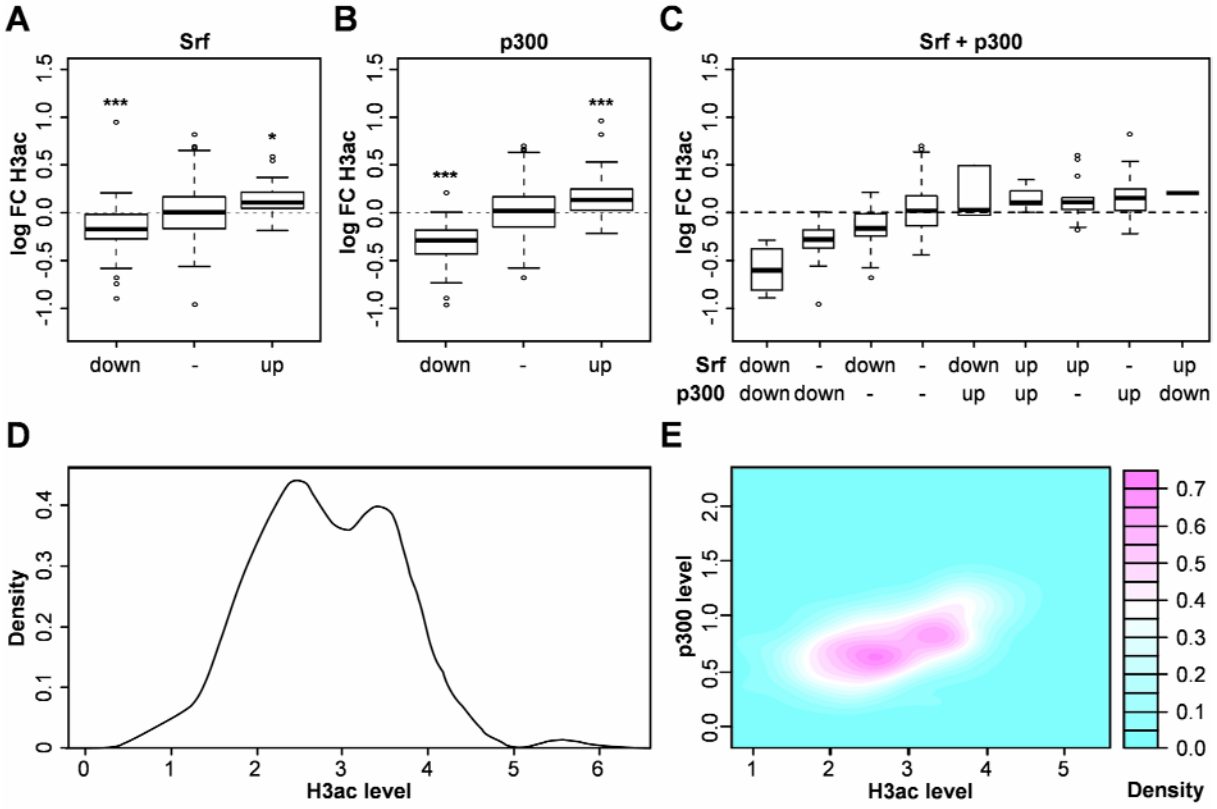


Fig. 5: (A - C) ANOVA models predicting changes in H3ac (combined time points). Boxplots illustrating the dependence of changes in H3ac from grouped significant changes in (A) Srf, (B) p300 and (C) Srf and p300. Significance levels indicating difference from zero according to the single-factor ANOVA models are given above each box, using the following coding: *:p<0.05, ***:p<0.001 Estimated coefficients and respective p-values for the two-factor ANOVA model can be found in Tab. 1. (D) Density of H3ac enrichment levels of region without a significant change in both Srf and p300 using a bandwidth of 0.253 and n=227 observations. (E) Two-dimensional density for H3ac and p300 enrichment levels. The height is indicated using the color key on the right.

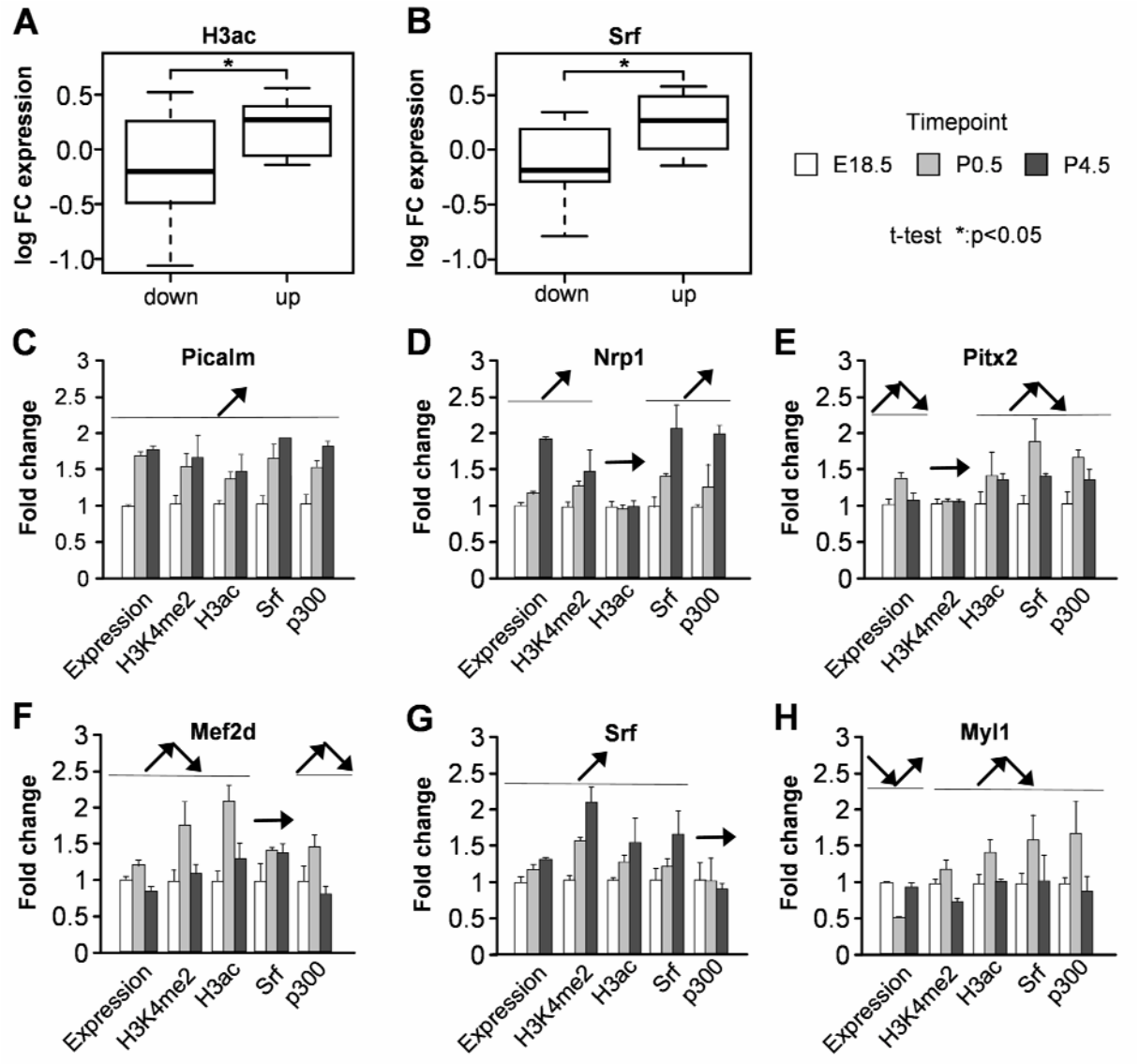


Fig. 6: (A + B) Gene expression changes grouped by significant up or down regulation of H3ac and Srf. Boxplots illustrating the dependence of changes in expression level from significant changes in (A) H3ac and (B) Srf. (C - H) Representative examples of variable gene regulation. (C) All factors and histone modifications correlate with gene expression. (D – G) All factors and histone modifications correlate with gene expression with the exception of (D) H3ac, (E) H3K4me2, (F) Srf and (G) p300. (H) Anti-correlation between Srf, p300, histone modifications and gene expression.

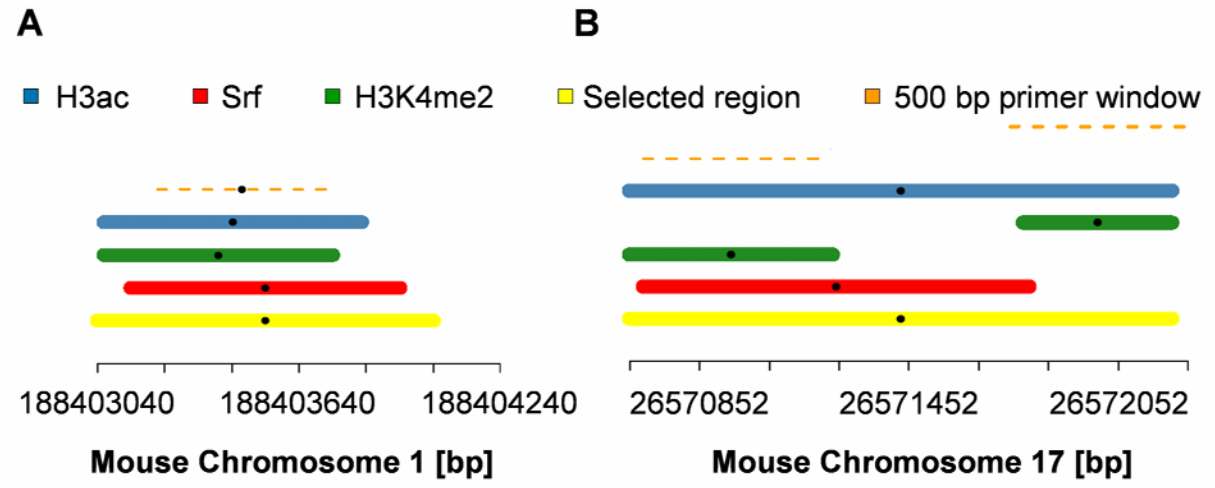
Tab. 1: Estimates and p-values of two-factor ANOVA. The p-values are based on *F* statistic and reflect the significance of the estimates difference from zero. All estimates are in log space.

	unchanged	**p300 up**	**p300 down**	**Srf up**	**Srf down**
Coefficient	0.03483	0.12557	-0.32705	0.06896	-0.18069
p-value	2.23×10^{-2}	1.9×10^{-3}	4.56×10^{-14}	1.84×10^{-1}	1.42×10^{-4}

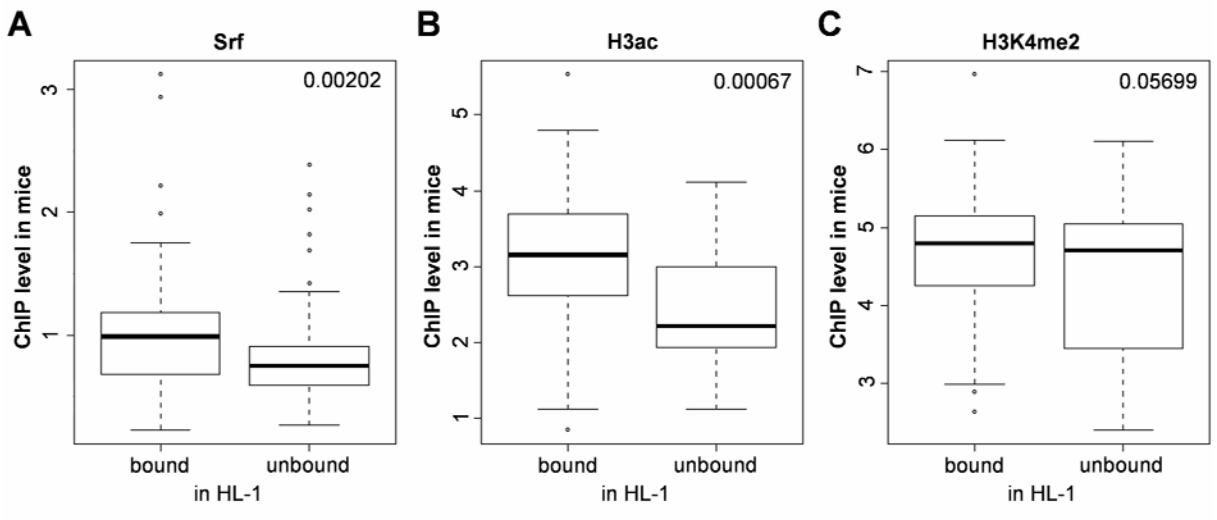
Tab. 2: Estimates and p-values of linear regression model. The p-values are based on *F* statistic and reflect the significance of the estimates difference from zero. All estimates are in log space.

	Intercept	**Srf**	**p300**
Coefficient	0.00622	0.31083	0.32681
p-value	5.68×10^{-1}	5.25×10^{-12}	1.12×10^{-23}

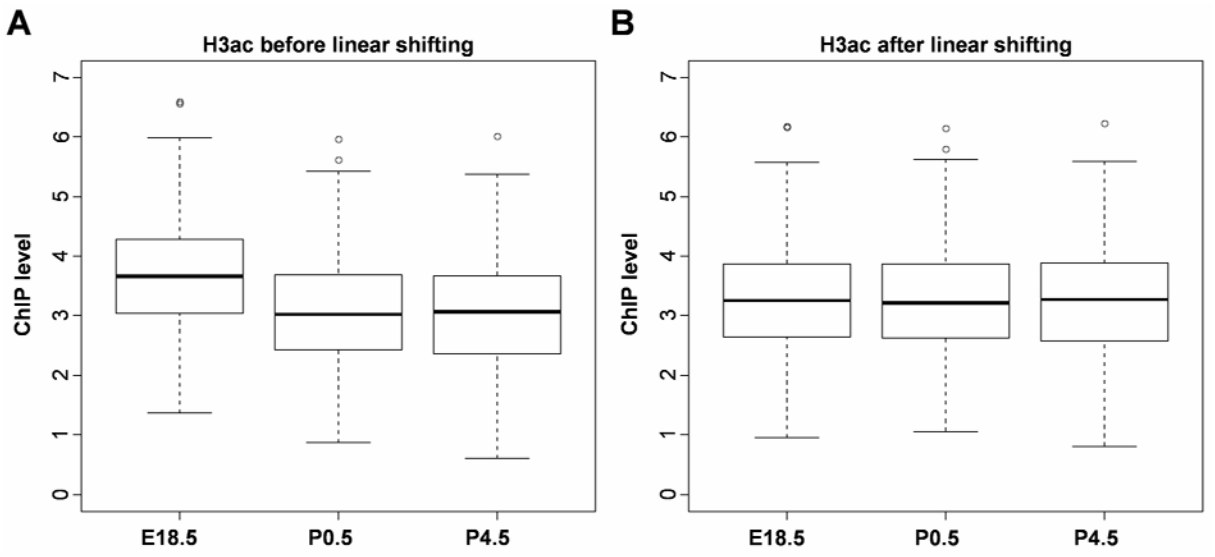
Supplemental information



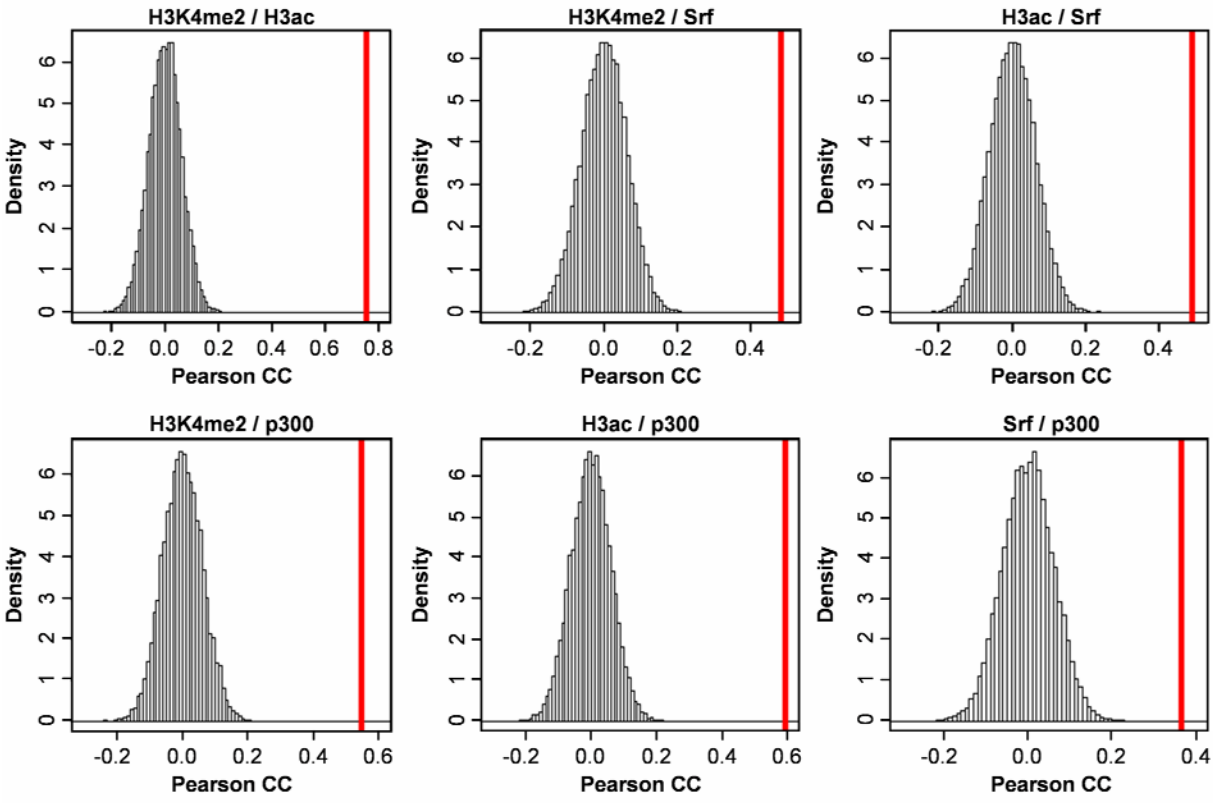
Supplemental Fig. S1: Two examples of selected regions based on ChIP peaks gathered in HL-1 cells. The selected region (yellow) spans all overlapping individual ChIP peaks (blue, green and red for H3ac, H3K4me2 and Srf, respectively). Fixed windows of 500bp length (orange dashed line) were positioned in the middle of the ChIP peaks for primer design. (A) A single 500bp primer window was associated with this selected region on chromosome 1. (B) Two primer windows were used to span all interesting ChIP peaks associated to this selected region on chromosome 17.



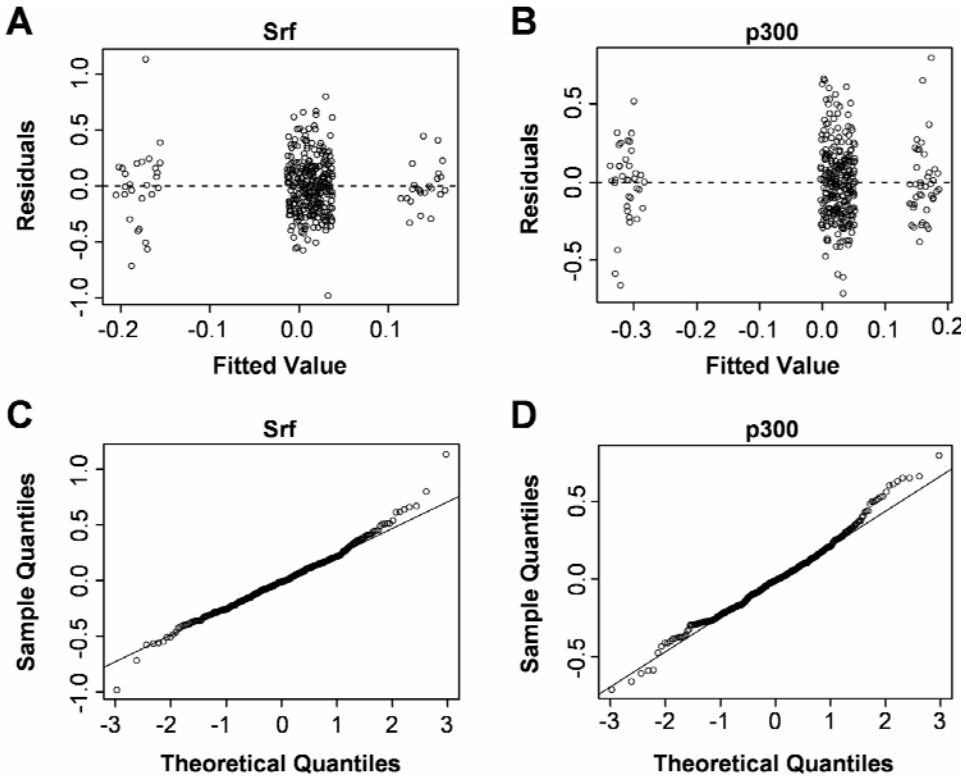
Supplemental Fig. S2: Boxplots comparing ChIP-qPCR enrichment in mouse hearts with regions showing enrichment of Srf, H3ac or H3K4me2 using ChIP-chip/seq in HL-1 cells. For each factor the regions that are bound in HL-1 cells also have a higher average enrichment in our analysis. T-test p-values for the difference in mean are indicated in the upper right corner.



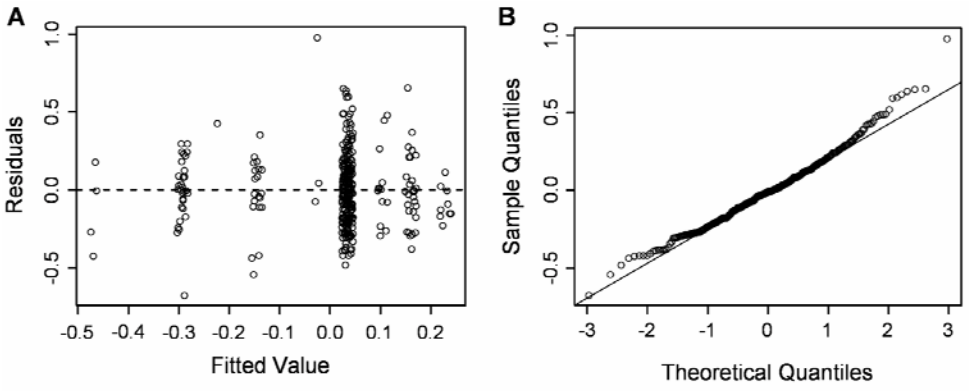
Supplemental Fig. S3: Boxplots of average enrichment over all regions for H3ac in every stage. (A) Measurements after ΔCP normalization showing a distinct trend between the individual stages. (B) Measurements after additional linear shifting which removes the trends.



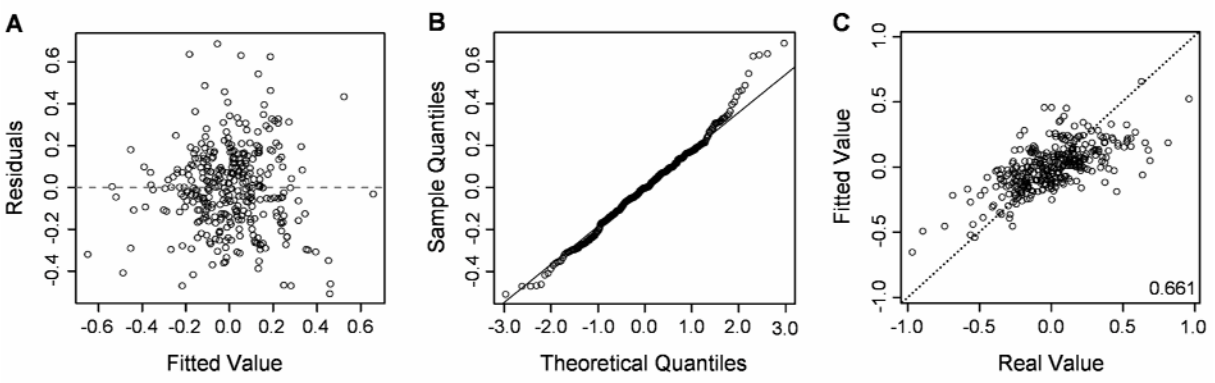
Supplemental Fig. S4: Distribution of Pearson correlation coefficients resulting from the random experiments. The correlation coefficients observed in the real data are indicated by a red vertical line. Data shown for combined time points.



Supplemental Fig. S5: Quality check of single-factor ANOVAs (combined time points). Residuals against fitted values for (A) Srf and (B) p300. Q-Q normal plot for (C) Srf and (D) p300.



Supplemental Fig. S6: Quality check of two-factor ANOVAs (combined time points). (A) Residuals against fitted values. (B) Q-Q normal plots.



Supplemental Fig. S7: Quality check of quantitative linear regression model (combined time points). (A) Residuals against fitted values for the linear model. (B) Q-Q normal plot for the linear model. (C) Model fit against measured values for the linear model. Pearson correlation coefficients are indicated in the lower right corner.

Supplemental Tab. 1: Antibodies used in ChIP and Western Blot (WB) and their respective amount used in the experiments.

Primary Antibodies	**Company**	**Used Amount**
Anti-Srf, rabbit polyclonal	Santa Cruz Biotech (sc-335X)	WB: 1:500, ChIP: 10µg
Anti-p300, rabbit polyclonal	Santa Cruz Biotech (sc-585x)	WB: 1:200, ChIP: 10µg
Anti-H3, rabbit polyclonal	Abcam (ab1791)	WB: 1:1000
Anti-H3ac (H3K9K14ac), rabbit polyclonal	Upsate (06-599)	WB: 1:500, ChIP: 5µg
Anti-H3K4me2, rabbit monoclonal	Abcam (ab32356)	WB: 1:250, ChIP: 10µl
Anti-GAPDH, mouse monoclonal	Ambion (AM4300)	WB: 1:5000
Secondary Antibodies		
Anti-mouse IgG conjugated with HRP	Sigma (A0168)	WB: 1:10000
Anti-rabbit IgG conjugated with HRP	Sigma (A2074)	WB: 1:10000

5 MANUSCRIPT 3

Evaluation of the LightCycler® 1536 Instrument for high-throughput quantitative real-time PCR

Jenny Schlesinger, Martje Tönjes, Markus Schueler, Qin Zhang, Ilona Dunkel and Silke R. Sperling. Methods. 2010 Apr;50(4):S19-22.
doi:10.1016/j.ymeth.2010.01.007
The original article is online available at
http://www.sciencedirect.com/science/article/pii/S1046202310000228

5.1 Synopsis

In context of a beta-side testing of the novel LightCycler® 1536 Real-Time PCR System, the aim of this project was the evaluation of this new system for its applicability as medium and high-throughput quantitative real-time PCR technology.

The quantitative real-time PCR is the widely used “gold standard” method in research and diagnostics to quantify nucleic acids, because of its high sensitivity, good reproducibility, broad dynamic quantification range, easy use and reasonable good value for money.[205-206] In general, it is routinely used for the verification of sequencing and microarray data. So far, the throughput was limited to 384 reactions per run, but now in the course of the rapid development of new technologies, the recently launched LightCycler® 1536 Real-Time PCR System offers the analysis of 1536 reactions per run with an approximate duration of 45 min.

In order to evaluate whether an up scaling from 348-well to 1536-well still provide reliable data; we determined mRNA expression levels of cardiac genes in human cardiac tissue samples using the new system with hydrolyses probe assays (TaqMan assays). We compared these data with already existing expression data obtained in an advanced gene expression study performed in our group.[47] The same cardiac genes in the same tissue samples of patients with heart malformations were measured in 384-well format using SYBR Green and the ABI PRISM 7900HT system.[47]

The workflow evaluating the LightCycler® 1536 Real-Time PCR System comprises the following steps: We designed for a comprehensive set of 50 cardiac genes the respective hydrolyses probe assays, which were subsequently tested for their efficiency using a 348-well format qPCR system. In the next step, the pipetting robot as well as the 1536 qPCR machine were set up and pre-tested to guarantee a proper performance. Subsequently, for a total of 42 cardiac genes the expression levels in 71 cardiac tissue samples were measured in triplicates on 1536-well plates. Finally, the results were analysed and the obtained expression profiles were compared with our previous data. The expression levels were found to be almost identical with both systems indicating that the data are reliable and robust. Therefore, we utilized the LightCycler® 1536 system for the generation of extensive qPCR data as presented in chapter 3 and 4 of this thesis. Chromatin immunoprecipitation (ChIP) followed by qPCR in a time-series of mouse hearts of three developmental stages was conducted to analyze the dynamic binding behavior of Srf and the histone acetyltransferase p300 in correlation to their occurrence with H3ac and H3K4me2.

To summarize, the novel LightCycler® 1536 instrument is a very applicable qPCR system for medium and high-throughput gene expression analysis. It allows to produce extensive qPCR data in a shorter time period and with reduced amounts of reagent and sample material compared to older 348-format systems.

5.2 Experimental contribution

For this work I designed and carried out all wet lab experiments except for sample preparation, RNA isolation and cDNA synthesis. In addition, I wrote the manuscript and included suggestions from co-authors.

Contributions of co-authors:
Conception: SR. Sperling
Performed the experiments: M. Tönjes, Q. Zhang, I. Dunkel
Bioinformatic analysis: M. Schueler

6 MANUSCRIPT 4

Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex

Lange M, Kaynak B, Forster UB, Tönjes M, Fischer JJ, Grimm C, **Schlesinger J**, Just S, Dunkel I, Krueger T, Mebus S, Lehrach H, Lurz R, Gobom J, Rottbauer W, Abdelilah-Seyfried S, Sperling S. Genes Dev. 2008 Sep 1;22(17):2370-84.
doi 10.1101/gad.471408
The original article is online available at
http://genesdev.cshlp.org/content/22/17/2370.long

6.1 Synopsis

The aim of this project was the detailed characterisation of DPF3, an epigenetic transcription factor important for cardiac and muscle development.

DPF3 (*D4, zinc and double PHD fingers, family 3*) is one of three members of the d4 gene family in mammalian.[207] These proteins are characterized by an N-terminal 2/3 domain, which is unique to this protein family and of currently unknown function. Furthermore they contain a C2H2-Krüppel-like zinc finger, which is generally described as DNA, -RNA and protein interaction domain, and they are marked by a C-terminal PHD finger which was previously called the d4 domain.[208] *DPF3* gives rise to two splice variants in human and mouse, which differ at their C-terminus. The isoform *DPF3a* contains only a single truncated PHD finger whereas the isoform *DPF3b* features a double PHD finger.

In previous genome-wide gene expression studies of congenital malformed hearts performed in our group, both splice variants of *DPF3* were found to be significantly up-regulated in human right ventricular tissue of patients with Tetralogy of Fallot compared with healthy controls.[47,209]

In our present study, we analysed the expression pattern of DPF3 during embryogenesis in mouse, chicken and zebrafish. In situ-hybridization revealed predominant expression in the developing heart and somites. Appropriately, Northern Blot analysis of multiple human tissues displayed specific expression of *DPF3* in cardiac and skeletal muscle.

To explore the role of DPF3 *in vivo*, we performed knockdown of *dpf3* in zebrafish using the Morpholino antisense technique. *Dpf3* morphants exhibit abnormal heart morphology, incomplete cardiac looping, poorly defined atrioventricular boundary, reduced cardiac contractility while the heart beat rate and the atrial and ventricular specification were widely normal. Moreover, we observed disturbed myofibril organisation in cardiac and skeletal muscle fibers caused by transcriptional deregulation of structural and regulatory proteins. Additionally, Dpf3 knockdown in murine C2C12 skeletal muscle cells using RNA interference caused disassembled muscular fibres.

PHD finger proteins are normally found in nuclear protein complexes involved in chromatin remodeling. Thus, we used tandem affinity purification technique with subsequent mass spectrometry analysis to isolate nuclear binding partners of DPF3a and DPF3b in HEK293T cells. Nearly all core components of the human BAF chromatin remodeling complex could be identified as potential interaction partners. Among these was BAF60c – a heart and somite specific subunit of the complex. To confirm the association of DPF3 with the BAF complex, we performed reverse tandem affinity purification with subsequent mass spectrometry analysis using BAF60c as bait.

Furthermore, we used the GST pull-down system to elucidate whether DPF3 recognizes histones and specific histone modifications as described for several PHD finger proteins. We could show that DPF3b contains the first PHD finger which binds acetylated lysines on histone 3 and 4 besides binding to mono- and dimethylated lysine 4 on histone 3. The binding of acetylated lysine residues is a feature previously only shown for bromodomains.[210] The truncated PHD finger of DPF3a is not capable to bind any histones.

Using ChIP we showed on a global scale that DPF3 binds distinct chromatin sites *in vivo*, which are essential for muscle development and function, and marked by acetylated and methylated histone tails. Additionally, we performed ChIP for BRG1, a core component of the BAF complex, revealing a high degree of overlapping binding sites for BRG1 and DPF3 suggesting that DPF3 potentially serves as an anchor between the BAF complex and modified histones.

Mef2a-deficient mice and zebrafish embryos are phenotypically similar to the observed myofibrillar disarray in *dpf3* knockdown embryos.[61,211] Consequently we performed promoter analysis using ChIP, siRNA knockdown and luciferase reporter gene assays and could identify Mef2a as transcriptional regulator of *Dpf3*.

To summarise, the cardiac and muscle specific transcription factor Dpf3 was found to interact with the human BAF chromatin remodeling complex. Moreover, Dpf3 contains the first double PHD finger binding methylated and acetylated lysine residues of histones and is essential for heart and skeletal muscle development. Thus, DPF3 might serve as a tissue specific anchor recruiting the BAF complex to chromatin target sites.

6.2 Experimental contribution

For this study I carried out the Brg1 ChIP experiment with subsequent real-time PCR analysis and I was involved to minor parts in performing the reverse tandem affinity purification with Baf60c, siRNA experiments, RNA isolation and real-time PCR measurements.

Contributions of co-authors:
Conception: SR. Sperling
Performed the experiments: M. Lange, B. Kaynak, UB. Forster, M. Tönjes, JJ. Fischer, C. Grimm, S. Just, I. Dunkel, S. Mebus, W. Rottbauer, S. Abdelilah-Seyfried
Bioinformatic analysis: T. Krueger
Mass spectrometry: J. Gobon
Electron microscopy: R. Lurz
Wrote the paper: M. Lange, SR. Sperling

7 DISCUSSION

The molecular mechanisms by which the genomic information mediates the synthesis of different cardiac relevant genes have been the focus of many studies over the last decades. It is well established that an evolutionarily conserved network of transcription factors directs cardiac development and function.[37] Histone modifications have emerged as prominent regulators of cardiac development and function as they control the accessibility of transcription factors to DNA and provide scaffolds for protein-protein interactions, thus adding a further level of regulating gene expression.[212] Just recently, studies revealed the high impact of miRNAs on regulating cardiac mRNA profiles via either translational repression or mRNA degradation.[166] However, most studies have focused on a single regulatory level independently and we still lack data showing the interaction between these three levels of regulation. This work presents for the first time a combinatorial analysis of these molecular mechanisms in respect to their regulation of cardiac transcription profiles. In addition, the epigenetic factor DPF3 was characterized in more detail as it couples the overall chromatin remodeling machinery to cardiac genes.

We initially focused on the four transcription factors Gata4, Mef2a, Nkx2.5 and Srf, which all belong to the evolutionarily conserved cardiac transcription factor network and are essential for heart development and function.[37,58,61,80,192] Previously, only few target genes have been identified and no study existed that analyzed these four factors together in one cell type. Using ChIP-chip we determined specific binding sites in the cardiomyocyte cell line HL-1. We found the positions of their binding sites to be symmetrically distributed around the transcriptional start site which might be a common eukaryotic feature for regulatory elements as suggested e.g. by Birney et al.[14] Beside already known targets, we identified hundreds of novel binding sites that mainly correspond to target genes involved in heart development and function what substantiated the reliability of out data. For example, the Gene Ontology (GO) term “heart looping” was significantly overrepresented among Nkx2.5 targets, a process, which is affected in Nkx2.5 knockout mice during embryonic development.[80] Interestingly, Nkx2.5 target genes were also found to be associated with “biomineral formation” and “bone mineralization” possibly indicating novel functions of Nkx2.5. A significant proportion of Srf target genes are involved in the regulation of heart contraction and embryonic heart tube development, and these processes are found to be disturbed in Srf mutant mice.[75] Srf is also a master regulator of the actin cytoskeleton[66-67] and a panel of its target genes are associated with actin cytoskeleton organization and biogenesis. Its transcriptional activity and its cell-type specificity are largely dependent on its interaction with other positive and

negative co-factors such as Myocardin as well as the two other studied factors Nkx2.5 and Gata4.[65,68] The physical association of Srf with Gata4 and Nkx2.5, in particular, was found to enhance transcription of cardiac genes.[73] Moreover, Srf itself regulates the expression of Nkx2.5 and Gata4 as shown by us and others.[75] A study by Boogerd et al. demonstrated that the physical interaction of Nkx2.5 and Gata4 results in a synergistically activation of atrial natriuretic factor gene expression.[213] The homeodomain of Nkx2.5 and the carboxy-terminal zinc finger of Gata4 were shown to be required for this interaction. Further, the physical interaction between the MADS box of SRF and the second zinc finger domain of GATA4 mediates activation of SRF-dependent gene expression[53] and physical interaction between the DNA-binding domain of MEF2 and the carboxy-terminal zinc finger of GATA4 has been described to increase the transcriptional activity of GATA4.[57] Although the studied transcription factors were known to physically interact and regulate each other's expression, only little was known about the relevance and the frequency of this *in vivo*. We show that a high number of genes are bound by two or more of these transcription factors indicating a complex and cooperative regulation of target gene expression. We further suggest that such synergistic interactions among tissue-specific transcription factors are an important mechanism to regulate developmental pathways.[213] As the comparison of observed to expected number of pairwise bound genes revealed a lower than expected number of Srf sites co-occupied by the other studied factors, we confirm a general role for Srf in the regulation of many cellular processes likely via its interaction with other co-factors than those measured.[65,68,214-215]

To explore the functional consequence of transcription factor binding, we performed RNAi knockdown for each factor and measured resultant mRNA profiles with genome-wide expression arrays. In accordance with a primarily activating function of all transcription factors, the majority of deregulated transcripts were down-regulated. However, we also found up-regulated genes. The role a transcription factor plays in a certain genomic content thereby likely depends on its interaction with co-factors. For example, depending on the availability of co-factors and signals from the cell, SRF can regulate its targets in an opposing manner.[216] In serum-stimulated cells, SRF activates important genes for cell growth such as c-fos, whereas in muscle cells it activates muscle genes that are again repressed in proliferating cells.[217-218] We made a similar observation in the Srf knockdown experiments. Structural muscle genes such as α-actin-1 (Actc1), α-actin-2 (Acta2), or myosin-heavy chain 7 (Myh7) were found to be significantly down-regulated whereas the insulin growth factor 1 (Igf1) was up-regulated. Similar to the analysis of common binding behavior, we studied deregulated genes and found that the different factors have a high proportion of common differentially expressed genes. Thus, only a small fraction of directly bound genes (~10%) was concomitantly differentially expressed, and vice versa, only a small number of differentially

expressed genes were directly bound pointing to a high proportion of indirect regulations. Further, the assumption that each transcription factor-binding event results in expression or repression of bound genes is misleading. Our observations are in line with other studies showing that the majority of genes bound by a particular transcription factor is not affected when the factor is knocked down.[193-195,219] There are several explanations for the observed small overlap. The most feasible are non-functional binding,[196,220] dosage dependent regulation by transcription factors[86-88] or redundant paralogs which can compensate the loss of one factor.[60,193,221-222] For instance, members of the Gata family were shown to be functionally redundant for the specification of cardiomyocytes in zebrafish.[221] Another study investigating functional consequences of Mef2c depletion in mice proposed that the redundant family members can compensate the essential role of Mef2c in later developmental stages of the heart.[222] A study in yeast analysing *cis*-regulatory elements demonstrated that slight differences in the effective transcription factor concentration could cause a switch between repressed and active gene transcription, underlining the importance of dosage dependent regulation.[223] In addition, the small overlap between transcription factor binding and deregulation of target genes adds to the hypothesis of combinatorial regulation of gene expression mediated by Gata4, Mef2a, Nkx2.5 and Srf and to potential buffering effects between these non-paralogs. This hypothesis was substantiated by further bioinformatical analyses revealing that genes, which were bound by multiple transcription factors, were also significantly less likely differentially expressed. Likewise, transcription factors, which had a high number of common binding sites, share only a small number of co-regulated genes upon RNAi-mediated knockdown.

In addition to the combinatorial regulation by transcription factors, it is well known that transcription factor binding is strongly dependent on accessibility of their binding sites often induced by the presence of specific histone modification patterns. The influences of these histone modifications during heart development and function has been addressed by several studies, but only little is known about their impact on the transcriptional activity of Gata4, Mef2a, Nkx2.5, and Srf. Hence, we aimed to elucidate to which extent H3ac, H4ac, H3K4me2, and H3K4me3 influence the gene expression of respective target genes. In accordance with other studies showing that transcriptionally active DNA regions are marked by these histone modifications,[224-226] we observed that around 80% of binding sites carry one or more of these modifications. Interestingly, we found that Gata4 and Srf bound targets were only significantly higher expressed when they were additionally marked by H3ac, while the expression levels of Nkx2.5 and Mef2a targets were significantly higher than the reference group, independent of whether H3ac was present or not. The correlation of H3ac and Srf target gene expression was further investigated and reconfirmed in a genome-wide manner using ChIP-seq technique. We could show that the decrease of target gene

expression upon Srf knockdown is significantly reduced for genes, which are additionally marked by H3ac. We conclude that H3ac can potentially buffer the loss of Srf. Further, using ChIP followed by qPCR after Srf knockdown, we found a significant decrease of H3ac marks at target genes, which are under normal conditions marked by both. In conclusion, our data provide evidence that Srf triggers the acetylation at particular target genes and moreover, that H3ac supports the activating function of Srf. This is in line with two other studies reporting that Srf directs tissue-specific transcription by recruiting co-activators containing HAT activity such as p300 and CREB-binding protein (CBP) leading to an increased histone acetylation.[198,200] The HAT p300 which acetylates lysine residues on histone 3,[90] can bind both SRF and Gata family members.[227-229] Moreover, p300 has been shown to acetylate lysine residues of Gata4 resulting in enhanced DNA binding affinity and an activation of Gata4-dependant transcription.[54] Thus, acetylation of H3 mediated via p300 provides an explanation for the observed high target gene expression of Gata4 and Srf. The observed correlation between Srf and H3ac gathered in cell culture were subsequently investigated in a time-series with mouse hearts of three developmental stages around birth. We studied the dynamic binding behavior of Srf and p300 in correlation to their occurrence with H3ac and H3K4me2 at heart and muscle relevant genes integrating expression levels of selected target genes. In accordance with other studies reporting that methylation of H3 recruits HAT complexes including p300 to induce histone acetylation at active chromatin regions,[94,202-204] we found the highest correlations for changes in H3K4me2/H3ac, H3ac/p300 and H3K4me2/p300. Studies mapping histone patterns show that H3K4me2 marks strongly correlate with H3ac at TSS of active genes.[230-232] Similarly, retinoid acid mediated Hox gene expression is activated in conjunction with high levels of H3ac and H3K4me2 and recruitment of p300.[233] It is suggestive that our observed correlation is achieved by the tissue-specific transcription factor Srf. Activation of smooth muscle cell (SMC) genes is mediated through binding of the Srf/Myocardin complex at GArG boxes marked with H3K4me2 and H4ac thereby recruiting other regulators like p300 what leads to further histone acetylation events including H3ac.[114] In line with this, we observed significantly correlation between changes in Srf/H3ac and Srf/H3K4me2. However, our findings gathered in cardiomyocytes slightly differ from results from SMCs. First, we found H3ac instead of H4ac to be the causative link between Srf and transcriptional activation of target genes. Second, we found only a low correlation between Srf and p300. As this correlation was lower than between Srf and H3ac, it indicates that an additional mechanism for Srf-triggered acetylation might exist probably independent of p300 recruitment. The histone acetyltransferase CBP (CREB-binding protein) sharing regions of very high-sequence similarity with p300 might be a potent candidate.[234] The analysis of the influence of Srf, p300 and the activating histone modifications on gene expression levels revealed a highly dynamic and cooperative regulation. In general, gene

expression correlated with the enrichment of the studied factors; thus, many genes were found that seem to be regulated by only a subgroup of these indicating a high variability of combinatorial regulation. The observed connection between Srf and histone modifications is further underlined by the fact that we found histone modifying enzymes such as histone demethylases (Jmjd1c, Jmjd2b, Jmjd3, Jmjd4 and Jmjd5) as important group of direct Srf targets.

In context of cardiac factors that contribute to chromatin environment, we further analyzed the PHD finger containing protein DPF3 that was found to be a regulator for heart development and up-regulated in patients with Tetralogy of Fallot (TOF).[209] Our study revealed DPF3 as a component of the BAF chromatin remodeling complex being in accordance with the fact that PHD finger containing proteins are often found in nuclear protein complexes involved in chromatin remodeling.[142] Protein pulldown experiments carried out in our group have shown that the interaction of DPF3 with the BAF complex is probably mediated via interaction with BRG1 and BAF60c (unpublished data). The responsible interaction domain resides within the N-terminal region, which is common for both splice variants and contains the 2/3 domain, a nuclear receptor interaction domain and the C2H2-Krüppel-like zinc finger. This is in line with a study investigating DPF2, the ubiquitously expressed member of the d4 family.[235] DPF2 was shown to function as adaptor protein between the SWI/SNF complex and RelB/p52 via its N-terminal domain, thereby being important for the transcriptional activation of the noncanonical NF-kappaB pathway. As Dpf3 is specifically expressed within the developing heart and somites and partially within the brain, we speculate that DPF3 serves as tissue-specific subunit of the BAF complex that controls the transition of cardiac precursor to differentiating cardiomyocytes. A similar example for tissue-specificity has been shown for DPF3 (also known as BAF45c) and the subunit DPF1 (also known as Baf45b), which are incorporated into the BAF complex during transition of proliferating neuronal precursor to differentiating neuronal cells in the developing mouse brain.[125] PHD finger domains are known to function as readers of histone modifications.[136] Indeed, we could show that the isoform DPF3b interacts with mono- and di-methylated lysine residues of histone 4. Moreover, we could show that DPF3b contains the first tandem PHD finger that also interacts with acetylated lysine residues of histone 3 and 4. So far, the recognition of acetylated histone marks was classified as a property only for bromodomains.[210] The investigation of the co-occurrence of DPF3b binding sites with those of BRG1 revealed highly overlapping bindings sites at muscle relevant genes that contain acetylated and/or methylated H3 and H4. Consequently, we favour the view that DPF3 serves as tissue-specific anchor between modified histone tails and the BAF complex. The three-dimensional structure of histone binding by the double PHD finger of DPF3b was just recently solved using nuclear magnetic resonance (NMR) methods.[236] The tandem PHD

finger of DPF3b was found to act as one functional unit. Each PHD finger consists of a two-strand antiparallel β-sheet followed by a carboxy-terminal α-helix that are stabilized by two zinc atoms coordinated by the Cys-4-His-Cys-3 motif similar to other known PHD structures.[136] It is proposed that the first PHD finger binds H3K14ac, whereas this interaction is subsequently inhibited by H3K4me3 recognized by the second PHD finger.[236] These opposing effects then promote the transcriptional activation of DPF3b target genes Pitx2 and Jmjd1.

As another regulatory level, we investigated the regulation by miRNAs. Regarding the modest proportion of differentially expressed target genes directly bound by Gata4, Mef2a, Nkx2.5 and Srf, these are good candidates to explain the high degree of indirect regulation. Indeed, focusing on Srf, which is known to regulate cardiac relevant miRNAs, we identified a panel of miRNAs being deregulated upon Srf knockdown. Subsequent target prediction of these deregulated miRNAs could explain three times more differentially expressed genes than Srf binding alone. The majority of deregulated miRNAs was found to be down-regulated supporting the role of Srf as a miRNA activator. Given the fact that miRNAs have widely been shown to down-regulate the levels of their target mRNAs,[237-238] we as well expected that the corresponding miRNA targets were negatively correlated in their expression. Surprisingly, although the majority of deregulated transcripts with predicted miRNA target sites were up-regulated thus being negatively correlated, a considerable amount was found to be positively correlated. A possible explanation would be both miRNAs and their targets are co-regulated by upstream factors, as suggested by several studies.[239-240] Furthermore, intronic miRNAs and corresponding mRNAs targets that are transcribed from the same genes tend to be positively correlated as shown by Wang et al.[241] In addition, effects on transcript abundance is often modest due to the repressive effect of miRNAs on the level of translation.[144] The influence on protein level, which is the final readout of the regulatory function of miRNAs, could not been examined in this study. Approaches that measure large-scale protein abundance using for example quantitative mass spectrometry are still rare and most of them focus on individual miRNAs and their corresponding protein output.[242-245] Computational target prediction algorithms are frequently used to identify potential targets. Most of these are based on sequence complementarities between the mature miRNA and the 3'UTR target site, binding energy of the miRNA-target duplex, and evolutionary conservation of the complementary sequence in the 3'UTR and target position in aligned UTRs of homologous genes.[246] Thus, comparison between validated and predicted targets still reveals many false-positive and false-negative predictions.[247] Therefore, the experimental validation of direct miRNA targets using for instance reporter gene systems is crucial.[248] Hence, further investigations are necessary to understand the essential impact of miRNAs as co- regulators of transcriptional profiles.

The presented study integrated a high number of high-throughput datasets, which needed to be validated to substantiate the reliability of produced data. In general, qPCR was routinely applied in this thesis, as it is the method of choice for rapid quantification of nucleic acids on a single genes basis due to its high sensitivity, good reproducibility, broad dynamic quantification range, easy use and reasonable good value for money.[205-206] In context of a beta-side testing, we evaluated the novel LightCycler® 1536 real-time PCR System as very applicable qPCR system for medium-throughput analyses, which makes it a valuable tool for high-throughput data verification.

A third epigenetic mechanism, which was not investigated in this thesis but which also affects chromatin structure and thus controls gene expression, is the methylation of DNA cytosine nucleotides.[249] In mammals, the 5'carbon of cytosine located next to a guanine is converted into 5-methylcytosine by DNA methyltransferases.[250] In particular, CG-rich regions present in satellite repeats, ribosomal DNAs, centromeric repeats and CpG islands are targets for DNA methylation.[251] In mammals, approximately 70% of those CpG-rich sequences are methylated and tightly associated with heterochromatin formation.[252-253] In line with this, a large number of promoter regions with CpG islands are unmethylated under normal conditions and an aberrant or increased methylation can result in pathogenetically relevant gene silencing.[254] In turn, DNA methylation of X-chromosomal genes is crucial as it is correlated with an inactive X-chromosome in female mammals.[255] Promoter methylation often shows correlation with a repressed chromatin state as it directly interferes with binding of transcriptional regulators and recruits methylcytosine binding proteins that cause a repressed chromatin environment. A first link between aberrant DNA-methylation and human heart disease was demonstrated by a study investigating genome-wide DNA-methylation level in patients with cardiomyopathy compared to healthy individuals.[256] Using a ChIP approach for methylated-DNA (MeDIP-chip), they identified changes in DNA methylation at specific genomic loci leading to changes in expression of corresponding genes. However, it still remains an open question whether changes in DNA-methylation are the cause or the consequence in the progression of heart failure.[256] Likewise, it is still unknown whether the presence of histone modifications facilitates binding of TFs and recruits other co-factors or whether an initial TF binding causes changes in histone modifications and the chromatin structure. It is highly suggestible that both are true and that both are highly intermingled. In line with this, findings investigating the function of BRG1 during differentiation of erythrocytes revealed that the transcription factor GATA1 recruits BRG1 to its enhancer elements resulting in a BRG1-dependant shift of nucleosomes, which facilitates the binding of the second transcription factor TAL1 leading to transcriptional activation of TAL1 target genes.[257] We have shown that the loss of Srf decreases H3ac at particular target genes pointing to an initial Srf binding with H3ac as consequence. Conversely, our data indicate that H3ac

potentially preserve Srf target gene expression after its knockdown. For the myocardin-induced activation of smooth muscle cell gene transcription, it was suggested that Srf binds weakly to the promoters of smooth muscle-specific genes.[258] Consequently, Srf interacts with myocardin and forms a complex with HATs and the SWI/SNF chromatin remodeling complex leading to a rearrangement of the nucleosomes that facilitates tight binding of Srf. The complex of Srf/myocardin/HATs can then further modify chromatin and facilitates the recruitment of the general transcripional machinery.[258] The direct measurement of the accessibility of the DNA using for example desoxyribonuclease (DNase) digestion followed by sequencing (DNase-seq) coupled with Srf and H3ac ChIP data could help to answer how the precise epigenetic regulation is established. In respect to the epigenetic factor DPF3, it is further suggestive to couple data of DNA accessibility gained with DNase-seq with its target sites during different developmental stages. So far, we have evidence that DPF3 recruits the BAF complex to heart and muscle relevant genes via its interaction with modified histones. It would be very interesting to elucidate if the recruitment of the BAF complex results in rearrangement of nucleosomes, changes in histone modifications or recruitment of other transcription factors. In the same line, it will be of high interest to figure out the precise mechanism how the combinatorial binding of Gata4, Mef2a, Nkx2.5, and Srf and their buffering effect is mediated. How do the four factors interact and which combinations of interaction are needed to direct correct gene regulation? The reduction or knockdown of several transcription factors at once could help to substantiate the suggested buffering effect between the studied transcription factors. Our data support an important role for the co-occurrence of Srf and H3ac at cardiac and muscle relevant genes in the course of the cardiac maturation process. However, further experiments are needed to elucidate which HATs are responsible for the proposed Srf triggered acetylation. Moreover, including ChIP data for Myocardin, which was to date not possible due to unreliable ChIP grade antibodies, would add to a full understanding of the analyzed regulatory circuits. So far, we investigated a subgroup of histone modifications, but many more exist, which could have an impact on the regulatory function of the studied system. In this study, we used whole hearts of different stages, yet the heart comprises multiple types like cardiomyocytes, endothelial cells, cardiac fibroblasts and vascular smooth muscle cells. Although cardiomyocytes mostly account for myocardial volume, cardiac fibroblasts are the most numerous cell type in the heart and play essential roles in its function.[259-260] Performing cardiomyocytes-specific ChIP time-series using primary cell culture would be of great interest but technically challenging. The usage of HDAC inhibitors like Trichostatin A in combination with ChIP experiments would be helpful to understand the precise interaction between Srf binding and acetylated histone marks. Others and we have demonstrated the importance of histone acetylation in the control of cardiac gene expression. Further studies are necessary to confirm the model in which acetylation

marks recruit DPF3 in conjunction with the BAF complex to target genes, whereas subsequent histone methylation leads to dissociation of DPF3-BAF, thus allowing the transcriptional machinery to initiate and activate gene transcription.[236] One possible mechanism for the transcriptional regulation of DPF3 target sites could be mediated via the recruitment of histone modifying enzymes such as HDACs and HATs. Knockdown of DPF3 followed by ChIP with antibodies against histone modifications and subsequent analysis of selected target genes would be suitable to substantiate this hypothesis. It would be also interesting to investigate if and to which extent the loss of Dpf3 provokes the phenotype also seen in Mef2a deficient mice. The generation of a conditional Dpf3 knockout mouse would be suitable. A mouse model over-expressing DPF3 would be helpful to better understand and probably find an association to the observed phenotype of patients with TOF. In addition, we still lack knowledge about the function of DPF3a, which contains only a truncated PHD finger and is not capable to bind any histones. Still, DPF3a interacts with the BAF complex probably via the interaction with other co-factors or even with DPF3b. Phosphorylation of proteins provides a well-known mechanism that transmits extracellular information to transcriptional regulators thereby changing their binding behaviour. Interestingly, prediction of phosphorylation sites for DPF3 showed potent casein kinase II (CKII) sites located in the DPF3a-specific C-term. CKII phosphorylates a plethora of different targets and participates in a variety of cellular processes.[261] So far, we have preliminary experimental evidence for the post-translational phosphorylation of DPF3a by CKII. These findings could explain the particular role of DPF3a in transcriptional regulation of heart and muscle development.

In summary, the studies presented in this thesis revealed high complexity of cardiac networks, which are regulated at many molecular levels comprising cooperative transcription factor binding, occurrence of activating histone modifications and cardiac relevant miRNAs. Using systematic *in vivo* approaches in a mammalian system, we could demonstrate the combinatorial effects regulating mRNA profiles. Our results indicate a fine-tuned regulation of cardiac gene expression directed by the combinatorial influence of Gata4, Mef2a, Nkx2.5 and Srf, while histone modifications and miRNAs modulate their functional consequence with a high degree of interdependency. Our constructed transcription network centered on Srf can build a basis for further gene focused studies and illustrates the different nodes of regulating cardiac gene expression. We identified the PHD finger containing protein DPF3 as one important epigenetic player in heart and muscle development, which associates with the BAF chromatin remodeling complex and binds histone acetylation and methylation marks. Thus, DPF3 might serve as a tissue-specific anchor that recruits the BAF complex to heart and muscle relevant genes via the interaction with modified histones thereby acting as co-regulator for DNA-binding transcription factors. Given the high degree of interdependency

found in this study, we highly believe that a systems-level view of the different regulatory levels is a mandatory step for a full understanding of a living system. A future identification of precise molecular mechanisms will be challenging, but will open new perspectives to understand heart development and complex cardiovascular disorders.

8 REFERENCES

1. Crick, F. Central dogma of molecular biology. *Nature* **227**, 561-3 (1970).
2. Hampsey, M. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol Mol Biol Rev* **62**, 465-503 (1998).
3. Kadonaga, J.T. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell* **116**, 247-57 (2004).
4. Lewis, B.A. & Reinberg, D. The mediator coactivator complex: functional and physical roles in transcriptional regulation. *J Cell Sci* **116**, 3667-75 (2003).
5. Thomas, M.C. & Chiang, C.M. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**, 105-78 (2006).
6. Wasserman, W.W. & Sandelin, A. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* **5**, 276-87 (2004).
7. Hahn, S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**, 394-403 (2004).
8. Asturias, F.J. RNA polymerase II structure, and organization of the preinitiation complex. *Curr Opin Struct Biol* **14**, 121-9 (2004).
9. Blazek, E., Mittler, G. & Meisterernst, M. The mediator of RNA polymerase II. *Chromosoma* **113**, 399-408 (2005).
10. Cramer, P. RNA polymerase II structure: from core to functional complexes. *Curr Opin Genet Dev* **14**, 218-26 (2004).
11. Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A. & Luscombe, N.M. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* **10**, 252-63 (2009).
12. Lemon, B. & Tjian, R. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**, 2551-69 (2000).
13. Sandelin, A. *et al.* Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet* **8**, 424-36 (2007).
14. Birney, E. *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816 (2007).
15. Levine, M. & Tjian, R. Transcription regulation and animal diversity. *Nature* **424**, 147-51 (2003).
16. Riethoven, J.J. Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods Mol Biol* **674**, 33-42 (2010).
17. Ogbourne, S. & Antalis, T.M. Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* **331 (Pt 1)**, 1-14 (1998).
18. Kornberg, R.D. Structure of chromatin. *Annu Rev Biochem* **46**, 931-54 (1977).
19. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60 (1997).
20. Wang, Y. *et al.* Linking covalent histone modifications to epigenetics: the rigidity and plasticity of the marks. *Cold Spring Harb Symp Quant Biol* **69**, 161-9 (2004).
21. Ho, L. & Crabtree, G.R. Chromatin remodelling during development. *Nature* **463**, 474-84.
22. Suzuki, M.M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**, 465-76 (2008).
23. Evans, S.M., Yelon, D., Conlon, F.L. & Kirby, M.L. Myocardial lineage development. *Circ Res* **107**, 1428-44 (2010).
24. Moorman, A.F., de Jong, F., Denyn, M.M. & Lamers, W.H. Development of the cardiac conduction system. *Circ Res* **82**, 629-44 (1998).
25. Garry, D.J. & Olson, E.N. A common progenitor at the heart of development. *Cell* **127**, 1101-4 (2006).

26. Martin-Puig, S., Wang, Z. & Chien, K.R. Lives of a heart cell: tracing the origins of cardiac progenitors. *Cell Stem Cell* **2**, 320-31 (2008).
27. Wu, S.M. *et al.* Developmental origin of a bipotential myocardial and smooth muscle cell precursor in the mammalian heart. *Cell* **127**, 1137-50 (2006).
28. Bruneau, B.G. The developmental genetics of congenital heart disease. *Nature* **451**, 943-8 (2008).
29. Srivastava, D. Making or breaking the heart: from lineage determination to morphogenesis. *Cell* **126**, 1037-48 (2006).
30. Harvey, R.P. Patterning the vertebrate heart. *Nat Rev Genet* **3**, 544-56 (2002).
31. Kelly, R.G. Building the right ventricle. *Circ Res* **100**, 943-5 (2007).
32. Christoffels, V.M. *et al.* Chamber formation and morphogenesis in the developing mammalian heart. *Dev Biol* **223**, 266-78 (2000).
33. Chen, H. *et al.* Overexpression of bone morphogenetic protein 10 in myocardium disrupts cardiac postnatal hypertrophic growth. *J Biol Chem* **281**, 27481-91 (2006).
34. Soonpaa, M.H., Kim, K.K., Pajak, L., Franklin, M. & Field, L.J. Cardiomyocyte DNA synthesis and binucleation during murine development. *Am J Physiol* **271**, H2183-9 (1996).
35. Clubb, F.J., Jr. & Bishop, S.P. Formation of binucleated myocardial cells in the neonatal rat. An index for growth hypertrophy. *Lab Invest* **50**, 571-7 (1984).
36. Leu, M., Ehler, E. & Perriard, J.C. Characterisation of postnatal growth of the murine heart. *Anat Embryol (Berl)* **204**, 217-24 (2001).
37. Olson, E.N. Gene regulatory networks in the evolution and development of the heart. *Science* **313**, 1922-7 (2006).
38. Hoffman, J.I. & Kaplan, S. The incidence of congenital heart disease. *J Am Coll Cardiol* **39**, 1890-900 (2002).
39. Anderson, R.H. & Tynan, M. Tetralogy of Fallot--a centennial review. *Int J Cardiol* **21**, 219-32 (1988).
40. Ho, S., McCarthy, K.P., Josen, M. & Rigby, M.L. Anatomic-echocardiographic correlates: an introduction to normal and congenitally malformed hearts. *Heart* **86 Suppl 2**, II3-11 (2001).
41. Di Felice, V. & Zummo, G. Tetralogy of fallot as a model to study cardiac progenitor cell migration and differentiation during heart development. *Trends Cardiovasc Med* **19**, 130-5 (2009).
42. Greenway, S.C. *et al.* De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* **41**, 931-5 (2009).
43. Nemer, G. *et al.* A novel mutation in the GATA4 gene in patients with Tetralogy of Fallot. *Hum Mutat* **27**, 293-4 (2006).
44. Clark, K.L., Yutzey, K.E. & Benson, D.W. Transcription factors and congenital heart defects. *Annu Rev Physiol* **68**, 97-121 (2006).
45. MacDonald, S.T. *et al.* Epiblastic Cited2 deficiency results in cardiac phenotypic heterogeneity and provides a mechanism for haploinsufficiency. *Cardiovasc Res* **79**, 448-57 (2008).
46. Sperling, S. *et al.* Identification and functional analysis of CITED2 mutations in patients with congenital heart defects. *Hum Mutat* **26**, 575-82 (2005).
47. Toenjes, M. *et al.* Prediction of cardiac transcription networks based on molecular data and complex clinical phenotypes. *Mol Biosyst* **4**, 589-98 (2008).
48. Cripps, R.M. & Olson, E.N. Twist is required for muscle template splitting during adult Drosophila myogenesis. *Dev Biol* **203**, 106-15 (1998).
49. Chien, K.R., Domian, I.J. & Parker, K.K. Cardiogenesis and the complex biology of regenerative cardiovascular medicine. *Science* **322**, 1494-7 (2008).
50. Nemer, M. Genetic insights into normal and abnormal heart development. *Cardiovasc Pathol* **17**, 48-54 (2008).
51. Alsan, B.H. & Schultheiss, T.M. Regulation of avian cardiogenesis by Fgf8 signaling. *Development* **129**, 1935-43 (2002).

52. Charron, F., Paradis, P., Bronchain, O., Nemer, G. & Nemer, M. Cooperative interaction between GATA-4 and GATA-6 regulates myocardial gene expression. *Mol Cell Biol* **19**, 4355-65 (1999).
53. Belaguli, N.S. *et al.* Cardiac tissue enriched factors serum response factor and GATA-4 are mutual coregulators. *Mol Cell Biol* **20**, 7550-8 (2000).
54. Dai, Y.S., Cserjesi, P., Markham, B.E. & Molkentin, J.D. The transcription factors GATA4 and dHAND physically interact to synergistically activate cardiac gene expression through a p300-dependent mechanism. *J Biol Chem* **277**, 24390-8 (2002).
55. Garg, V. *et al.* GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* **424**, 443-7 (2003).
56. Lee, Y. *et al.* The cardiac tissue-restricted homeobox protein Csx/Nkx2.5 physically associates with the zinc finger protein GATA4 and cooperatively activates atrial natriuretic factor gene expression. *Mol Cell Biol* **18**, 3120-9 (1998).
57. Morin, S., Charron, F., Robitaille, L. & Nemer, M. GATA-dependent recruitment of MEF2 proteins to target promoters. *EMBO J* **19**, 2046-55 (2000).
58. Molkentin, J.D., Lin, Q., Duncan, S.A. & Olson, E.N. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev* **11**, 1061-72 (1997).
59. Karamboulas, C. *et al.* Disruption of MEF2 activity in cardiomyoblasts inhibits cardiomyogenesis. *J Cell Sci* **119**, 4315-21 (2006).
60. Lin, Q., Schwarz, J., Bucana, C. & Olson, E.N. Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* **276**, 1404-7 (1997).
61. Naya, F.J. *et al.* Mitochondrial deficiency and cardiac sudden death in mice lacking the MEF2A transcription factor. *Nat Med* **8**, 1303-9 (2002).
62. Wang, L., Fan, C., Topol, S.E., Topol, E.J. & Wang, Q. Mutation of MEF2A in an inherited disorder with features of coronary artery disease. *Science* **302**, 1578-81 (2003).
63. Luxenburg, C., Pasolli, H.A., Williams, S.E. & Fuchs, E. Developmental roles for Srf, cortical cytoskeleton and cell shape in epidermal spindle orientation. *Nat Cell Biol* **13**, 203-14.
64. Sun, Q. *et al.* Defining the mammalian CArGome. *Genome Res* **16**, 197-207 (2006).
65. Niu, Z., Li, A., Zhang, S.X. & Schwartz, R.J. Serum response factor micromanaging cardiogenesis. *Curr Opin Cell Biol* **19**, 618-27 (2007).
66. Miano, J.M. Serum response factor: toggling between disparate programs of gene expression. *J Mol Cell Cardiol* **35**, 577-93 (2003).
67. Miano, J.M., Long, X. & Fujiwara, K. Serum response factor: master regulator of the actin cytoskeleton and contractile apparatus. *Am J Physiol Cell Physiol* **292**, C70-81 (2007).
68. Wang, D. *et al.* Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* **105**, 851-62 (2001).
69. Posern, G. & Treisman, R. Actin' together: serum response factor, its cofactors and the link to signal transduction. *Trends Cell Biol* **16**, 588-96 (2006).
70. Kaplan-Albuquerque, N., Van Putten, V., Weiser-Evans, M.C. & Nemenoff, R.A. Depletion of serum response factor by RNA interference mimics the mitogenic effects of platelet derived growth factor-BB in vascular smooth muscle cells. *Circ Res* **97**, 427-33 (2005).
71. Soulez, M., Tuil, D., Kahn, A. & Gilgenkrantz, H. The serum response factor (SRF) is needed for muscle-specific activation of CArG boxes. *Biochem Biophys Res Commun* **219**, 418-22 (1996).
72. Wei, L. *et al.* RhoA signaling via serum response factor plays an obligatory role in myogenic differentiation. *J Biol Chem* **273**, 30287-94 (1998).
73. Sepulveda, J.L., Vlahopoulos, S., Iyer, D., Belaguli, N. & Schwartz, R.J. Combinatorial expression of GATA4, Nkx2-5, and serum response factor directs early cardiac gene activity. *J Biol Chem* **277**, 25775-82 (2002).

74. Schratt, G. *et al.* Serum response factor is crucial for actin cytoskeletal organization and focal adhesion assembly in embryonic stem cells. *J Cell Biol* **156**, 737-50 (2002).
75. Balza, R.O., Jr. & Misra, R.P. Role of the serum response factor in regulating contractile apparatus gene expression and sarcomeric integrity in cardiomyocytes. *J Biol Chem* **281**, 6498-510 (2006).
76. Arsenian, S., Weinhold, B., Oelgeschlager, M., Ruther, U. & Nordheim, A. Serum response factor is essential for mesoderm formation during mouse embryogenesis. *Embo J* **17**, 6289-99 (1998).
77. 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* **102**, 1082-7 (2005).
78. Miano, J.M. *et al.* Restricted inactivation of serum response factor to the cardiovascular system. *Proc Natl Acad Sci U S A* **101**, 17132-7 (2004).
79. Takeuchi, J.K. *et al.* Tbx5 specifies the left/right ventricles and ventricular septum position during cardiogenesis. *Development* **130**, 5953-64 (2003).
80. Lyons, I. *et al.* Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev* **9**, 1654-66 (1995).
81. Plageman, T.F., Jr. & Yutzey, K.E. T-box genes and heart development: putting the "T" in heart. *Dev Dyn* **232**, 11-20 (2005).
82. Basson, C.T. *et al.* Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genet* **15**, 30-5 (1997).
83. Srivastava, D. & Olson, E.N. A genetic blueprint for cardiac development. *Nature* **407**, 221-6 (2000).
84. Srivastava, D. *et al.* Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nat Genet* **16**, 154-60 (1997).
85. Riley, P.R., Gertsenstein, M., Dawson, K. & Cross, J.C. Early exclusion of hand1-deficient cells from distinct regions of the left ventricular myocardium in chimeric mouse embryos. *Dev Biol* **227**, 156-68 (2000).
86. Bruneau, B.G. *et al.* A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell* **106**, 709-21 (2001).
87. Takeuchi, J.K. *et al.* Tbx20 dose-dependently regulates transcription factor networks required for mouse heart and motoneuron development. *Development* **132**, 2463-74 (2005).
88. Pu, W.T., Ishiwata, T., Juraszek, A.L., Ma, Q. & Izumo, S. GATA4 is a dosage-sensitive regulator of cardiac morphogenesis. *Dev Biol* **275**, 235-44 (2004).
89. Spivakov, M. & Fisher, A.G. Epigenetic signatures of stem-cell identity. *Nat Rev Genet* **8**, 263-71 (2007).
90. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).
91. Jenuwein, T. & Allis, C.D. Translating the histone code. *Science* **293**, 1074-80 (2001).
92. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature* **403**, 41-5 (2000).
93. Ahn, S.H. *et al.* Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in S. cerevisiae. *Cell* **120**, 25-36 (2005).
94. Koch, C.M. *et al.* The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res* **17**, 691-707 (2007).
95. Bhaumik, S.R., Smith, E. & Shilatifard, A. Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* **14**, 1008-16 (2007).
96. Roth, S.Y., Denu, J.M. & Allis, C.D. Histone acetyltransferases. *Annu Rev Biochem* **70**, 81-120 (2001).
97. Denslow, S.A. & Wade, P.A. The human Mi-2/NuRD complex and gene regulation. *Oncogene* **26**, 5433-8 (2007).
98. Shahbazian, M.D. & Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* **76**, 75-100 (2007).

99. Chang, S. *et al.* Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. *Mol Cell Biol* **24**, 8467-76 (2004).
100. Zhang, C.L. *et al.* Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell* **110**, 479-88 (2002).
101. Yao, T.P. *et al.* Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361-72 (1998).
102. Kook, H. *et al.* Cardiac hypertrophy and histone deacetylase-dependent transcriptional repression mediated by the atypical homeodomain protein Hop. *J Clin Invest* **112**, 863-71 (2003).
103. Song, K. *et al.* The transcriptional coactivator CAMTA2 stimulates cardiac growth by opposing class II histone deacetylases. *Cell* **125**, 453-66 (2006).
104. Antos, C.L. *et al.* Dose-dependent blockade to cardiomyocyte hypertrophy by histone deacetylase inhibitors. *J Biol Chem* **278**, 28930-7 (2003).
105. Kong, Y. *et al.* Suppression of class I and II histone deacetylases blunts pressure-overload cardiac hypertrophy. *Circulation* **113**, 2579-88 (2006).
106. Haberland, M., Montgomery, R.L. & Olson, E.N. The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet* **10**, 32-42 (2009).
107. Shilatifard, A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* **75**, 243-69 (2006).
108. Allis, C.D. *et al.* New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-6 (2007).
109. Martin, C. & Zhang, Y. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* **6**, 838-49 (2005).
110. Fuks, F. DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* **15**, 490-5 (2005).
111. Wang, H. *et al.* Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Mol Cell* **8**, 1207-17 (2001).
112. Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D. & Patel, D.J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**, 1025-40 (2007).
113. Hawkins, R.D. & Ren, B. Genome-wide location analysis: insights on transcriptional regulation. *Hum Mol Genet* **15 Spec No 1**, R1-7 (2006).
114. McDonald, O.G. & Owens, G.K. Programming smooth muscle plasticity with chromatin dynamics. *Circ Res* **100**, 1428-41 (2007).
115. Davis, F.J., Gupta, M., Camoretti-Mercado, B., Schwartz, R.J. & Gupta, M.P. Calcium/calmodulin-dependent protein kinase activates serum response factor transcription activity by its dissociation from histone deacetylase, HDAC4. Implications in cardiac muscle gene regulation during hypertrophy. *J Biol Chem* **278**, 20047-58 (2003).
116. Lu, J., McKinsey, T.A., Zhang, C.L. & Olson, E.N. Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. *Mol Cell* **6**, 233-44 (2000).
117. Sparrow, D.B. *et al.* MEF-2 function is modified by a novel co-repressor, MITR. *Embo J* **18**, 5085-98 (1999).
118. Youn, H.D., Grozinger, C.M. & Liu, J.O. Calcium regulates transcriptional repression of myocyte enhancer factor 2 by histone deacetylase 4. *J Biol Chem* **275**, 22563-7 (2000).
119. Lusser, A. & Kadonaga, J.T. Chromatin remodeling by ATP-dependent molecular machines. *Bioessays* **25**, 1192-200 (2003).
120. Eberharter, A. & Becker, P.B. ATP-dependent nucleosome remodelling: factors and functions. *J Cell Sci* **117**, 3707-11 (2004).
121. Bao, Y. & Shen, X. SnapShot: chromatin remodeling complexes. *Cell* **129**, 632 (2007).

122. Stankunas, K. *et al.* Endocardial Brg1 represses ADAMTS1 to maintain the microenvironment for myocardial morphogenesis. *Dev Cell* **14**, 298-311 (2008).
123. Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A* **106**, 5181-6 (2009).
124. Yan, Z. *et al.* BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells* **26**, 1155-65 (2008).
125. Lessard, J. *et al.* An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55**, 201-15 (2007).
126. Hang, C.T. *et al.* Chromatin regulation by Brg1 underlies heart muscle development and disease. *Nature* **466**, 62-7.
127. Takeuchi, J.K. *et al.* Chromatin remodelling complex dosage modulates transcription factor function in heart development. *Nat Commun* **2**, 187.
128. Lickert, H. *et al.* Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **432**, 107-12 (2004).
129. Narlikar, G.J., Fan, H.Y. & Kingston, R.E. Cooperation between complexes that regulate chromatin structure and transcription. *Cell* **108**, 475-87 (2002).
130. Wu, J.I., Lessard, J. & Crabtree, G.R. Understanding the words of chromatin regulation. *Cell* **136**, 200-6 (2009).
131. Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-6 (1999).
132. Lange, M. *et al.* Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev* **22**, 2370-84 (2008).
133. Shen, W. *et al.* Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails. *Biochemistry* **46**, 2100-10 (2007).
134. Jacobson, R.H., Ladurner, A.G., King, D.S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-5 (2000).
135. Maurer-Stroh, S. *et al.* The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci* **28**, 69-74 (2003).
136. Li, H. *et al.* Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91-5 (2006).
137. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-20 (2001).
138. Lusser, A., Urwin, D.L. & Kadonaga, J.T. Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nat Struct Mol Biol* **12**, 160-6 (2005).
139. Whetstine, J.R. *et al.* Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**, 467-81 (2006).
140. Huyen, Y. *et al.* Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **432**, 406-11 (2004).
141. Koga, H. *et al.* A human homolog of Drosophila lethal(3)malignant brain tumor (l(3)mbt) protein associates with condensed mitotic chromosomes. *Oncogene* **18**, 3799-809 (1999).
142. Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86-90 (2006).
143. Pena, P.V. *et al.* Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100-3 (2006).
144. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-97 (2004).
145. Sevignani, C., Calin, G.A., Siracusa, L.D. & Croce, C.M. Mammalian microRNAs: a small world for fine-tuning gene expression. *Mamm Genome* **17**, 189-202 (2006).

146. Lee, R.C., Feinbaum, R.L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-54 (1993).
147. Reinhart, B.J. *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403**, 901-6 (2000).
148. Friedman, R.C., Farh, K.K., Burge, C.B. & Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**, 92-105 (2009).
149. Bentwich, I. *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* **37**, 766-70 (2005).
150. Lagos-Quintana, M. *et al.* Identification of tissue-specific microRNAs from mouse. *Curr Biol* **12**, 735-9 (2002).
151. Liu, N. & Olson, E.N. MicroRNA regulatory networks in cardiovascular development. *Dev Cell* **18**, 510-25 (2010).
152. Kim, V.N. & Nam, J.W. Genomics of microRNA. *Trends Genet* **22**, 165-73 (2006).
153. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-60 (2004).
154. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F. & Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231-5 (2004).
155. Yi, R., Qin, Y., Macara, I.G. & Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-6 (2003).
156. Chendrimada, T.P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740-4 (2005).
157. Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**, 102-14 (2008).
158. Fazi, F. & Nervi, C. MicroRNA: basic mechanisms and transcriptional regulatory networks for cell fate determination. *Cardiovasc Res* **79**, 553-61 (2008).
159. Chen, J.F. *et al.* The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet* **38**, 228-33 (2006).
160. Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**, 214-20 (2005).
161. Zhao, Y. *et al.* Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell* **129**, 303-17 (2007).
162. Ivey, K.N. *et al.* MicroRNA regulation of cell lineages in mouse and human embryonic stem cells. *Cell Stem Cell* **2**, 219-29 (2008).
163. Liu, N. *et al.* microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev* **22**, 3242-54 (2008).
164. Divakaran, V. & Mann, D.L. The emerging role of microRNAs in cardiac remodeling and heart failure. *Circ Res* **103**, 1072-83 (2008).
165. Latronico, M.V., Catalucci, D. & Condorelli, G. Emerging role of microRNAs in cardiovascular biology. *Circ Res* **101**, 1225-36 (2007).
166. Thum, T., Catalucci, D. & Bauersachs, J. MicroRNAs: novel regulators in cardiac development and disease. *Cardiovasc Res* **79**, 562-70 (2008).
167. Thum, T. *et al.* MicroRNAs in the human heart: a clue to fetal gene reprogramming in heart failure. *Circulation* **116**, 258-67 (2007).
168. Tan, Y. *et al.* Transcriptional inhibiton of Hoxd4 expression by miRNA-10a in human breast cancer cells. *BMC Mol Biol* **10**, 12 (2009).
169. Gonzalez, S., Pisano, D.G. & Serrano, M. Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. *Cell Cycle* **7**, 2601-8 (2008).
170. White, S.M., Constantin, P.E. & Claycomb, W.C. Cardiac physiology at the cellular level: use of cultured HL-1 cardiomyocytes for studies of cardiac muscle cell structure and function. *Am J Physiol Heart Circ Physiol* **286**, H823-9 (2004).

171. Anisimov, S.V., Tarasov, K.V., Riordon, D., Wobus, A.M. & Boheler, K.R. SAGE identification of differentiation responsive genes in P19 embryonic cells induced to form cardiomyocytes in vitro. *Mech Dev* **117**, 25-74 (2002).
172. Boheler, K.R. *et al.* Differentiation of pluripotent embryonic stem cells into cardiomyocytes. *Circ Res* **91**, 189-201 (2002).
173. Hescheler, J. *et al.* Embryonic stem cells: a model to study structural and functional properties in cardiomyogenesis. *Cardiovasc Res* **36**, 149-62 (1997).
174. Klug, M.G., Soonpaa, M.H., Koh, G.Y. & Field, L.J. Genetically selected cardiomyocytes from differentiating embronic stem cells form stable intracardiac grafts. *J Clin Invest* **98**, 216-24 (1996).
175. Rudnicki, M.A., Reuhl, K.R. & McBurney, M.W. Cell lines with developmental potential restricted to mesodermal lineages isolated from differentiating cultures of pluripotential P19 embryonal carcinoma cells. *Development* **107**, 361-72 (1989).
176. Wobus, A.M., Kleppisch, T., Maltsev, V. & Hescheler, J. Cardiomyocyte-like cells differentiated in vitro from embryonic carcinoma cells P19 are characterized by functional expression of adrenoceptors and Ca2+ channels. *In Vitro Cell Dev Biol Anim* **30A**, 425-34 (1994).
177. Claycomb, W.C. *et al.* HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte. *Proc Natl Acad Sci U S A* **95**, 2979-84 (1998).
178. Seymour, E.M. *et al.* HL-1 myocytes exhibit PKC and K(ATP) channel-dependent delta opioid preconditioning. *J Surg Res* **114**, 187-94 (2003).
179. Aggeli, I.K., Beis, I. & Gaitanaki, C. ERKs and JNKs mediate hydrogen peroxide-induced Egr-1 expression and nuclear accumulation in H9c2 cells. *Physiol Res* **59**, 443-54.
180. Kimes, B.W. & Brandt, B.L. Properties of a clonal muscle cell line from rat heart. *Exp Cell Res* **98**, 367-81 (1976).
181. Sperling, S.R. Systems biology approaches to heart development and congenital heart disease. *Cardiovasc Res* **91**, 269-78.
182. He, A., Kong, S.W., Ma, Q. & Pu, W.T. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc Natl Acad Sci U S A* **108**, 5632-7.
183. Jiang, H., Daniels, P.J. & Andrews, G.K. Putative zinc-sensing zinc fingers of metal-response element-binding transcription factor-1 stabilize a metal-dependent chromatin complex on the endogenous metallothionein-I promoter. *J Biol Chem* **278**, 30394-402 (2003).
184. Visel, A., Rubin, E.M. & Pennacchio, L.A. Genomic views of distant-acting enhancers. *Nature* **461**, 199-205 (2009).
185. Johnson, D.S. *et al.* Systematic evaluation of variability in ChIP-chip experiments using predefined DNA targets. *Genome Res* **18**, 393-403 (2008).
186. Valouev, A. *et al.* Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat Methods* **5**, 829-34 (2008).
187. Ji, H. *et al.* An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat Biotechnol* **26**, 1293-300 (2008).
188. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* **15**, 1034-50 (2005).
189. Morrison, T.B., Weis, J.J. & Wittwer, C.T. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques* **24**, 954-8, 960, 962 (1998).
190. Holland, P.M., Abramson, R.D., Watson, R. & Gelfand, D.H. Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci U S A* **88**, 7276-80 (1991).
191. Livak, K.J., Flood, S.J., Marmaro, J., Giusti, W. & Deetz, K. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for

detecting PCR product and nucleic acid hybridization. *PCR Methods Appl* **4**, 357-62 (1995).
192. 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* **280**, 32531-8 (2005).
193. Gitter, A. *et al.* Backup in gene regulatory networks explains differences between binding and knockout results. *Mol Syst Biol* **5**, 276 (2009).
194. Hu, Z., Killion, P.J. & Iyer, V.R. Genetic reconstruction of a functional transcriptional regulatory network. *Nat Genet* **39**, 683-7 (2007).
195. Harbison, C.T. *et al.* Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99-104 (2004).
196. Gao, F., Foat, B.C. & Bussemaker, H.J. Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinformatics* **5**, 31 (2004).
197. McDonald, O.G., Wamhoff, B.R., Hoofnagle, M.H. & Owens, G.K. Control of SRF binding to CArG box chromatin regulates smooth muscle gene expression in vivo. *J Clin Invest* **116**, 36-48 (2006).
198. Cao, D. *et al.* Modulation of smooth muscle gene expression by association of histone acetyltransferases and deacetylases with myocardin. *Mol Cell Biol* **25**, 364-76 (2005).
199. Manabe, I. & Owens, G.K. Recruitment of serum response factor and hyperacetylation of histones at smooth muscle-specific regulatory regions during differentiation of a novel P19-derived in vitro smooth muscle differentiation system. *Circ Res* **88**, 1127-34 (2001).
200. Qiu, P. & Li, L. Histone acetylation and recruitment of serum responsive factor and CREB-binding protein onto SM22 promoter during SM22 gene expression. *Circ Res* **90**, 858-65 (2002).
201. Schlesinger, J. *et al.* The cardiac transcription network modulated by Gata4, Mef2a, Nkx2.5, Srf, histone modifications, and microRNAs. *PLoS Genet* **7**, e1001313 (2011).
202. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859-72 (2005).
203. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019-31 (2009).
204. Martin, D.G., Grimes, D.E., Baetz, K. & Howe, L. Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. *Mol Cell Biol* **26**, 3018-28 (2006).
205. Bustin, S.A. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**, 169-93 (2000).
206. Ding, C. & Cantor, C.R. Quantitative analysis of nucleic acids--the last few years of progress. *J Biochem Mol Biol* **37**, 1-10 (2004).
207. Ninkina, N.N. *et al.* Cerd4, third member of the d4 gene family: expression and organization of genomic locus. *Mamm Genome* **12**, 862-6 (2001).
208. Chestkov, A.V., Baka, I.D., Kost, M.V., Georgiev, G.P. & Buchman, V.L. The d4 gene family in the human genome. *Genomics* **36**, 174-7 (1996).
209. Kaynak, B. *et al.* Genome-wide array analysis of normal and malformed human hearts. *Circulation* **107**, 2467-74 (2003).
210. Mujtaba, S., Zeng, L. & Zhou, M.M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**, 5521-7 (2007).
211. Wang, Y.X. *et al.* Requirements of myocyte-specific enhancer factor 2A in zebrafish cardiac contractility. *FEBS Lett* **579**, 4843-50 (2005).
212. van Weerd, H., Koshiba-Takeuchi, K., Kwon, C. & Takeuchi, J.K. Epigenetic Factors and Cardiac Development. *Cardiovasc Res*.
213. Boogerd, C.J. *et al.* Functional analysis of novel TBX5 T-box mutations associated with Holt-Oram syndrome. *Cardiovasc Res* **88**, 130-9 (2010).

214. Pipes, G.C., Creemers, E.E. & Olson, E.N. The myocardin family of transcriptional coactivators: versatile regulators of cell growth, migration, and myogenesis. *Genes Dev* **20**, 1545-56 (2006).
215. Small, E.M. *et al.* Myocardin-related transcription factor-a controls myofibroblast activation and fibrosis in response to myocardial infarction. *Circ Res* **107**, 294-304.
216. Creemers, E.E., Sutherland, L.B., Oh, J., Barbosa, A.C. & Olson, E.N. Coactivation of MEF2 by the SAP domain proteins myocardin and MASTR. *Mol Cell* **23**, 83-96 (2006).
217. Treisman, R. Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors. *Cell* **46**, 567-74 (1986).
218. Wang, Z. *et al.* Myocardin and ternary complex factors compete for SRF to control smooth muscle gene expression. *Nature* **428**, 185-9 (2004).
219. Yang, A. *et al.* Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Mol Cell* **24**, 593-602 (2006).
220. Li, X.Y. *et al.* The role of chromatin accessibility in directing the widespread, overlapping patterns of Drosophila transcription factor binding. *Genome Biol* **12**, R34.
221. Holtzinger, A. & Evans, T. Gata5 and Gata6 are functionally redundant in zebrafish for specification of cardiomyocytes. *Dev Biol* **312**, 613-22 (2007).
222. Vong, L.H., Ragusa, M.J. & Schwarz, J.J. Generation of conditional Mef2cloxP/loxP mice for temporal- and tissue-specific analyses. *Genesis* **43**, 43-8 (2005).
223. Gertz, J. & Cohen, B.A. Environment-specific combinatorial cis-regulation in synthetic promoters. *Mol Syst Biol* **5**, 244 (2009).
224. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-12 (2009).
225. Heintzman, N.D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311-8 (2007).
226. Morse, R.H. Epigenetic marks identify functional elements. *Nat Genet* **42**, 282-4 (2010).
227. Ramirez, S., Ait-Si-Ali, S., Robin, P., Trouche, D. & Harel-Bellan, A. The CREB-binding protein (CBP) cooperates with the serum response factor for transactivation of the c-fos serum response element. *J Biol Chem* **272**, 31016-21 (1997).
228. Blobel, G.A., Nakajima, T., Eckner, R., Montminy, M. & Orkin, S.H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc Natl Acad Sci U S A* **95**, 2061-6 (1998).
229. Kakita, T. *et al.* p300 protein as a coactivator of GATA-5 in the transcription of cardiac-restricted atrial natriuretic factor gene. *J Biol Chem* **274**, 34096-102 (1999).
230. Bernstein, B.E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169-81 (2005).
231. Hsu, M. *et al.* Complex developmental patterns of histone modifications associated with the human beta-globin switch in primary cells. *Exp Hematol* **37**, 799-806 e4 (2009).
232. Kim, T.H. *et al.* A high-resolution map of active promoters in the human genome. *Nature* **436**, 876-80 (2005).
233. Gillespie, R.F. & Gudas, L.J. Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. *J Mol Biol* **372**, 298-316 (2007).
234. Backs, J. & Olson, E.N. Control of cardiac growth by histone acetylation/deacetylation. *Circ Res* **98**, 15-24 (2006).
235. Tando, T. *et al.* Requiem protein links RelB/p52 and the Brm-type SWI/SNF complex in a noncanonical NF-kappaB pathway. *J Biol Chem* **285**, 21951-60 (2010).
236. Zeng, L. *et al.* Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**, 258-62 (2010).
237. Lim, L.P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769-73 (2005).

238. Linsley, P.S. *et al.* Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. *Mol Cell Biol* **27**, 2240-52 (2007).
239. Tsang, J., Zhu, J. & van Oudenaarden, A. MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell* **26**, 753-67 (2007).
240. Nunez-Iglesias, J., Liu, C.C., Morgan, T.E., Finch, C.E. & Zhou, X.J. Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer's disease cortex reveals altered miRNA regulation. *PLoS One* **5**, e8898 (2010).
241. Wang, Y.P. & Li, K.B. Correlation of expression profiles between microRNAs and mRNA targets using NCI-60 data. *BMC Genomics* **10**, 218 (2009).
242. Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455**, 64-71 (2008).
243. Cheng, J. *et al.* The impact of miR-34a on protein output in hepatocellular carcinoma HepG2 cells. *Proteomics* **10**, 1557-72.
244. Vinther, J., Hedegaard, M.M., Gardner, P.P., Andersen, J.S. & Arctander, P. Identification of miRNA targets with stable isotope labeling by amino acids in cell culture. *Nucleic Acids Res* **34**, e107 (2006).
245. Somel, M. *et al.* MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res* **20**, 1207-18 (2010).
246. John, B. *et al.* Human MicroRNA targets. *PLoS Biol* **2**, e363 (2004).
247. Huang, Y. *et al.* A study of miRNAs targets prediction and experimental validation. *Protein Cell* **1**, 979-86 (2010).
248. Kuhn, D.E. *et al.* Experimental validation of miRNA targets. *Methods* **44**, 47-54 (2008).
249. Klose, R.J. & Bird, A.P. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* **31**, 89-97 (2006).
250. Cheng, X. & Blumenthal, R.M. Mammalian DNA methyltransferases: a structural perspective. *Structure* **16**, 341-50 (2008).
251. Poetsch, A.R. & Plass, C. Transcriptional regulation by DNA methylation. *Cancer Treat Rev*.
252. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**, 6-21 (2002).
253. Rountree, M.R. & Selker, E.U. DNA methylation and the formation of heterochromatin in Neurospora crassa. *Heredity* **105**, 38-44.
254. Jones, P.A. & Baylin, S.B. The epigenomics of cancer. *Cell* **128**, 683-92 (2007).
255. Grant, S.G. & Chapman, V.M. Mechanisms of X-chromosome regulation. *Annu Rev Genet* **22**, 199-233 (1988).
256. Movassagh, M. *et al.* Differential DNA methylation correlates with differential expression of angiogenic factors in human heart failure. *PLoS One* **5**, e8564 (2010).
257. Hu, G. *et al.* Regulation of nucleosome landscape and transcription factor targeting at tissue-specific enhancers by BRG1. *Genome Res* (2011).
258. Zhou, J. *et al.* The SWI/SNF chromatin remodeling complex regulates myocardin-induced smooth muscle-specific gene expression. *Arterioscler Thromb Vasc Biol* **29**, 921-8 (2009).
259. Baudino, T.A., Carver, W., Giles, W. & Borg, T.K. Cardiac fibroblasts: friend or foe? *Am J Physiol Heart Circ Physiol* **291**, H1015-26 (2006).
260. Camelliti, P., Borg, T.K. & Kohl, P. Structural and functional characterisation of cardiac fibroblasts. *Cardiovasc Res* **65**, 40-51 (2005).
261. Jia, H. *et al.* Casein kinase 2 promotes Hedgehog signaling by regulating both smoothened and Cubitus interruptus. *J Biol Chem* **285**, 37218-26 (2010).

9 APPENDIX

9.1 Summary

For the development and maintenance of all eukaryotic organisms, the process of gene transcription needs to be regulated by a set of different cellular levels comprising transcriptional, epigenetic and post-transcriptional mechanisms. Taking the mammalian heart as model, we present a systems biology approach investigating the regulation of mRNA profiles by the interplay between tissue-specific transcription factors (TF), co-occurring histone modifications and miRNAs.

The analysis of ChIP-chip and RNAi mediated knockdown experiments for the four key cardiac TFs Gata4, Mef2a, Nkx2.5 and Srf revealed a high number of common regulated genes as well as binding sites. Strikingly, we found a high number of genes bound by multiple TFs that were significantly less likely differentially expressed in the respective knockdown, indicating a complex and cooperative regulation of gene expression whereby these four non-paralogous can potentially compensate each other's function. Co-occurring histone 3 acetylation (H3ac) were found to have a high impact on the expression of Srf and Gata4 target genes. The correlation between Srf and H3ac was further investigated using ChIP-seq and ChIP-qPCR after Srf knockdown. Our results provide evidence that Srf triggers the acetylation at particular target genes and moreover, that H3ac buffers the activating function of Srf in its knockdown. In addition, we identified a panel of miRNAs deregulated in Srf knockdown that can explain three times more differentially expressed genes than Srf binding events alone can do.

To get a better understanding of the Srf-driven regulation of transcription we investigated the binding of Srf and the histone acetyltransferase p300 and the co-occurrence of H3ac and histone 3 lysine 4 di-methylation. We selected regulatory regions important for heart and muscle development and screened their binding behaviour in a time-series of mouse hearts of three developmental stages around birth using ChIP-qPCR. We observed a strong correlation between all four factors pointing to a common regulatory mechanism, which is triggered by Srf and has H3ac as outcome that is to a certain degree independent from p300. In addition, the correlation of expression levels of target genes with the dynamic changes of the studied factors indicated a high variability of combinatorial regulation.

For our comprehensive ChIP experiments in mouse hearts, we applied the novel LightCycler® 1536 real-time PCR System, which we evaluated beforehand for its applicability in context of a beta-side testing of the Roche Company. We found that the new system is a very applicable qPCR system that offers the analysis of 1536 reactions per run within a short

time period and with low consumption of reagents and sample material, providing a very suitable PCR system for medium and high-throughput analyses.

Finally, the functional role of DPF3 was characterized in more detail as it emerged as likely cardiac relevant due to its up-regulation in patients with congenital heart disease. We identified DPF3 as novel chromatin remodeling factor important for heart and muscle development. Tandem Affinity Purification followed by MassSpec Analysis revealed DPF3 as a component of the human BAF chromatin remodeling complex. In addition, the isoform DPF3b contains the first double PHD finger that binds methylated lysine residues of histone 4 as well as acetylated lysine residues of histones.

To summarize, this work revealed a fine-tuned regulation of cardiac gene expression directed by the combinatorial influence of key cardiac TFs, while histone modifications and miRNAs modulate their functional consequence with a high degree of interdependency. We gained more insights into the functional role of DPF3, which might serve as tissue-specific anchor that recruits the BAF complex to heart and muscle relevant genes via its interaction with modified histones.

9.2 Zusammenfassung

Für die Entwicklung und Aufrechterhaltung aller eukaryotischen Organismen muss die zeitliche und zellspezifische Expression von Genen durch einen umfassenden Satz von verschiedenen zellulären Mechanismen reguliert werden. Zum besseren Verständnis der molekularen Grundlagen der Transkription, untersuchten wir das Zusammenspiel von drei Ebenen, die an der Regulation von kardialen Transkriptionsnetzwerken beteiligt sind. Diese beinhalteten die Bindung herzspezifischer Transkriptionsfaktoren (TF) an Zielgene, das gleichzeitige Auftreten von Histonmodifikationen, sowie den Einfluss von miRNAs auf post-transkriptioneller Ebene.

Die Durchführung und Analyse von umfangreichen ChIP-chip und RNAi-vermittelten Knockdown Experimenten mit den für die Herzentwicklung wichtigen TF Gata4, Mef2a, Nkx2.5 und Srf ergab eine hohe Anzahl von gemeinsam regulierten Genen sowie Bindestellen. Interessanterweise waren Zielgene, die von mehreren TF gebunden wurden, am seltensten im jeweiligen Knockdown dereguliert. Dies lässt auf eine komplexe und kooperative Regulation der Genexpression schließen und deutet auf einen potentiellen Puffereffekt durch kombinatorische Bindung hin, wobei die Genexpression auch beim Ausfall eines einzelnen TF weiter aufrechterhalten werden kann. Eine gleichzeitige Acetylierung von Histon 3 (H3ac) zeigte einen signifikanten Einfluss auf die Expression von Gata4- und Srf Zielgenen. Weitere ChIP-seq und ChIP-qPCR Experimente nach gezieltem Srf Knockdown bestätigten den Zusammenhang zwischen Srf und H3ac. Zudem deuteten unsere Daten

sowohl einen Puffereffekt der H3ac auf die Expression von Srf-Zielgenen in dessen Knockdown, als auch eine Srf-gesteuerte Acetylierung an ausgewählten Zielgenen an. Unsere Untersuchung hinsichtlich des Einflusses von miRNAs ergab, dass bis zu 45% aller differentiell exprimierten Gene durch den Einfluß von miRNAs erklärt werden können, wodurch diese vielversprechende Kandidaten für die Regulation von indirekten Zielgenen sind.

Mittels ChIP-qPCR von Mäuseherzproben verschiedener Entwicklungsstadien kurz vor und nach der Geburt, analysierten wir den Zusammenhang zwischen der Srf-getriebenen Transkriptionsregulation in Korrelation mit der Histon-Acetyltransferase p300 und dem gleichzeitigen Auftreten von H3ac und Histon-3 Lysin-4-Dimethylierung an ausgewählten herz- und muskelspezifischen Zielgenen. Wir beobachteten eine deutliche zeitliche Korrelation zwischen allen vier Faktoren, was auf einen gemeinsamem Mechanismus hindeutet, der von Srf ausgelöst wird und eine H3ac zur Folge hat, die zu einem gewissen Grad p300-unabhängig zu seien scheint. Darüber hinaus zeigte die Korrelation zwischen Expression von Zielgenen und dynamischer Bindung eine hohe Variabilität durch kombinatorische Regulierung.

Für unsere umfangreichen ChIP Experimente mit Mäuseherzproben verwendeten wir das neue LightCycler® 1536 Real-Time PCR System, welches wir zuvor auf seine Anwendbarkeit im Rahmen einer Evaluierung geleitet durch die Firma Roche, getestet haben. Das neue System ermöglichte die Analyse von 1536 Reaktionen pro Lauf innerhalb eines kurzen Zeitraums mit geringem Proben- und Materialverbrauch, und ist damit ein gut geeignetes PCR-System für Mittel- und Hochdurchsatzanalysen.

Der für die Herz- und Muskelentwicklung wichtige epigenetische TF DPF3, welcher in Patienten mit angeborenen Herzfehlen stark hochreguliert ist, konnte mittels Affinitätsaufreinigung als Komponente des humanen BAF Chromatin-Remodeling-Komplexes identifiziert werden. Die Isoform DPF3b enthält den ersten bekannten Doppel-PHD Finger, der sowohl methylierte als auch acetylierte Histone erkennt.

Zusammenfassend kann gesagt werden, dass unsere durchgeführten Analysen zu Regulationsmechanismen von kardialen Transkriptionsnetzwerken ein hohes Maß an Komplexität und Flexibilität ergaben, wobei die verschiedenen genetischen, epigentischen und post-transkriptionellen Regulationsebenen vielfach miteinander verknüpft sind und wechselwirken. Des Weiteren identifizierten wir DPF3 als gewebespezifische Untereinheit des BAF Komplexes, der durch gezielte Interaktion mit Histonmodifikationen diesen eventuell zu herz- und muskelspezifischen Genen rekrutieren kann.

9.3 Abbreviations

Ac	Acetylation
ANOVA	Analysis of variance
ASD	Atrial septal defect
AVSD	Atrioventricular septal defect
BAF	Brahma-associated factor
bp	Base pair
CBP	CREB-binding protein
CHD	Congenital heart defect/disease
ChIP	Chromatin immunoprecipitation
ChIP-chip	Chromatin immunoprecipitation followed by microarray analysis
ChIP-seq	Chromatin immunoprecipitation followed by Next-Generation Sequencing
CKII	Casein kinase 2
Cys	Cysteine
DNA	Desoxyribonucleic acid
DNase	Desoxyribunuclease
E9.5	Embryonic day 9.5 in mouse
ENCODE	ENCyclopedia of DNA Elements
ESC	Embryonic stem cell
FHF	First heart field
GO	Gene Ontology
H	Histone
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
His	Histidine
IFT	Inflow tract
K	Lysine
kb	Kilo base
Kme/me2/me3	Mono-, di-, tri-methylation of lysine
Me	Methylation
miRNA	MicroRNA
mRNA	Messenger RNA
NGS	Next-Generation Sequencing
NMR	Nuclear magnetic resonance
nt	Nucleotide
OFT	Outflow tract
p	p-value
P0.5	Postnatal day 0.5 in mouse
PCR	Polymerase chain reaction
PHD finger	Plant-homeodomain
Pol II	RNA polymerase II
qPCR	Quantitative real-time PCR
RISC	RNA-induced silencing complex
RNA	Ribunucleic acid
RNAi	RNA interference
RNase	Ribonuclease
rRNA	Ribosomal RNA

SHF	Secondary heart field
siRNA	Short interfering RNA
SMC	Smooth muscle cell
TAP-MS	Tandem affinity purification followed by mass spectrometry
TF	Transcription factor
TOF	Tetralogy of Fallot
TSS	Trancriptional start site
UTR	Untranslated region
VSD	Ventricular septal defect

9.4 Curriculum Vitae

For reasons of privacy protection, a complete CV is not included in the electronic version of the thesis.

Publications

Schueler M*, Zhang Q*, **Schlesinger J**, Tönjes M, Sperling SR. Dynamics of Srf, p300 and histone modifications during cardiac maturation in mouse. *In preparation*. * Authors contributed equally.

Schlesinger J*, Schueler M*, Grunert M*, Fischer JJ*, Zhang Q, Krueger T, Lange M, Tönjes M, Dunkel I, Sperling SR (2011) The Cardiac Transcription Network Modulated by Gata4, Mef2a, Nkx2.5, Srf, Histone Modifications, and MicroRNAs. PLoS Genet 7: e1001313. * Authors contributed equally.

Schlesinger J, Tönjes M, Schueler M, Zhang Q, Dunkel I, Sperling SR (2010) Evaluation of the LightCycler 1536 Instrument for high-throughput quantitative real-time PCR. Methods 50: S19-22.

Lange M, Kaynak B, Forster UB, Tonjes M, Fischer JJ, Grimm C, **Schlesinger J**, Just S, Dunkel I, Krueger T, Mebus S, Lehrach H, Lurz R, Gobom J, Rottbauer W, Abdelilah-Seyfried S, Sperling S (2008) Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev 22: 2370-2384.

Publications in Public Science

Schlesinger J, Schueler M, Sperling SR (2011) Komplex und flexibel - die Regulation der kardialen Genexpression. Laborwelt.

Tönjes M, **Schlesinger J**, Schüler M, Dunkel I, Bethune J, Sperling SR (2010) High-Throughput Cardiac Gene Expression Analysis Using the Universal ProbeLibrary and the Novel LightCycler® 1536 Real-Time PCR System. Biochemica.

Tönjes M, **Schlesinger J**, Schueler M, Dunkel I, Bethune J, Sperling SR (2010) Hochdurchsatz-Expressionsanalyse kardialer Gene. Biospektrum.

Selected Conference Talks and Poster Presentations

“The Cardiac Transcription Network Driven by the Interplay of Transcription Factors, Histone Modifications and MicroRNAs”. HeartRepair Annual Conference 19 May 2010, Amsterdam, Netherlands

“The Cardiac Transcription Network Driven by the Interplay of Transcription Factors, Histone Modifications and MicroRNAs”. Weinstein Cardiovascular Development Conference 20-22 May 2010, Amsterdam, Netherlands

“The epigenetic transcription factor DPF3 translocates during development from the cytoplasm to the nucleus”. HeartRepair Annual Conference 5-8 April 2009, Berlin, Germany

“Funktional analysis and identification of interaction partners of the transcription factor DPF3.” 20th International Congress of Genetics, July 2008, Berlin, Germany

9.5 Danksagung (Acknowledgements)

Hiermit möchte ich mich herzlich bei allen bedanken, die mich im Laufe der Zeit hier am Max-Planck-Institut bei der Durchführung meiner Arbeit unterstützt und stets mit Rat und Tat zur Seite gestanden haben.

An erster Stelle möchte ich mich bei Prof. Dr. Silke R. Sperling bedanken für die Möglichkeit diese Arbeit in ihrer Arbeitsgruppe anfertigen zu können, für ihre gute wissenschaftliche Betreuung und die Begutachtung dieser Arbeit, sowie die Ermöglichung an verschiedenen internationalen Konferenzen teilnehmen zu können. Ich danke Herrn Prof. Dr. Hans Lehrach für die Möglichkeit meine Arbeit in seiner Abteilung am MPI anfertigen zu können. Prof. Dr. Thomas Schmülling danke ich für die Bereitschaft meine Arbeit am Fachbereich Biologie, Chemie, Pharmazie der Freien Universität Berlin zu begutachten.

Ein ganz besonderer Dank geht an alle früheren und heutigen Mitglieder des „Sperling Labs“, die für eine einzigartige Arbeitsatmosphäre und einen guten Zusammenhalt gesorgt haben. Markus Schüler und Marcel Grunert danke ich für ihre Unterstützung und die oft erheiternden Stunden der Zusammenarbeit. Markus, du warst eine große Stütze und unglaublich guter Motivator. Dank Dir, weiß ich, dass ich auch im nächsten Leben niemals Bioinformatik studieren kann. Marcel, du warst mein Retter bei vielen Computerproblemen. Ein Dank an dein ansteckendes Lachen und die vielen gemeinsamen Heimfahrten. Ilona Dunkel für den Fakt, dass sie die beste TA der Welt ist. Ich dank Dir für dein Durchhaltevermögen, die viele Hilfe im Labor und den regen Austausch über Erziehungsmethoden. Cornelia Dorn danke ich für ihre unglaubliche Gelassenheit, das Verbot mit dem “i” and our „denglisch spells“. Markus, Marcel und Conny, unser letzes gemeinsames Jahr war einzigartig und unsere kleinen Abstecher in den Keller nach den Mittagspausen werde ich nicht vergessen. Ein herzliches Dankeschön an meine ehemalige Kollegin und Freundin Martje Tönjes für einfach alles; deine Persönlichkeit, dein positives Denken und den unermüdlichen Drang anderen zu helfen. Qin Zhang, thanks for your warm personality, your help in the lab, your encouragements, and your delicious Chinese food. Barbara Gibbas für ihre großartige Hilfe bei allen administrativen Angelegenheiten, sowie die 1a-Versorgung mit Leckerlies auf dem „K“-Tisch. Martin Lange für seine geduldige Art und all die vielen Tipps. Katharina Rost für lustige Vesperminuten. Alan Punnoose einfach, weil er ein netter Kerl war. Jenny Fischer und Dr. Christina Grimm für die Einführung in das Lab mit all seinen Tücken.

Die Zeit hier war eine unvergeßliche und es haben mich viele Mitarbeiter am Institut auf verschiedene Weise unterstützt. Die Mitglieder der Arbeitsgruppen von Harald Seitz und Diego Walther möchte ich besonders hervorheben für die nette Zusammenarbeit hier auf dem Gang in Turm 4. Silke Stahlberg, Robert Wild und Linda Hallen danke ich für die vielen netten Unterhaltungen, auch abseits des Wissenschaftsalltags. Rudi Lurz war ein echter

Schatz, stets zu Diensten, wenn ich Probleme mit den Mikroskopen hatte. Danke! Pia Kuss und Florian Witte bin ich zu Dank verpflichtet für die Einführung in die Welt der Immunhistologie. Ich bedanke mich bei Silke Stahlberg und Stephan Klatt für die Zusammenarbeit als Chair in der STA. Ich danke Mirjam Peetz vom Tierhaus für ihre Unterstützung. Prof. Dr. Martin Vingron danke ich für sein Engagement die Interessen der Studenten am Institut umzusetzen.
Der Firma Roche danke ich für die tolle Zusammenarbeit und die Möglichkeit ihr neues LightCycler® 1536 Real-Time PCR System zu testen.

Meinen Freunden danke ich für die nötige Ablenkung und den Rückhalt außerhalb der Forschung. Mein aller größter Dank geht an Guido Witt und meine Eltern für Eure Liebe und Euren Glauben an mich.

9.6 Selbständigkeitserklärung

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfaßt habe und keine anderen als die angegebenen Quellen und Hilfsmittel in Anspruch genommen habe. Ich versichere, dass diese Arbeit in dieser oder anderer Form keiner anderen Prüfungsbehörde vorgelegt wurde.

Jenny Schlesinger **Berlin, den 20.09.2011**