universität wien

MASTERARBEIT

Titel der Masterarbeit

„The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with Their Respective Genes During Early Cardiomyogenesis“

verfasst von

Hannah Paar, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, April 2014

Studienkennzahl lt. Studienblatt:	A 066834
Studienrichtung lt. Studienblatt:	Masterstudium Molekulare Biologie
Betreut von:	Ao. Univ. Prof. Dr. Georg Weitzer

Table of Contents

Abstract

Cardiovascular disease is still the leading cause of death worldwide. The inherent regenerative potential of the human heart is not sufficient to prevent scar formation in the case of a cardiac event. The discovery of somatic cardiac stem cells has for the first time provided the opportunity to develop treatment strategies which could increase the regeneration ability of diseased and scarred heart tissue. In order to successfully differentiate somatic cardiac stem cells into fully functional cardiomyocytes, the transcriptional network regulating cardiac differentiation has to be investigated.

In order to gain new insights into this transcriptional network regulating early cardiomyogenesis, we concentrated on the transcription factors Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin. We performed chromatin immunoprecipitation with both embryonic stem cells (ESCs) and cardiovascular progenitor cells (CVPCs) at carefully selected time points during differentiation in order to investigate the physical interaction of the five transcription factors of interest with their respective genes. The embryoid and cardiac body in vitro model systems were used to obtain data which reflect the transcriptional regulation during early embryogenesis and cardiomyogenesis, respectively.

Nanog is a well-established marker for pluripotency, but so far not much is known about its role during differentiation. Brachyury and Mesp1 are both markers for early mesoderm and have been shown to regulate each other's expression. Nkx2.5 is a marker for the cardiac lineage and is expressed throughout cardiac development, including the adult heart. Desmin is mainly known for its function as an intermediate filament in muscle cells. However, evidence has been found to support the theory that Desmin also plays an important role in gene regulation during early cardiac development.

We could demonstrate that all five investigated transcription factors physically interact with their respective genes in a very complex and dynamic way during the early stages of cardiomyogenesis. Furthermore, the conducted experiments underline the differences between ESCs and CVPCs and indicate that some genes are regulated in a different manner in those two cell types. Furthermore, our data provide indications of possible new regulatory DNA segments in the 5'UTRs of the investigated genes.

Zusammenfassung

Herz-Kreislauf Erkrankungen sind noch immer die häufigste Todesursache weltweit. Das angeborene Regenerationspotential des menschlichen Herzens ist nicht ausreichend, um Narbenbildung im Falle eines akuten kardialen Ereignisses zu verhindern. Die Entdeckung somatischer Herzstammzellen bietet zum ersten Mal die Möglichkeit Behandlungsstrategien zu entwickeln, welche die Regenerationsfähigkeit von krankem und vernarbtem Herzgewebe erhöhen könnten. Um somatische Herzstammzellen erfolgreich in voll funktionsfähige Kardiomyozyten zu differenzieren, muss das transkriptionelle Netzwerk, welches die Differenzierung von Herzzellen reguliert, untersucht werden.

Wir haben uns auf die Transkriptionsfaktoren Nanog, Brachyury, Mesp1, Nkx2.5 und Desmin konzentriert, um neue Erkenntnisse über dieses transkriptionelle Netzwerk, das die frühe Herzentwicklung reguliert, zu gewinnen. Wir untersuchten die Interaktion dieser fünf Transkriptionsfaktoren mit ihren jeweiligen Genen mittels Chromatin Immunopräzipitation zu sorgfältig ausgewählten Zeitpunkten während der Differenzierung von embryonalen Stammzellen (ESZ) und kardiovaskulären Vorläuferzellen (KVZ). Die in-vitro Modellsysteme „Embryoid Bodies“ und „Cardiac Bodies“ wurden verwendet, um Daten zu erhalten, welche die transkriptionelle Regulierung während der Embryogenese beziehungsweise der Kardiomyogenese widerspiegeln.

Nanog ist ein fest etablierter Marker für Pluripotenz, aber über seine Rolle während der Differenzierung ist bisher nicht viel bekannt. Brachyury und Mesp1 sind Marker für frühes Mesoderm und es wurde gezeigt, dass beide das Gen des Anderen regulieren. Nkx2.5 ist ein Marker für KVZ und wird, angefangen bei der Kardiogenese bis hin zum erwachsenen Herzen, exprimiert. Desmin ist hauptsächlich für seine Funktion als Intermediärfilament in Muskelzellen bekannt. Es wurden jedoch auch Beweise dafür gefunden, dass Desmin zusätzlich eine wichtige Rolle in der Genregulierung während der frühen Herzentwicklung spielt.

Wir konnten zeigen, dass während der frühen Stadien der Kardiomyogenese alle fünf untersuchten Transkriptionsfaktoren auf eine sehr komplexe und dynamische Weise mit ihren jeweiligen Genen interagieren. Außerdem unterstreichen die durchgeführten Experimente die Unterschiede zwischen ESZ und KVZ und zeigen, dass einige Gene in unterschiedlicher Art und Weise in jenen zwei Zellarten reguliert werden. Weiters liefern

unsere Daten Hinweise auf mögliche neue regulierende DNA Segmente in den 5 `UTRen der untersuchten Gene.

Abbreviations

Abbreviation	Meaning
A	Adenine
ADP	Adenosin diphosphate
Akt	Protein kinase B
ANF	Anti nuclear factor
AP	Anterior-posterior
AVE	Anterior visceral endoderm
Axin2	Axis inhibition protein 2
b-HLH	Basic helix loop helix
ß-ME	Beta mercaptoethanol
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BMP	Bone morphogenic protein
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
C	Cytosine
CB	Cardiac body
Cdx	Caudal homeobox protein
ChIP	Chromatin immunoprecipitation
CVD	Cardiovascular disease
DVE	Distal visceral endoderm
CVPCs	Cardiovascular progenitor cells
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
E	Embryonic day
EB	Embryoid body
EMT	Epithelial-mesenchymal transition
ENCODE	Encyclopedia of DNA Elements
Epi	Epiblast
ESCs	Embryonic stem cells
Esrrb	Estrogen related receptor beta
Exe	Extraembryonic ectoderm
FACS	Fluorescence activated cell sorting
Fgf	Fibroblast growth factor
Fox	Forkhead box protein
G	Guanine
Gcnf	Germ cell nuclear factor
gp130	Glycoprotein 130
GSK3β	Glycogen synthase kinase 3 beta
Hand	Heart- and neural crest derivatives-expressed protein
hESCs	Human embryonic stem cells
HDAC	Histone deactetylase
ICD	International Statistical Classification of Diseases and Related Health Problems
ICM	Inner cell mass
IF	Intermediate filament
Isl-1	LIM-homeodomain transcription factor Islet 1
Jak	Janus kinase
Klf4	Kruppel-like factor 4
Lef	Lymphoid enhancer-binding factor
LIF	Leukemia inhibitory factor
MAPK	Mitogen-activated protein kinases
M-CAT	Malonyl CoA-acyl carrier protein transacylase

MCE	Minimal cardiac enhancer
MCPs	Multipotent cardiovascular progenitors
MEK	MAPK/ERK kinase
Mef2	Macrolide-efflux protein 2
mESCs	Murine embryonic stem cells
Mesp1/2	Mesoderm posterior homolog 1/2
miRNA	MicroRNA
mRNA	Messenger RNA
MyoD	Myoblast determination protein
NBT	Nitro blue tetrazolium
Nkx2.3/5/6	NK2 transcription factor related, locus 3/5/6
Notch	Neurogenic locus notch homolog protein?
Ntl1	Nitrate transporter protein
NuRD	Nucleosome Remodeling Deacetylase
Oct4	Octamer-binding transcription factor 4
PCR	Polymerase chain reaction
PE	Primitive endoderm
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
Pitx	Paired-like homeodomain transcription factor
PKA	Protein kinase A
PKC	Protein kinase C
PrE	Primitive ectoderm
PS	Primitive streak
PTM	Posttranslational modification
RNA	Ribonucleic acid
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
Shh	Sonic hedgehog
Smad	Similar to mothers against decapentaplegic
Sox2	SRY-related high-mobility group (HMG)-box protein 2
Sp1/3	Specificity protein 1/3
SRF	Serum response factor
Stat3	Signal transducer and activator of transcription 3
T	Thymidine
Tbx1/3/5	T-box 1/3/5 transcription factor
Tcf3	T-cell factor 3
TGFß	Transforming growth factor beta
Tnn2	Tenascin N protein 2
TF	Transcription factor
VE	Visceral endoderm
3‘/5‘UTR	3‘/5‘ untranslated region
WHO	World Health Association
YAP	Yes-associated protein
Zeb	Zinc finger E-box-binding homeobox protein
Zfp281	Zinc finger protein 281

1. Introduction

1.1. The Cardiovascular Disease

Especially in industrialized nations, the prevalence of cardiovascular disease (CVD) has reached epidemic proportions and it is the number one cause of death globally (WHO, 2013). In America one in three adults is estimated to have some form of CVD. Although advances in the field of medicine have succeeded in reducing mortality rates, patients with CVD still have an increased risk for further cardiac events, like unstable angina, myocardial infarction and heart failure. In the United States, CVD costs more than $108 billion each year, making it also an economic issue (Dickson et al., 2013).

CVDs defined by the ICD codes include all diseases of the circulatory system as well as congenital CVD and refer to people with hypertension, heart disease, stroke, peripheral artery disease and diseases of the veins (Go et al., 2013).

The underlying cause for the severity of many of those diseases is the fact that the human heart has a very low inherent potential of regeneration, although the assumption that there is no regeneration potential whatsoever has been proven wrong (Beltrami et al., 2003). Recent studies indicate that about 50% of cardiomyocytes are replenished during the life of an average human being (Xin et al., 2013). Today research focuses more and more on understanding the process of cardiomyogenesis in the hope that this knowledge will help to develop new methods of increasing the regeneration ability of diseased and scarred heart tissue (Chugh et al., 2012).

1.2. Stem Cells

A stem cell is defined by its ability to self-renew, differentiate into at least one cell type and reconstitute its tissue of origin functionally. There are still discrepancies in the classification of stem cells, but a functional and widely used classification is based on the developmental potential of the cells (Can, 2008):

- Totipotent cells: can give rise to all embryonic and extraembryonic tissues necessary for mammalian development, e.g. zygote
- Pluripotent cells: can give rise to all cell types of the embryonic tissues, including germ cells, e.g. embryonic stem cells (ESCs), induced pluripotent stem cells, inner cell mass of the blastocyst-stage embryo

- Multipotent cells: can give rise to all cells of a certain tissue, e.g. hematopoietic stem cells
- Unipotent cells: can give rise to a single cell type, e.g. spermatogonial stem cells

It is believed that almost every tissue harbors stem cells that are responsible for tissue homeostasis and at least partial regeneration of injured tissue (Can, 2008; Steinhoff, 2013).

1.2.1. Embryonic Stem Cells

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of mammalian embryos before implantation and can be maintained indefinitely in culture (Evans and Kaufman, 1981). This unlimited capacity to self-renew in vitro together with their ability to give rise to all somatic and germ cell lineages of the developing embryo, except extraembryonic tissue, opens wide research opportunities (Nichols and Smith, 2009).

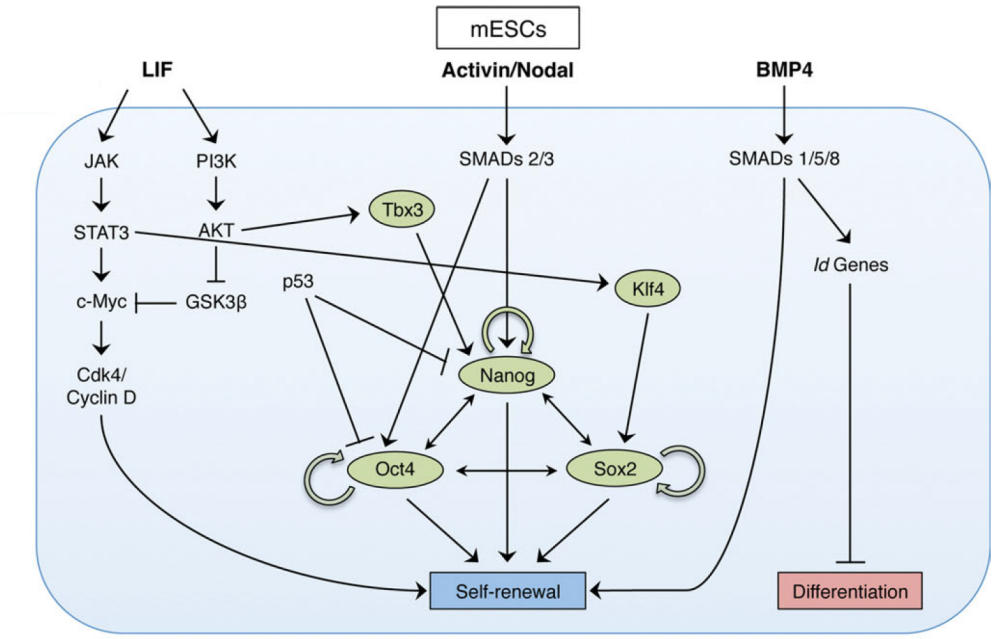


Figure 1.1. Mouse ESC self-renewal and survival pathways. LIF and BMP4 are essential for mESC maintenance. Nanog, Oct4 and Sox2 form a positive autoregulatory loop that ensures continued self-renewal (Saunders et al., 2013).

Mouse ESCs (mESCs) harvested on embryonic day 3.5 (E3.5) display "ground state" pluripotency and remain pluripotent when cultured under specific conditions (Nichols and Smith, 2009). The cytokines leukemia inhibitory factor (LIF) and bone morphogenic protein 4 (BMP4) have been proven to be sufficient for ESCs to self-renew and remain in an undifferentiated state under serum-free conditions and when cultured on mouse embryonic fibroblasts, respectively (Smith et al., 1988; Tesar et al., 2007; Williams et al., 1988; Ying et al., 2003). While LIF ensures self-renewal through the activation of the Jak/Stat3 and PI3K/Akt signaling pathways, BMP activates factors responsible for the inhibition of differentiation through the upregulation of Smad proteins. In ESCs, pluripotency is promoted by the core pluripotency network consisting of the transcription factors (TFs) Nanog, Oct4 and Sox2 as well as by other factors. Klf4 and Tbx3 are upregulated by LIF through the Jak/Stat3 and PI3K/Akt pathways and in turn activate Sox2 and Nanog (Saunders et al., 2013). Mouse ESCs are highly heterogeneous regarding their overall expression profiles (MacArthur et al., 2012).

Human ESCs (hESCs) more closely resemble mouse epiblast stem cells (mEpiSCs) derived after implantation than mESCs and require basic fibroblast growth factor (bFgf) and insulin or insulin-like growth factor (IGF) for survival instead of LIF or BMP4 (Saunders et al., 2013). bFgf acts through activating the MAPK and the Activin/Nodal signaling pathways, while IGF activates the Ras and PI3K pathways. In hESCs, upregulation of Nanog is achieved by Smad 2 and 3, which activate Activin/Noal signaling as well as directly bind and upregulate the Nanog gene (Niwa et al., 2009).

1.2.2. Cardiovascular Progenitor Cells

Until the discovery of replicating myocytes, the heart was believed to be a post-mitotic organ that exerts its function until death with the same or less amount of cells that are present at birth (Hosoda et al., 2011). In recent years evidence for the existence of somatic stem cells in the adult heart, also referred to as heart stem cells or cardiovascular progenitor cells (CVPCs), has been found by many groups (Anversa et al., 2007; Beltrami et al., 2001; Bergmann et al., 2009; Hansson et al., 2009; Kajstura et al., 2008, 2010; Laugwitz et al., 2008; Messina et al., 2004; Wu et al., 2008), but their isolation and stable culture have been proven difficult. CVPCs were not only identified in numerous species, but also in different regions of the heart (Hoebaus et al., 2013): epicardial tissue (Limana et al., 2007), ventricular tissue (Galvez et al., 2008) and heart auricles (Gambini et al., 2011). It is well established that stem cells need a specialized niche in order to preserve

their survival and replication potential (Hosoda et al., 2011). It is believed that the niche is composed of "a subset of tissue cells and extracellular substrates that can indefinitely house one or more stem cells and control their self renewal and progeny production in vivo" (Spradling et al., 2001).

Our lab succeeded in creating a surrogate stem cell niche for the in vitro propagation of stable clonal CVPC lines. With the help of ESCs and LIF-secreting fibroblasts, CVPCs from postnatal murine hearts with a selection marker could be isolated and maintained as stable clonal CVPC lines that self-renew in a LIF-dependent manner. Those cell lines express not only stemness factors like Oct4, Sox2 and Nanog, but also early myocardial TFs like Nkx2.5, Gata4 and Isl-1. Unlike ESCs, they can only give rise to cells of the mesoderm and differentiate into either cardiomyocytes, endothelial cells or smooth muscle cells (Hoebaus et al., 2013).

1.3. Early Mouse Development

1.3.1. Early Events in Mouse Embryogenesis

During the early stages of mouse embryo development, from the fertilized egg to gastrulation, a series of lineage specification events and the simultaneous development of the embryonic axes need to be tightly regulated. Three tasks must be achieved during the early phase of development: (1.) the number of cells has to increase by proliferation, (2.) the number of cell types needs to increase by differentiation and (3.) polarity has to be generated to allow the establishment of the future body axis (Takaoka and Hamada, 2012).

After three rounds of cell division the eight-cell stage is reached, which is followed by compaction of the early mouse embryo, in which an increase in cell-cell contact promotes cell adhesion. This process generates apical-basal polarity in each of the eight cells of the early morula (Pauken and Capco, 2000; Plusa et al., 2005; Takaoka and Hamada, 2012; Vinot et al., 2005).

In the next step the cells of the morula undergo further cleavage to generate the blastocyst. During this fourth division the orientation of the mitotic spindle is random; therefore, both symmetric and asymmetric divisions can occur (Johnson and Ziomek, 1981).

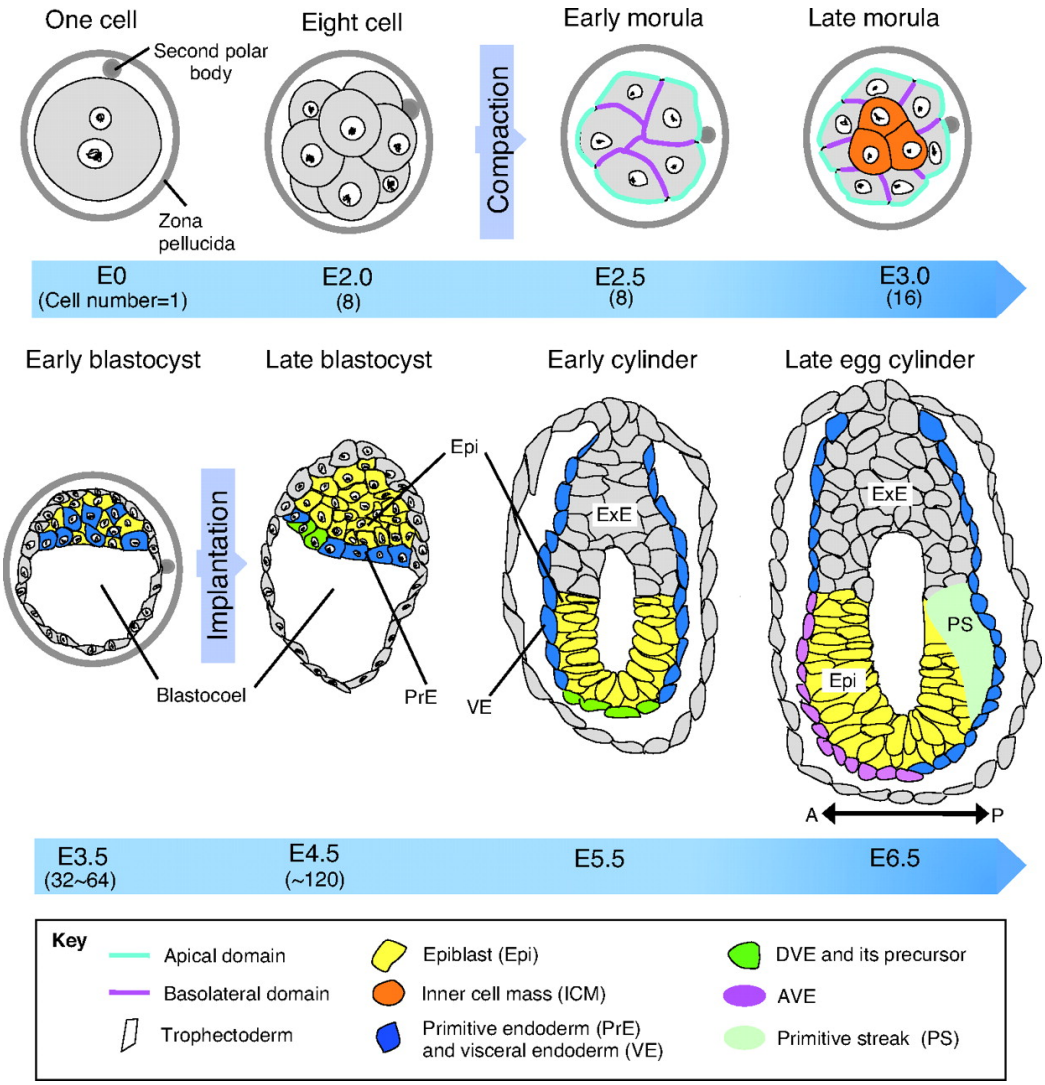


Figure 1.2. Embryonic development in the mouse: from fertilization to gastrulation. Most distinct stages and events from the one cell stage to the late morula on E3.0 (top) and from the early blastocyst to the late egg cylinder stage on E6.5 (bottom). Cell types are color-coded. AVE, anterior visceral endoderm; DVE, distal visceral endoderm; Epi, epiblast; Exe, extraembryonic ectoderm; PrE, primitive ectoderm; PS, primitive streak; VE, visceral endoderm (Takaoka and Hamada, 2012).

The next division resulting in the 32-cell stage generates outside cells and inside cells and this inside-outside polarity leads to Hippo signaling that controls cell contact-mediated inhibition of proliferation and is activated only in the inside-cells (Nishioka et al., 2009). By E3.5 The first cell fate decision has occurred and there are two cell types in the blastocyst: trophectoderm (TE) and inner cell mass (ICM) cells. Cdx2 can be used as a

marker for TE, while Oct3/4 and Nanog expression are indication of ICM (Niwa et al. 2005; Dietrich & Hiiragi 2007; Takaoka & Hamada 2012).

At E4.5 the embryo hatches from the zona pellucida and is implanted in the uterus (Tam and Behringer, 1997). By then, primitive endoderm (PrE) and epiblast cells in the ICM have segregated and are clearly distinguishable by their positions and molecular markers. The PrE cells form a layer on the surface of the ICM and face the blastocoel cavity, while the epiblast is located inside the ICM. In the blastocyst the generation of the ICM and the blastocoel cavity creates the abembryonic axis, which is perpendicular to the border between the ICM and the blastocoel and is the source of future dorsal-ventral polarity (Takaoka and Hamada, 2012).

Next, the anterior-posterior (AP) polarity of the embryo is established: By E6.5 the anterior visceral endoderm (AVE) is formed from the distal visceral endoderm (DVE) on the future anterior side of the embryo. The epiblast is specified to future anterior identity by signals secreted from the AVE, but epiblast cells located farther from the AVE escape the signals and instead form the primitive streak on the opposite side of the embryo (Beddington and Robertson, 1998).

1.3.2. Early Events in Mouse Cardiogenesis

The heart is formed soon after gastrulation and is the first functional organ generated during embryonic development. It is believed that the first ancestor of the primitive heart arose about 500 million years ago and the essential regulatory mechanisms involved in cardiac differentiation have been evolutionarily highly conserved (Bondue et al., 2008). The heart is thought to have evolved through the modification of an ancestral gene regulatory network consisting of TFs of the Nkx, Hand, Gata and Mef families (Bondue and Blanpain, 2010).

The first cardiac progenitors arise during the early stages of gastrulation, when the epiblast cells are incorporated within the cranial portion of the primitive streak. In the mouse the prospective cardiogenic mesoderm migrates away from the primitive streak at E6 in order to form the anterior and lateral plate mesoderm, where the primary heart field (the cardiac crescent) is established at E7 (Bondue et al., 2008). The secondary heart field is formed by cells of the pharyngeal mesoderm, which are located medial and anterior to the primary heart field. By E8 a linear heart tube has been formed through migration of cells from the primary heart field to the midline. This tube is the scaffold for further heart growth. Posterior and anterior expansion of the heart tube is achieved by migration of

cells from the secondary heart field, resulting in the arterial and venous poles. Due to uneven growth and remodeling, the linear heart tube undergoes a rightward looping at E8.5, resulting in the formation of primitive ventricles and atria. This leads to the venous pole moving anteriorly, giving rise to the structure of the future cardiac chambers at E10.5. Heart maturation depends on septation formation in the ventricles and atria and valve formation by E15 (Xin et al., 2013).

The first cardiac contractions in the mouse embryo can be observed as early as E8.5 and E22 in humans. It is necessary for the heart to beat and support circulation before morphogenesis is complete, as the survival of the embryo depends on it (Sedmera, 2011).

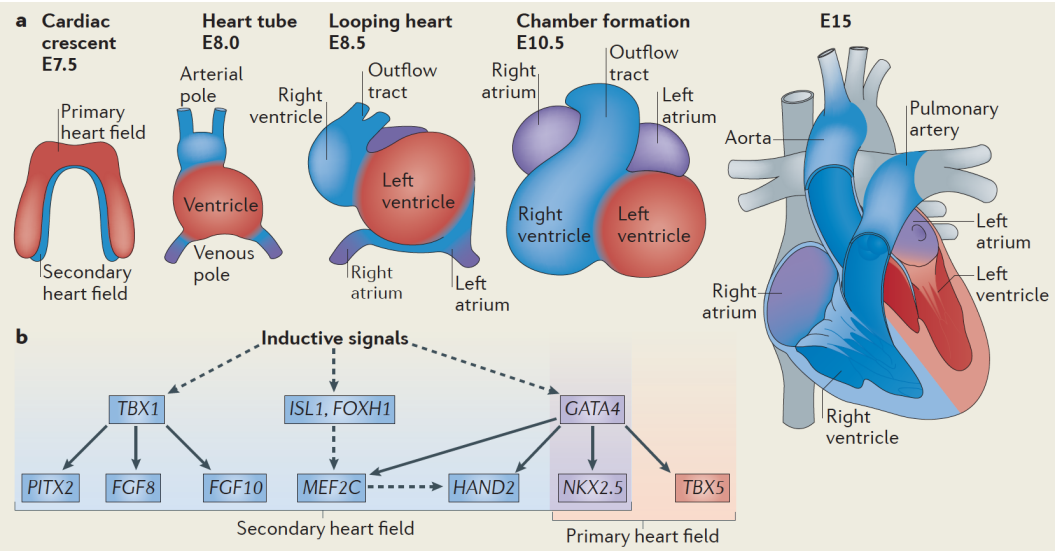


Figure 1.3. Heart development in the mouse: a) the primary and secondary heart fields are formed at E7.5. At E8.0 the linear heart tube is generated, as well as the arterial and venous poles. Rightward looping of the linear heart tube leads to the formation of primitive ventricles and atria at E8.5. This results in the venous pole moving anteriorly and the proper positioning of the future cardiac chambers for further development at E10.5. By E15 septation formation in the ventricles and atria, as well as valve formation have been completed. b) Transcriptional regulation and signaling pathways involved in cardiogenesis (Xin et al., 2013).

Multipotent cardiovascular progenitors (MCPs) differentiate into the various cell lineages that constitute the mature heart, including cardiomyocytes, pacemaker cells, vascular cells and smooth muscle cells. There are two sources of MCPs: The primary heart field MCPs, which give rise to the left ventricle, and cells of both atria and the second heart field

MCPs, which give rise to the right ventricle, atrial cells and cells of the vascular outflow tract. Not much is known about the cellular mechanisms regulating the specification of these early MCPs and their migration from the primitive streak to the anterolateral pole of the embryo (Bondue and Blanpain, 2010).

1.3.3. The Gene Regulatory Network in Cardiogenesis and the Regulation of Cardiomyocyte Proliferation

In cardiogenesis, the gene regulatory events and cellular movements are not only temporally, but also spatially tightly regulated by complex signaling networks. To allow the formation of the heart, intrinsic as well as extrinsic signals and signaling between the primary and secondary heart field are necessary (Xin et al., 2013).

Factors that induce cardiac development in vertebrates are among others BMP, Notch, Wnt and Shh. They activate a set of genes that encode transcriptional activators expressed in the primary and secondary heart fields, which are key regulators of genes that connect upstream signaling to muscle-specific genes and genes encoding proteins involved in heart growth and patterning. TFs found in the primary and secondary heart field are Nkx2.5 and Gata4. Tbx5 is only expressed in the primary heart field and ISL1 and Tbx1 only in the secondary heart field. Among the most important downstream targets of ISL1 and Tbx1 are Mef2c, Hand2, Pitx2, Fgf8 and Fgf10 (Figure 1.3) (Xin et al., 2013). Mutations in essential TFs like Nkx2.5, Tbx5 and Gata4 were shown to cause severe congenital heart defects (Franco and Kelly, 2011).

The evolutionarily conserved Hippo signaling pathway is known for its role in cell proliferation and organ size (Pan, 2010). Several adaptors and kinsases of this pathway have been shown to phosphorylate the transcriptional co-activator Yes-associated protein (YAP) and thereby prevent its entry into the nucleus. This inactivation of YAP is important for maintaining organ size during development. The interaction of unphosphorylated YAP with DNA-binding TFs like TEAD, Smad1, Runx4, Tbx5 and p73 leads to gene expression (Xin et al., 2013). The Hippo pathway is also believed to negatively regulate Wnt signaling and thereby restrain heart size and growth (Heallen et al., 2011).

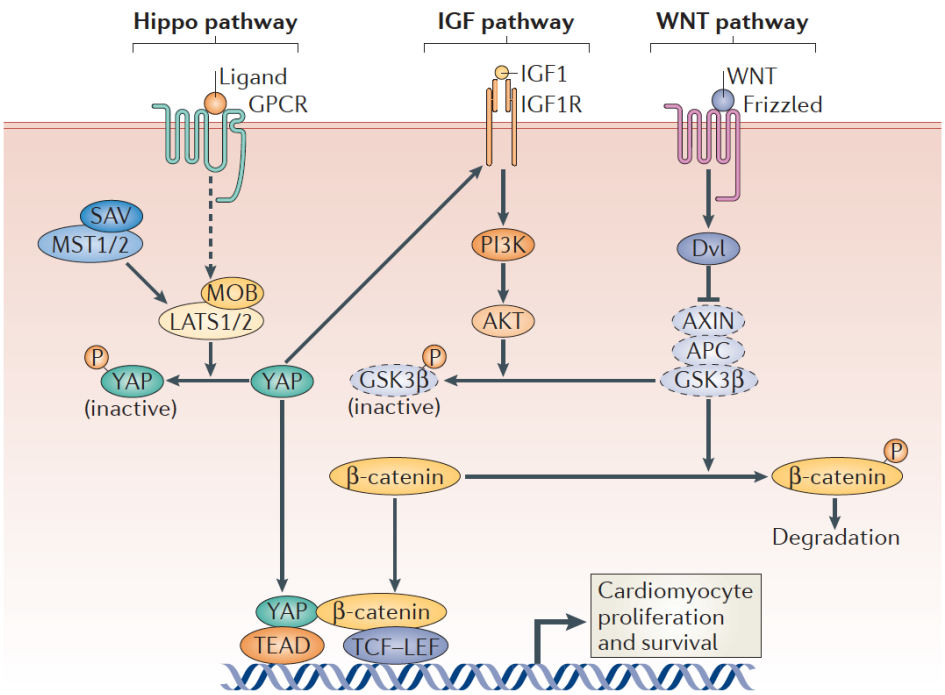


Figure 1.4. Regulation of cardiomyocyte proliferation. Activation of the Hippo pathway leads to the phosphorylation and nuclear exclusion of YAP (Yes-associated protein) to prevent cardiomyocyte proliferation. Unphosphorylated YAP results in cardiomyocyte proliferation, because YAP can activate the expression of genes in the WNT signaling pathway through its interaction with β-catenin. APC, adenomatosis polyposis coli; AXIN, axis inhibition protein; GPCR, G protein-coupled receptor; DVL, Dishevelled; GSK3β, glycogen synthase kinase 3 β; IGF1R, IGF1 receptor; LEF, lymphoid enhancer-binding factor; MOB, Mps one binder; TCF, T cell factor. Dashed lines indicate the involvement of additional factors, and dashed icons represent components of the complex that phosphorylates β-catenin to promote its degradation (Xin et al., 2013).

1.4. The Embryoid Body Model System

When ESCs differentiate in colonies called Embryoid bodies (EBs), they can generate a broad spectrum of cell types representing derivates of all three germ layers (endoderm, mesoderm and ectoderm). Studies have confirmed that molecular and cellular processes involved in the establishment of lineages (hematopoietic, neural, endothelial and skeletal muscle) are very similar in the EB model system and the early embryo, making this a valuable model for research (Kouskoff et al., 2005). Apart from the fact that they facilitate the investigation of several aspects of mouse embryogenesis, EBs provide an

easy method to test the effects of null mutations on early embryogenesis and are a powerful tool for characterizing the function of precursor cells (Desbaillets et al., 2000).
Several methods have been established for the generation of EBs: culture in the absence of feeder cells and LIF, in suspension in bacterial dishes, in spinner flasks, in a semisolid methylcellulose medium or as hanging drops (Desbaillets et al., 2000). This model system recapitulates the complex assembly of cell adhesion and intercellular signaling more accurately than the monolayer differentiation model (Bratt-leal et al., 2009).
EB differentiation begins with the formation of ESC aggregates. The first differentiation step is the formation of a layer of primitive endoderm (PE) on the surface of the EB that shows an epithelial morphology. The PE cells further differentiate into visceral and parietal endoderm and deposit a basement membrane rich in laminin and collagen IV. As differentiation progresses, cells of all three germ layers arise (Bratt-leal et al., 2009).
It was demonstrated that comparable to in vivo embryogenesis, differentiation in embryoid bodies relies on the establishment of signals and morphogen gradients. An example is the local activation of the Wnt pathway that leads to the establishment of anteroposterior polarity and the formation of a primitive streak-like region in the embryoid body (ten Berge et al., 2008).

1.5. Transcriptional Regulation by Transcription Factors

It is well established that the regulation of DNA transcription requires a highly complex network. Proteins that bind DNA also interact with other proteins and form complexes, which in turn create higher order chromatin structures, are involved in specific cellular processes like gene regulation, and change over time in different cell types and genetic backgrounds (Furey, 2013).
Most DNA-binding proteins are able to interact with various DNA sequences giving rise to a sequence “motif” that describes their binding preferences. Such a motif can be more specifically defined as a position weight matrix, which describes the preferred nucleotides (probabilities) at each position in a binding site. These probabilities are derived from the frequency at which a nucleotide is present in already defined binding sites found across the genome. Binding affinity refers to the strength of an interaction and is believed to increase together with binding specificity with high probability nucleotides. Interestingly a study showed that binding intensities were not highly correlated with predicted binding sites in the genome, but did show a high correlation with data derived from chromatin environment data models, indicating a strong influence of chromatin accessibility on

protein-DNA binding. It has been shown that chromatin is highly dynamic and has rather variable differences even across a population of similar cells (Furey, 2013).

1.5.1. Nanog

Nanog together with Oct4 and Sox2 constitutes the core of the pluripotency network, which regulates pluripotent stem cells. They govern not only their own expression by binding each other's promoter region but also regulate other TFs involved in maintaining pluripotency (Boyer et al., 2005; Evans and Kaufman, 1981; Martin, 1981; Mitsui et al., 2003). Nanog especially is responsible for blocking differentiation and establishing a pluripotent ground state during somatic cell reprogramming (Saunders et al., 2013).

1.5.1.1.Nanog Functions in Stem Cell Pluripotency

Nanog is necessary for ICM development, demonstrated by Nanog$^{-/-}$ embryos' inability to form viable epiblasts. Interestingly, conditional Nanog deletion in cultured mESCs made the cells more prone to differentiation, but did not abolish their pluripotent state. This seems to indicate that Nanog is necessary for the establishment of ground state pluripotency, but not for its maintenance (Saunders et al., 2013). On the other hand it was also shown that in Nanog overexpressing cells, Nanog can maintain self-renewal of ESCs independently of the LIF/gp130/Stat3 pathway (Mitsui et al., 2003).

Oct4 and Sox2 protein levels remain relatively stable in undifferentiated ESCs (Saunders et al., 2013), but other important regulators, among them Nanog, exhibit significant temporal expression fluctuations at single-cell level. ESCs with strongly fluctuating Nanog levels are transiently primed for differentiation without definitely and irreversibly committing. This strongly indicates that Nanog acts as a gatekeeper by suppressing spontaneous differentiation but ensuring directed differentiation under appropriate conditions with persistent stimuli. During short-lived Nanog downregulation, the cells seem to adopt a promiscuous co-expression of pluripotency and early lineage marker genes. The fluctuations into a Nanog-low state last about 24h, which is consistent with the observation that Nanog protein levels decline greatly within 24h of Nanog downregulation and with the halftime of the Nanog protein of about two hours (MacArthur et al., 2012).

Recent studies may be able to explain this phenomenon of seemingly spontaneous fluctuation. The reason might lie in the variable allelic expression of Nanog. Nanog expression was monoallelic from the two-cell blastomere stage to the early blastocyst stage, but from then on it became increasingly biallelic, coinciding with establishment of

the pluripotent ground state in the ICM. A subgroup of cells also underwent allelic switching of expression, which might be the reason for Nanog's heterogeneous expression in mESCs (Miyanari and Torres-Padilla, 2012; Saunders et al., 2013).

1.5.1.2. Genetic and Proteomic Features of Nanog

The gene encoding the homeodomain (HD) protein Nanog is located on Chromosome 6 of the mouse genome. The Nanog protein can form functional dimers through its tryptophan-rich domain. Dimerization seems to be essential for mESC self-renewal and pluripotency, as other TFs prefer the interaction with the dimer (Mullin et al. 2008; Saunders et al. 2013). It was shown that the Kruppel-like zink finger TF Zfp281 interacts with the Nanog dimer and seems to play an important role in the regulation of Nanog expression levels as it was found to be required for Nanog binding its own promoter (Saunders et al., 2013).

Both in the human and in the mouse genome various Nanog pseudogenes were found. Their exact role and function is not well understood yet (Booth and Holland, 2004; Zhang et al., 2006). In mESCs Nanog has also been shown to exist in 3 different protein variants (Nanog, Nanog b and Nanog c) created by alternative splicing and the usage of alternate promoters. All three variants can interact and show similar characteristics, but Nanog b and c exhibit reduced functions in ESC maintenance (Das et al., 2011).

Nanog has been shown to preferentially bind the TAAT(G/T)(G/T) motif (Jauch et al., 2008) and is also able to interact with the consensus sequence (C/G)(G/A)(C/G)C(G/C)ATTAN(G/C) (Mitsui et al., 2003).

1.5.1.3. Transcriptional Regulation of Nanog

Nanog has many and complex functions in ESC self-renewal and pluripotency, therefore it is not surprising that it is regulated extensively and promiscuously by a multitude of TFs that have either activating or repressing functions at the Nanog locus (Saunders et al., 2013). In mouse, Nanog expression is primarily monoallelic and fluctuates greatly among ESCs in standard serum/LIF culture conditions (Ying et al. 2008; Miyanari & Torres-Padilla 2012) . It was discovered that most of the positive regulation of Nanog expression is promoted by the proximal promoter region (-289 to +117 relative to the transcription initiation site), which also contains an Oct-Sox enhancer (Kuroda et al., 2005; Rodda et al., 2005). It was proposed that Oct4 and Sox2 are the major regulators of Nanog expression with additional TFs fine-tuning expression in response to specific environmental cues (Saunders et al., 2013). An enhancer element (between -5203 and -

4192 relative to the transcription initiation site) was found, which is active in conditions that promote the appearance of early mesoderm-specified progenitors and to which Stat3 and Brachyury are able to bind and activate transcription (Suzuki et al., 2006).

The Wnt-signaling responsive transcriptional regulator Tcf3 is an example of a negative regulator that binds to an upstream regulatory region in the Nanog locus in order to downregulate Nanog expression and therefore enable differentiation (Pereira et al., 2006). Esrrb on the other hand is an example of a positive regulator that acts together with Oct4 and also directly binds the Nanog locus (van den Berg et al., 2008). Other known factors that modulate Nanog expression are Zfp 143, Klf4, Cdx2, Gcnf, Sp1, Sp3 and many more. Apart from individual TFs, Nanog is also regulated by various signal transduction cascades, like the before mentioned Jak/Stat3 or PI3K/Akt pathways and others like the Fgf/MEK, GSK3β and TGFβ (Saunders et al., 2013).

Moreover, changes in chromatin structure by chromatin remodeling complexes also play an important role in Nanog regulation. The tumor suppressor p53 was shown to bind the Nanog promoter and upon RA-induced differentiation, recruit HDAC complexes to the gene and repress Nanog transcription. Nanog can also regulate its own expression both negatively and positively. Binding of Nanog to its own promoter together with Sox2 or Oct4 activates, while binding together with Zfp281, bound to the NuRD complex, represses Nanog expression. It was shown that the negative autoregulatory feedback loop that prevents Nanog overexpression functions independently of Sox2 and Oct4 (Mitsui et al., 2003).

1.5.1.4. Nanog Targets

Together with Oct4, Sox2, Klf4 and Esrrb, Nanog was found to bind a variety of genes that contain super enhancers. These super enhancers consist of a large cluster of constituent enhancers (few hundred bp to 50 kb), stimulate higher transcriptional activity than typical enhancers and are assumed to control the expression of genes essential for cell identity in mammalian cell types (Whyte et al., 2013).

Brachyury expression can be directly activated by BMP proteins through BMP-responsive elements in the 5'UTR of the Brachyury locus. Nanog can inhibit this activation and therefore downregulate Brachyury expression by inhibiting BMP signaling at the level of the formation of active Smad/p300 complexes, thus inhibiting mesoderm formation (Suzuki et al., 2006).

1.5.1.6. Nanog Post-translational Modifications

In mESCs Nanog has been shown to have several phosphorylation sites in different cellular contexts. Phosphorylation of Nanog seems to have a positive effect on stability and prevents proteasome-mediated degradation and may play an important role in ESC self-renewal (Moretto-Zita et al., 2010). Another post-translational modification used to maintain appropriate Nanog levels is ubiquitination (Buckley et al., 2012). Although some few post-translational modifications of Nanog have been described, it can be expected that many more with diverse and complex functions exist (Saunders et al., 2013).

1.5.1.5. Micro RNAs and Nanog

Tay et al. predicted two miR-296 targets and six miR-470 targets in the coding sequence of the Nanog gene (miR-296: -425 and -683; miR-470: -271, -290, -363, -457, -615, -727 bases of the 5' end of the predicted target from the transcription initiation site). Transfection of mESCs with precursor miRNA oligomers only slightly affected the mRNA level of the endogenous target, but had more substantial effects on the protein. It is presumed that these miRNAs induce post-transcriptional downregulation of its target, probably by somehow inhibiting translation. Therefore, they might play a role in ESC differentiation by regulating protein levels (Tay et al., 2008).

1.5.2. Brachyury

The Brachyury gene is located on Chromosome 17 of the mouse genome and is part of a large family of a so-called T-box genes, which act as transcriptional regulators especially in embryonic development (Müller and Herrmann, 1997). T-box TFs are mainly known for their ability to bind DNA and regulate transcription. Some can activate as well as repress transcription through activation and repression domains in their C-terminal domains. While some genes may be regulated by members of the T-box family alone, most target genes seem to depend on a combination of various TFs (Naiche et al., 2005).

1.6.2.1. Brachyury in Early Development

The first Brachyury (Greek for "short tail) mutant was observed in 1927, while the Brachyury gene was not identified until 1990. Those early mutants were heterozygous and had a short tail. Further experiments showed that mice homozygous for the mutation are nonviable and die at day 11 of gestation due to severe defects in allantois formation (Nibu et al., 2013). It was also shown that these embryos do not develop a notochord and

show severe abnormalities in the development of the neural tube and somites (Herrmann, 1991). When the Brachyury gene was finally identified in 1990 (Herrmann et al., 1990), experiments showed that it is not only expressed in the notochord, but also in primitive-streak mesodermal cells (Wilkinson et al., 1990).

Today it is known that in early mouse embryogenesis, Brachyury is expressed in the primitive streak, tailbud and notochord and plays a vital role in early development in general as well as in gastrulation. Similar expression patterns and functions were found in other species, including Xenopus and zebrafish, leading to the conclusion that Brachyury function is highly conserved (Evans et al., 2012).

Shortly after implantation, Brachyury is expressed in the posterior-proximal side of the epiblast. At the onset of gastrulation it marks the primitive streak and as somitogenesis proceeds, it is localized at the tail tip and notochord (Tanaka, 2009). The notochord is a crucial embryonic structure as it provides support, patterns the central nervous system and is responsible for the development of various organs, including the heart (Nibu et al., 2013) .

Brachyury homologs were found highly conserved in many species, including Xenopus, Ciona, Drosophila, sponges and several other simple organisms. Studies in many of those organisms indicate that the original function of Brachyury was to promote cell movement and adhesion, which are essential for both morphogenesis and tumorigenesis. In the course of evolution, Brachyury ultimately became a driver of mesoderm specification and notochord formation (Nibu et al., 2013).

1.5.2.2. Genetic and Proteomic Features of Brachyury

The T-box binding site (TBE) is the best described DNA binding site for T-box TFs, including Brachyury. Brachyury was shown to bind this palindromic sequence as a dimer, with each monomer binding half of the sequence (5'-AGGTGTGAAATT-3'). The Brachyury dimer is stabilized by a hydrophobic patch and a salt bridge (Naiche et al., 2005). The T box interacts with the DNA through deeply embedding a carboxy terminal helix into an enlarged minor groove without bending the DNA (Müller and Herrmann, 1997). It was also shown that Brachyury is able to bind only one half of the palindromic sequence such as (T/C)TTCACACCT. The fact that so far all tested T-box proteins can bind to the palindromic Brachyury binding site suggests that DNA binding is not enough to promote specificity (Tada and Smith, 2001). Promoter specificity of T-box TFs is provided by interactions with other TFs and different binding kinetics (Naiche et al.,

2005). Evans et al. showed with a ChIP-chip approach that the DNA correlating with most of the obtained binding peaks was not enriched for the canonical T-box binding site 5'-TCACACCT-3', but they did identify enrichment for the simple sequence repeat (AC)n. They did however find sequences very similar to the canonical T-box binding site in the proximity of many of the (AC)n repeats. Further experiments showed that Brachyury binds only weakly to the (AC)n motif alone and that it is likely that this motif stabilizes binding to an adjacent or even distant T-box site in rodents (Evans et al., 2012). It was shown that many TFs bind distal enhancer elements and are linked to a far away promoter region by chromatin looping. This way they can interact with other proteins involved in transcriptional regulation (Farnham, 2009).

1.5.2.3. Transcriptional Regulation of Brachyury

Brachyury is thought to be a direct target of Wnt/ β-catenin signaling in the nascent mesoderm, as its expression was strongly reduced in the absence of β-catenin and therefore Wnt signaling (Bondue and Blanpain, 2010). Somewhere between -396 and -204 bp upstream of the transcription initiation site of the mouse Brachyury gene, a BMP-responsive element was found. This and further experiments suggest that Brachyury is a direct transcriptional target of BMP signaling (Suzuki et al., 2006). Furthermore, Mesp1 was shown to inhibit Brachyury expression through a negative feedback loop (Bondue and Blanpain, 2010).

1.5.2.4. Brachyury Targets

Recent studies indicate that mouse Brachyury binds a variety of targets in differentiating mESCs, including components of signal transduction pathways that direct cell fate in early embryogenesis (Wnt and Fgf signal transduction pathway). Among the 396 identified putative targets were several TFs of the homeobox, winged-helix, paired box, zink-finger and odd-paired family, all of which were also found to be controlled by the zebrafish Brachyury homolog Ntl. Targets not found in zebrafish, but in mouse and hESC-derived mesoderm cells, were among others Axin2, a negative regulator of the Wnt pathway, gamma-catenin/Jup, FgF8 and Wnt3A (Evans et al., 2012).

Murine Brachyury was found to bind together with Stat3 an imperfect T-box palindrome at -4912 bp in the Nanog promoter. This interaction is part of a negative feedback loop that involves Brachyury, Nanog and BMP. Brachyury expression leads to an up-regulation of Nanog, which leads to a suppression of the BMP signal for cell differentiation (Suzuki et al., 2006). It is notable that the experiment leading to those

conclusions was conducted under normal cell culture conditions for undifferentiated stem cells (with serum and LIF) (Tanaka, 2009).

Brachyury has been shown to positively regulate its own expression, once transcription is induced by other TFs. This autoregulatory loop has been detected in Xenopus and human cells so far, but has been ruled out in mice (Nibu et al., 2013). Schmidt et al. showed in T/T mutant ESCs with a T promoter construct linked to a lacZ reporter gene, which is only active in the primitive streak and the tail bud but not in the node and notochord, that lacZ is expressed under the control of the T promoter until 13.5 days of gestation. This allows the conclusion that functional Brachyury protein is not required cell autonomously for the maintenance of its own expression in the primitive streak. Further data suggest that Brachyury might be necessary for maintaining its own expression after the early somite stage (Schmidt et al., 1997).

1.5.2.5. Brachyury and Cancer

Brachyury is overexpressed in many tumor types and seems to be involved in epithelial-mesenchymal transition (EMT). Apart from Brachyury's role in early development, the protein has recently been established as a diagnostic marker to identify chordomas, a rare form of bone cancer, which originates from remnants of the notochord. These tumor cells not only share histological features with notochord cells, but also the expression of characteristic genes. As Brachyury has long been known for its role in notochord formation, it is now investigated as a possible cause of chordoma. One possible way for Brachyury to promote cancer is through its ability to activate Snail. Snail is a transcriptional repressor that is involved in EMT and was found to be regulated by Brachyury in various species (Nibu et al., 2013).

1.5.3. Mesp1

Mesp1 was proposed to be a key regulator of cardiovascular progenitors in vertebrates and represents the earliest marker for cardiovascular progenitors. Mesp1 and Mesp2 seem to be essential for early cardiac mesoderm formation and multipotent cardiovascular progenitor migration (Bondue and Blanpain, 2010), although they seem to be able to substitute for each other during early stages of gastrulation (Bondue et al., 2011). In the absence of both Mesp1 and Mesp2, no mesoderm is formed and cardiac specification is not possible (Saga et al., 2000). Therefore it is not surprising that Mesp1 directly regulates the expression of many TFs involved in cardiovascular specification (Bondue

et. al., 2011), although the exact mechanisms by which Mesp1 regulates cardiac development are still widely unclear (Bondue et. al., 2008).

1.5.3.1. Mesp1 in Early Development

Mesp1 was first found in 1996 in the posterior part of the mesoderm at E7 to E7.5. Later it was found to be first expressed in the nascent mesoderm at the beginning of gastrulation at E6.5 in the early mesodermal cells that enter the primitive streak. It is rapidly downregulated when these cells exit the primitive streak and are incorporated into the heart field and the head mesenchyme (Saga et al., 1996, 1999). After E8.5, Mesp1 is expressed in the presomitic mesoderm (Saga et al., 1999). Furthermore, it was shown that Mesp1 expressing cells later also contribute to the development of the endocardium (Kitajima et al., 2006; Saga et al., 2000). Mesp1 is expressed in the precursor of MCPs of both heart fields. The data collected so far suggest that almost all cardiac cells are derived from cells that at one point expressed Mesp1 (Bondue and Blanpain, 2010). Furthermore, it is believed that Mesp1 may play a role in the development of some muscles and bones of the face (Harel et al., 2009; McBratney-Owen et al., 2008; Yoshida et al., 2008), as well as in the generation of certain cells of the embryonic liver and hematopoietic stem cells (Asahina et al., 2009).

Mesp1 loss during mouse development leads to a defect in migration of the cardiac progenitor cells, resulting in cardiac malformation (Saga et al., 1999). Experiments showed that transient expression of Mesp1 promotes and accelerates cardiac differentiation, while continuous expression using a Dox-inducible system has an inhibitory effect (Bondue et al., 2008). It was also shown that Mesp1 not only promotes cardiac differentiation, but also the generation of endothelial and smooth muscle cells during ESC differentiation (Bondue and Blanpain, 2010). Forced expression of Mesp1 in fibroblasts was not sufficient to induce cardiac cells (Ieda et al., 2010), which suggests that Mesp1 needs other factors, only present in primitive streak like cells, to promote cardiovascular differentiation (Bondue and Blanpain, 2010).

1.5.3.2. Genetic and Proteomic Features of Mesp1

In the mouse, the Mesp1 gene is located in the mid region of chromosome 7 and encodes a TF of the b-HLH family (Saga et al., 1996). Like other members of the b-HLH TF family, Mesp1 was shown to bind to the classical E-box element with the sequence CAC(G/A)TG (Warren et al. 2011; Chan et al. 2013), more generally CANNTG (Nguyen et al., 2003).

1.5.3.3. Transcriptional Regulation of Mesp1

Two possible enhancer elements called P1-EME (Mesp1 early mesoderm enhancer) and P1-PSME (Mesp1 presomitic mesoderm enhancer) were located between 3,1 and 4,4 kb upstream of the transcription initiation site of the Mesp1 gene. It is however believed that there are other not yet known regulatory elements controlling Mesp1 expression (Oginuma et al., 2008). Van den Ameele et al. also found a DNA segment they called T3 (-3.8 kb) in this highly conserved P1-EME region, which was occupied by Eomesodermin. Furthermore, they also showed that Eomesodermin can bind another DNA segment on the Mesp1 gene called T1 (+0.1 kb). A third element called T2 (-2.4 kb) containing another putative Eomesodermin binding site was not occupied (van den Ameele et al., 2012). These data, as well as experiments done by other groups, suggest that the T-box TF Eomesodermin is a direct activator of Mesp1 (van den Ameele et al., 2012; Costello et al., 2011). The experiments of David et al. revealed that Mesp1 positive populations arose from Brachyury positive cells and they showed that a T responsive element (-2.4 bk) in the 3.4 kb proximal Mesp1 promoter is bound and activated by Brachyury. Furthermore Mesp1 mRNA expression was highly induced in Brachyury overexpressing cells, while Eomesodermin overexpression did not have this effect, suggesting that Eomesodermin is not a direct activator of Mesp1 (David et al., 2011).

Recently a Tcf/Lef-Oct4 site (-31 to -17 relative to the transcription initiation site), with a sequence identical to the already established Sox2-Oct4 composite site, was identified in the Mesp1 promoter. This site promoted activation of the Mesp1 promoter by Oct4 and canonical Wnt signaling. Loss of this interaction impaired Mesp1 expression and the cardiac program. It was proposed that this site helps maintain pluripotency through Sox2-Oct4 and induce lineage related genes through Oct4's recruitment of Tcf/Lef (Li et al., 2013).

Mesp1 has a biphasic effect on its own expression: First Mesp1 promotes its own expression, followed by a long lasting repression of its mRNA expression. It was shown that Mesp1 directly binds its own promoter. Moreover, Mesp1 also regulates the expression of its closest homolog Mesp2. The exact mechanisms by which Mesp1 achieves these goals are still widely unclear (Bondue et al., 2008).

1.5.3.4. Mesp1 Targets

Mesp1 is thought to activate TFs like Tbx5, Nkx2.5, Hand2, Gata4, Mef2c and Myocardin and repress other genes that regulate early steps of primitive streak formation

and early endoderm cell fate specification, like Brachyury, Fgf8 and Sox17 by binding conserved basic helix-loop-helix binding sites within their genomic regions (Bondue and Blanpain, 2010; Bondue et al., 2008). It was shown that Mesp1 binds an E-box in the described cardiac enhancer of the Nkx2.5 gene. Myocardin, a transcriptional coactivator that regulates SRF, is believed to be one of the key downstream effectors of Mesp1 by mediating cardiac and smooth muscle cell fate specification, as the loss of Myocardin function blocks the promoting effect of Mesp1 (Bondue and Blanpain, 2010). Furthermore, it was shown that Mesp1 can promote the expression of other cardiac structural genes like Myh6, Myl1, Myl2, Myl7 and Tnn2, which are involved in cardiomyocyte differentiation. Mesp1 also regulates a variety of TFs involved in epithelial-mesenchymal transition (EMT) like Snail1, Twist1, FoxC1/2 and Zeb1/2, which are responsible for the downregulation of E-cadherin and the upregulation of mesenchymal markers (Lindsley et al., 2008).

1.5.4. Nkx2.5

The murine Nkx2.5 gene encodes a homeobox TF and is located on chromosome 17. Nkx2.5 has been shown to function in murine heart formation and development. It is one of the earliest expressed cardiac genes and a well-established marker for the cardiomyocyte lineage (Warren et al., 2011).

It is evolutionarily highly conserved and homologs or related genes were found in all species with a heart, including the Drosophila TF tinman. It is part of the regulatory network responsible for cardiac and visceral mesoderm specification (Tanaka et al., 1999).

1.5.4.1. Nkx2.5 in Early Development

In the mouse Nkx2.5 expression can first be detected in the precardiac mesoderm and the adjacent pharyngeal endoderm at E7.5. By E8.5 it is expressed homogeneously in the myocardium of the atria and the ventricles and its expression is maintained throughout development and in the adult heart (Tanaka et al., 1999). In EBs, Nkx2.5 positive cells could be isolated as early as day 5 of differentiation (Kim et al., 2009).

An Nkx2.5 knock out is lethal in mice and results in developmental arrest in the initial stage of looping around E10.5. Experiments with a tamoxifen-inducible Nkx2.5 gene knockout demonstrated that even induced deletions at E12.5 and E19 lead to premature death, whereas deletion after terminal differentiation of cardiomyocytes (beginning at 2 weeks of age) leads to moderate cardiac contractile defects (Warren et al., 2011). In

humans, Nkx2.5 mutations are associated with congenital heart disease (Kasahara et al., 2001).

1.5.4.2. Genetic and Proteomic Features of Nkx2.5

The Nkx2.5 protein contains a highly conserved 60-amino-acid homeodomain, which interacts with DNA through a helix-turn-helix DNA-binding motif, and two other conserved regions, whose functions are still unknown (Reamon-Buettner et. al., 2004). The homeodomain is also essential for dimerization (Kasahara et al., 2001). Nkx2.5 can bind DNA as a monomer (Kasahara et al., 2001), but also as either a homodimer or a heterodimer with other TFs like Tbx5 (Bruneau et al., 2001), SRF (Chen et al., 1996) or Zac1 (Yuasa et al., 2010). Nkx2.5 has been shown to bind the consensus sequence 5'-TNAAGTG-3', with the core binding sequence being 5'-AAG-3', but is also thought to bind yet unknown nonconsensus sequences (Warren et al., 2011). It was also demonstrated that Nkx2.5 can bind the short core sequence 5'-TAAT-3' with reduced affinity (Chen and Schwartz, 1995).

1.5.4.3. Transcriptional Regulation of Nkx2.5

Searcy et al. found that 3 kb of the 5' flanking sequence is sufficient to activate Nkx2.5 expression in the cardiac crescent at E7.25 and they identified a 505 bp regulatory element, from -3059 to -2554 upstream of the Nkx2.5 gene, containing multiple TF binding sites (Searcy et al., 1998).

Lien et al. were able to identify a cardiac-specific enhancer that "fully recapitulates the expression of Nkx2.5 from the onset of cardiogenesis in the cardiac crescent and heart tube through the stage of looping" (Lien et al., 1999). Later in cardiac development, this particular enhancer element is active only in the right ventricular chamber, suggesting other modular regulatory elements controlling Nkx2.5 expression in the left ventricle and atrial chamber. The essential region of this enhancer spans from -9435 to -9277 upstream of Nkx2.5 and contains an essential high-affinity Gata-binding site. Though Gata is known to activate the cardiac-specific enhancer, it cannot be sufficient and other yet unknown TFs are needed, as Gata factors are expressed homogeneously in the heart after E11.5, when activity of the Nkx2.5 enhancer becomes restricted to the right ventricle (Lien et al., 1999). Another enhancer was detected -3400 to -4800 upstream of the gene. Furthermore, various negative regulatory elements have been described between -8922 and -8039 bp (Lien et al., 1999).

Tanaka et al. found three TATA-less promoters that produce together with alternative splicing at least three different RNA transcripts in three-week old mice. The data suggest that the regulation of the Nkx2.5 gene is a highly complex process, ensuring expression in anatomically distinct areas of the heart as well as time specific regulation during different developmental stages. Furthermore, the gene seems to require many "cis-acting elements spread in a large area of the genome to recapitulate the homogeneous pattern of expression of this gene throughout development" (Tanaka et al., 1999). The cause for this may be that during evolution new regulatory modules developed to achieve more complexity (Tanaka et al., 1999).

Nkx2.5 transcription itself is activated through major upstream pathways, like the BMP pathway (Zhang and Bradley, 1996). Positive autoregulation of a gene's promoter is a way to ensure tissue-specificity. There are multiple NKE elements in the Nkx2.5 promoter, so a positive feedback loop could be expected. Surprisingly, evidence for the existence of a negative feedback loop was found, which supports the conclusion that Nkx2.5 levels have to be tightly controlled to allow normal cardiac development. In the developing stomach, Nkx2.5 seems to be regulated by a positive feedback loop, which leads to the conclusion that the manner in which Nkx2.5 regulates its own expression is diverse for different tissues (Tanaka et al., 1999).

1.5.4.4. Nkx2.5 Targets

Among the known targets of Nkx2.5 is the ANF gene (Warren et al., 2011). Nkx2.5 together with Gata4 can activate ANF transcription by binding to an adjacent site in the promoter (Lien et al., 1999). It was also demonstrated that Nkx2.5 can interact with Gata4 through its HD domain (Kasahara et al., 2001), and its binding to a distal upstream enhancer is necessary for the activation of the Gata6 promoter in the developing heart (Lien et al., 1999). Furthermore, Nkx2.5 can interact with Nkx2.3 and Nkx2.6 and is believed to also be able to interact with other NK2 class proteins (Kasahara et al., 2001).

1.5.4.5. Nkx2.5 Post-translational Modifications

Although relatively little is known about post-translational modifications of the Nkx2.5 protein and possible effects of such, it was recently shown that excessive Nkx2.5 O-GlcNAcylation leads to Nkx2.5 downregulation in the heart tissue of diabetic mice and may play a role in diabetic cardiomyopathy. Furthermore, it may interfere with cardiomyocyte survival pathways. In another study it was revealed that excessive O-GlcNac conditions suppress the differentiation of Nkx2.5-lineage cardiomyocyte

precursor cells from ESCs (Kim et al., 2012). It was also shown that under normal conditions, O-GlcNAcylation decreases in differentiating EBs when compared to undifferentiated ESCs (Kim et al., 2009).

Moreover, Serin 163 of the Nkx2.5 protein was shown to be phosphorylated by casein kinase II, which plays a role in nuclear localization and transcriptional activity (Kim et al., 2009).

1.5.5. Desmin

The Desmin gene is located on Chromosome 1 of the mouse genome. Desmin is a muscle-specific type III intermediate filament (IF), which plays a role in dilated cardiomyopathy, myofibrillar myopathies and muscle wasting. It is one of the earliest markers of muscle differentiation (Li and Capetanaki, 1993). Desmin is mainly located to the Z-disk of striated muscles and the dense bodies of smooth muscle cells and it is essential for the structural integrity and function of muscles (Winter et al., 2013). Desmin is expressed in all muscle tissues, cardiac, skeletal and smooth (Capetanaki et al., 1997) as well as in endothelial cells (Costa et al., 2004).

1.5.5.1. Desmin in Early Mouse Development

Desmin is expressed at low levels in myoblasts and its expression is greatly increased at the onset of differentiation. In embryonic development, Desmin can first be detected at E8.25 in the neuroectoderm, where it is transiently expressed with Vimentin and Keratin. It can then be found in the heart rudiment at E8.5 and in the myotome at E9. Furthermore, Desmin is expressed at low levels in satellite cells and replicating myoblasts (Li and Capetanaki, 1993). Further studies indicate that the Desmin gene is activated at least half a day before the protein can be detected (Capetanaki et al., 1997).

Desmin is expressed in muscle tissue together with Myf5 and before MyoD (Li and Capetanaki, 1993). Compared to other muscle cell types, cardiac cells have the highest proportion of Desmin and a large number of mitochondria. In all muscle types, Desmin is also concentrated around the nucleus and has been found to remain bound to the nuclear fraction (Costa et al., 2004).

It was demonstrated that Desmin is absolutely necessary for in vitro muscle differentiation in embryoid bodies. Surprisingly, many Desmin knockout mice were viable and fertile, even though they showed a multisystem disorder involving cardiac, skeletal and smooth muscles with the most severe effects found in the heart. It is

suspected that Vimentin, Nestin or Synemin were able to compensate the loss of Desmin to some extent (Capetanaki et al., 1997).

1.5.5.2. Genetic and Proteomic Features of Desmin

The Desmin protein was first purified in 1977 and the Desmin gene was characterized in 1989. Since then several isoforms have been described (Costa et al., 2004).

There is a single Desmin gene that encodes a 53-kDa protein, which is built of an arginine-rich head domain, a rod domain with two α-helical coiled coils and a tail domain (Winter et al., 2013). Desmin can polymerize into 10 nm filaments due to its rod domain: The α-helical coiled coils allow the formation of parallel dimers, which further interact and form antiparallel tetramers called protofilaments (Geisler et al. 1992; Winter et al. 2013). Eight of those protofilaments can associate and form a unit-length filament, and several unit-length filaments can form the mature Desmin intermediate filament (Herrmann et al., 1999; Kiss et al., 2011).

1.5.5.3. Transcriptional Regulation of Desmin

Like Nkx2.5 and several downstream muscle structural genes, Desmin is controlled by modular regulatory elements (Schwartz and Olson, 1999).

Capetanaki et al. discovered several potential cis-acting regulatory elements in the 5'UTR of the Desmin gene: three E boxes (E1, E2 and E3 at -79, -832 and -936), one Mef2 binding site (at -864) and a region with homology to an M-CAT motif (at -587), as well as several GC boxes upstream and downstream of E1. Further analysis revealed that the first 81 bp upstream of the transcription initiation site, containing E1 and a TATA box at -36, were enough to induce muscle specific expression and are considered to be the Desmin promoter. The highest expression level was achieved by a construct containing up to 897 bp (Li and Capetanaki, 1993). The region between -798 and -976 was described to behave like a classical enhancer, which is responsible for muscle-specific high Desmin expression levels (Li and Capetanakil, 1994). It was shown that both MyoD and Myogenin bind the E1 and E2 boxes of the Desmin promoter and enhancer respectively. A negative element was identified between -81 and -578 (Li and Capetanaki, 1993). It was demonstrated that binding of Mef2 to the identified binding site is crucial for embryonic development and that Mef2 cooperation with mHLH regulators is necessary for normal transcription of Desmin during embryonic skeletal muscle development (Kuisk et al., 1996).

1.5.5.4. Desmin Targets

Desmin can interact with many other filaments and proteins, including Vimentin (Herrmann and Aebi, 2000), Synemin (Titeux et al., 2001; Xue et al., 2004), Nestin (Sjöberg et al., 1994), Paranemin (Breckler and Lazarides, 1982; Schweitzer et al., 2001), Syncoilin (McCullagh et al., 2007), Lamin B (Georgatos et al., 1987), Spectrin (Langley and Cohen, 1986), Plectin (Konieczny et al., 2008), Titin (Fürst et al., 1989), Desmoplakin (Getsios et al., 2004; Lapouge et al., 2006), Calponin (Mabuchi et al., 1997), and Myospryn (Kouloumenta et al., 2007). Many of these interactions are needed for Desmin's interaction with cellular structures and organelles, such as the nucleus, the myofibrils, mitochondria and the sarcolemma (Winter et al., 2013). Desmin can also form heteropolymers with IF proteins of the same class, like Vimentin (Costa et al., 2004).

Moreover, Desmin could play a role in the regulation of gene expression. It has also been shown to be able to interact with DNA through a DNA binding domain conserved in all type III filaments (Wang et al., 2001). Furthermore it is believed that Desmin can bind Lamins that are located in the nucleus (Cartaud et al., 1995; Tolstonog et al., 2002). It has also been shown to be present in the nucleus and form dimers with MyoD, thus possibly regulating gene expression (Li et al., 1994). It is suspected that Desmin and MyoD also regulate the expression of each other's genes (Wang et al., 2001). It has been described for other proteins (e.g. actin, microtubule) that there are different populations, depending on their location, with different properties and functions. The same could be true for Desmin, with the subsarcolemmal Desmin having a somewhat different function than the nuclear Desmin. It also stands to reason that Desmin may have slightly different roles in skeletal, cardiac and smooth muscle cells (Costa et al., 2004).

1.5.5.5. Desmin Post-translational Modifications

Desmin was shown to be the target of various posttranslational modifications (PTMs), including phosphorylation, ADP ribosylation, ubiquitination, glycation, oxidation and nitration. All of those modifications, except for ubiquitination, induce the disassembly of Desmin intermediate filaments with varying efficiency. Phosphorylation and ADP ribosylation most likely act through the addition of negatively charged phosphate groups to the positively charged arginine residues in the head domain. Some phosphorylations are not enough to disassemble Desmin completely and may produce altered binding properties that change the interaction of Desmin with other proteins. It was also shown that actions involving cAMP and PKA are important in myogenesis and involve

phosphorylation of Desmin and Vimentin and that those two intermediate filaments are phosphorylated during all stages of myogenesis. This modification seems to be necessary for myoblast fusion (Winter et al., 2013).

1.5.5.6. Desmin and Disease

In many myopathies and cardiomyopathies, abnormally accumulated granular and filamentous aggregates of Desmin have been described, but the causative role of these aggregates is still widely unclear (Capetanaki et al., 1997). Symptoms of Desmin-related myopathies include skeletal muscle weakness, arrhythmias and heart failure. The before mentioned accumulation of Desmin can be caused by mutations in Desmin or the chaperone αB-crystallin. Overexpression experiments demonstrated that a strong overexpression resulted in accumulation of Desmin and the formation of aggregates, as seen in human desminopathy patients. It is still not fully understood if this accumulation is the cause or the consequence of the disease (Costa et al., 2004).

In many Desmin-related myopathies, Desmin seems to be hyperphosphorylated due to PKC overexpression. This may lead to loss of the filamentous structure, which in turn affects myofibril integrity and the localization of the organelles. Furthermore, the disassembled Desmin forms aggregates that lead to atrophy via the ubiquitin-proteasome system. Although many mutations in the Desmin gene that lead to the loss of a PTM site have been linked to myopathy, it is still mostly unclear if the loss of the PTM site is the cause for the disease or if the mutation alters the structure of the protein in a way that causes disease (Winter et al., 2013).

1.6. Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) is a technique that makes it possible to map protein-DNA interactions in vivo (Vinckevicius and Chakravarti, 2012). ChIP experiments yield complex data sets that provide a snapshot of varying chromatin states and protein binding events taking place in millions of cells that are influenced by genetic and environmental factors (Furey, 2013). In the classical approach, DNA and proteins are transiently crosslinked, the chromatin is sheared and the protein of interest is precipitated with an antibody. The last steps are the reversion of the crosslink and amplification of the pulled down DNA by PCR with primer pairs specific for a certain region in the genome. In the case of the protein interacting with the specific genomic region, more DNA is

pulled down and amplified, which results in a higher signal (Vinckevicius and Chakravarti, 2012).

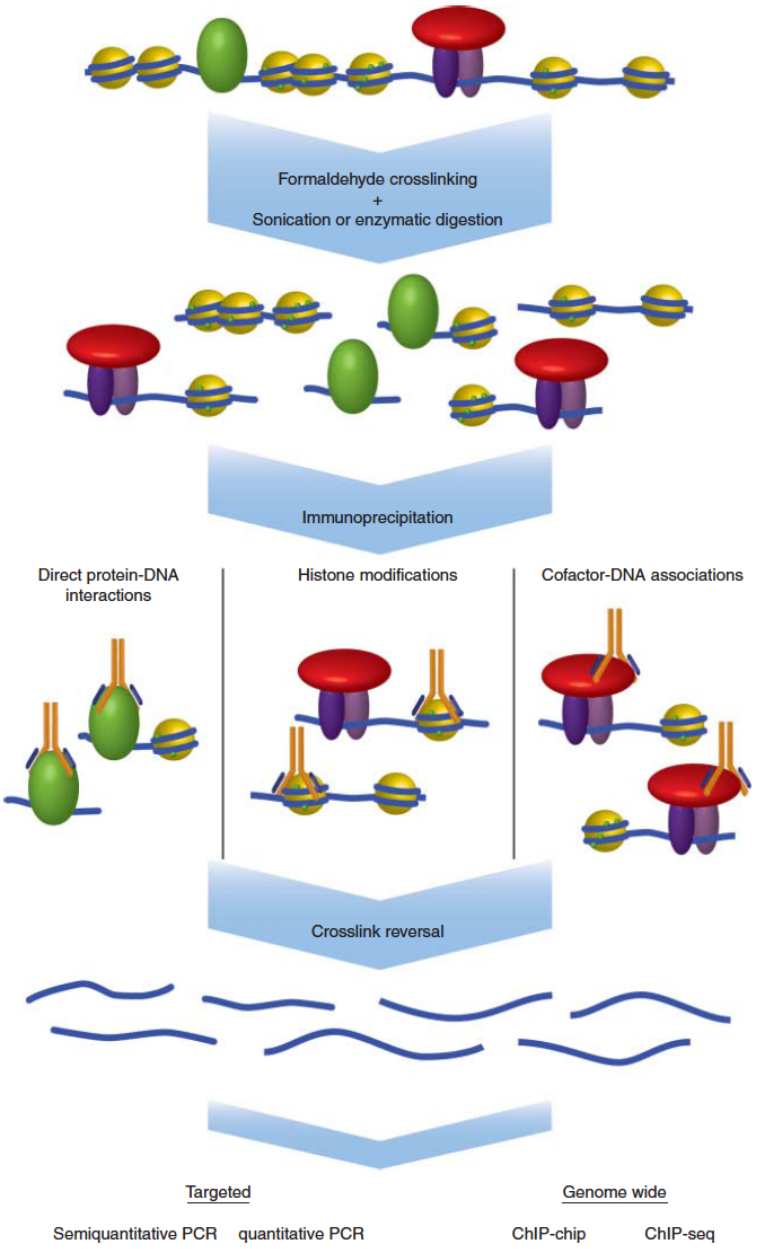


Figure 1.5. Essential steps in chromatin immunoprecipitation. DNA and proteins are cross-linked and the chromatin is sheared into pieces that are several hundred base pairs in length. Associated DNA fragments are isolated with an antibody against the protein of interest, the crosslinking is reversed and the DNA can be analyzed with a number of analysis techniques (Vinckevicius and Chakravarti, 2012).

In recent years, more and more adaptions to the protocol have been developed in order to examine interaction patterns on a genome wide scale, the best known being ChIP-chip (microarrays) and ChIP-seq (deep sequencing) (Vinckevicius and Chakravarti, 2012).

What makes Chromatin immunoprecipitation such a valuable tool is the fact that it is possible to monitor dynamic interactions over time in vivo, which has greatly enhanced the understanding of transcriptional initiation as well as maintenance (Vinckevicius and Chakravarti, 2012). In the living cell DNA binding proteins are essential for various major cellular processes such as DNA transcription, splicing, replication and repair. Proteins involved in these processes can be TFs directly binding certain DNA sequences as well as histone proteins that are part of nucleosomes. Advances in ChIP methodology are constantly expanding the number, types and resolution of protein-DNA interactions discovered (Furey, 2013).

For a successful ChIP experiment, a highly specific antibody to the bound protein or modification is essential. Recent studies within the ENCODE and modENCODE projects testing over 200 antibodies have shown that antibody quality varies greatly, even between independently prepared lots of the same antibody. The studies demonstrated that 25% of tested antibodies failed specificity tests and 20% failed immunoprecipitation experiments (Egelhofer et al., 2011).

Advances have been made in decreasing the amount of cells needed for a successful experiment as well as in increasing the precision of the obtained binding information (Furey, 2013). As DNA fragments produced by sonication are a few hundred base pairs long, but protein binding sequences are typically only 6-20 bp, 5' end digestion of the crosslinked DNA with lamda exonuclease (ChIP-exo) results in a 90-fold greater precision. It also largely eliminates contaminating DNA not bound by the target factor (Rhee and Pugh, 2011).

In order to determine if multiple proteins bind to the same locus simultaneously or if they are present on different chromosomes in the same cell or in different cells, sequential ChIP with different antibodies in successive experiments can be performed (Markham et al., 2007).

Despite all those advances in ChIP methodology, several challenges still remain. For example, the detection of TF binding to highly repetitive sequences is still problematic with all ChIP protocols and further improvements are needed to obtain reliable data in these regions (Chung et al., 2011; Furey, 2013). Another issue that cannot be resolved with ChIP alone is the interpretation of signal strength that is influenced by the strength

of the interaction, which can depend on variations in genotype and on the percentage of cells in the experiment that have the binding event. Furthermore, signals also reflect indirect binding where one protein interacts with another protein that is bound to DNA.
Competition ChIP assays have made it possible to elucidate the difference between stable binding and turnover. It has been proposed that sites stably bound by the same factor (resident sites) are associated with efficient transcriptional regulation, whereas sites with a high turnover (treadmilling sites) are associated with lower transcriptional efficiency (Furey, 2013).

1.7. Working Hypothesis

In recent years a lot of research has focused on the transcriptional network regulating events during early cardiogenesis, but there are still many aspects that are not fully understood yet.
In our experiments we concentrated on the TFs Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin. Nanog is mainly known for its role in maintaining pluripotency in undifferentiated stem cells (Saunders et al., 2013). Brachyury and Mesp1 are two of the earliest markers for the mesodermal cell lineage and are both expressed soon after the differentiation process starts (Bondue and Blanpain, 2010; Evans et al., 2012). Nkx2.5 is first expressed soon after Mesp1 and is a well-established marker for CVPCs (Tanaka et al., 1999). Desmin is believed to also act as a TF regulating gene expression during cardiogenesis, in addition to its function as an intermediate filament in muscle cells (Costa et al., 2004). Due to the fact that (1.) these five TFs are expressed in close succession during differentiation and (2.) are important for either the maintenance of stem cells or the regulation of early cardiogenesis, we assumed that there is a good chance that they also regulate each other's gene expression in CVPCs. There were already some data available describing the interaction of some of the investigated TFs. Previous work of our group has shown that Desmin is not only present in the nucleus but also occupies the regulatory DNA segments of the Nkx2.5 gene. Furthermore, it was shown by other groups that Nanog and Brachyury regulate each other's gene expression at least indirectly (Suzuki et al., 2006). Brachyury was also shown to activate Mesp1 gene expression, and a negative regulatory feedback loop was described between those two TFs (David et al., 2011). Mesp1 was also shown to regulate the Nkx2.5 gene by binding the described cardiac enhancer, but the exact mechanisms have not been described so far (Bondue and Blanpain, 2010).

We used the EB and CB model system to simulate early processes during embryogenesis and cardiogenesis, respectively, in vitro. We performed chromatin immunoprecipitation at carefully selected time points during differentiation to investigate the physical interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with their respective genes. We investigate the physical interaction at already described regulatory DNA segments, as well as at not yet described DNA segments, where an accumulation of putative TF binding sites was found. We assumed that these DNA segments were most likely to contain new regulatory elements.

We hope that our data will help to further elucidate the transcriptional network regulating events during early cardiogenesis. This understanding is a prerequisite for the development of new strategies to efficiently differentiate somatic cardiac stem cells into fully functional cardiomyocytes and therefore find new ways of treating cardiovascular disease.

2. Material

2.1. General Chemicals for Molecular Biology and Material

Acetic Acid	AppliChem, D
Acrylamid	Merck, D
Agarose	Sigma Aldrich, USA
β-Mercaptoethanol	AppliChem, D
BCIP	Sigma Aldrich, USA
Bis N,N-Methylene-bis-acrylamide	Biorad, USA
Bovine Serum Albumin	Roth, D
Bovine Serum Albumin Fraction V	PAA, A
Bromphenolblue	Sigma Aldrich, USA
Coomassie Brilliant Blue G250	Fluka, CH
Dimethylformamid	Merck, D
Dithiothreitol (DTT)	Acros, B
Ethanol	Sigma Aldrich, USA
Ethidiumbromide	Fluka, CH
EDTA	Acros, B
EGTA	Acros, B
Formaldehyde	Merck, D
Glycerin	Merck, D
Glycine	AppliChem, D
Glycogen	Carl Roth, D
HEPES	AppliChem, D
Hydrochloric Acid (32%)	AppliChem, D
Instant Milk	Maresi, A
Lithium chloride	AppliChem, D
Magnesium Chloride Hexahydrate	Acros, B
Methanol	Merck, D
NBT	Sigma Aldrich, USA
Nitrocellulose Membrane 0.2µm	BioRad, USA
Nonidet P 40 Substitute (NP-40)	Fluka, D

Phenylmethansulfonylfluoride (PMSF)	Fluka, CH
2-Propanol	Carl Roth, D
Protease inhibitor conplete mini	Roche, CH
SDS	BioRad, USA
Sodium acetate (NaOAc)	Merck, D
Sodium dhloride	Sigma Aldrich, USA
Sodium deoxycholate (NaDoc)	AppliChem, D
Temed	AppliChem, D
Tris Ultrapure	AppliChem, D
Triton X-100	AppliChem, D

2.2. Chemicals for Tissue Culture

β-Mercaptoethanol	Loba, A
Ampothericin B	Invitrogen, USA
D (+)-Glucose Anhydrous	AppliChem, D
DMEM powder	Gibco, USA
DMEM	Gibco, USA
DMSO (Dimethylsulfoxide)	Sigma Aldrich, USA
Fetal Bovine Serum	HyClone, USA
Fetal Bovine Serum	Gibco, USA
Fetal Bovine Serum	Sigma Aldrich, USA
Gelatine	Difco, USA
L-(+)-Glutamin	Acros, B
Mitomycin C	AG Scientific, USA
Natriumhydrogencarbonate	Life Technologies, USA
Penicillin/ Streptomycin	Gibco, USA
Potassium Chloride	Sigma Aldrich, USA
Potassiumhydrogenphosphate	Fluka, CH
Sodiumhydrogencarbonate	Life Technologies, USA
Streptomycin	Sigma Aldrich, USA
Trypsin	Life Technologies, USA

2.3. Enzymes

PCR Master-Mix Gold S	Peqlab, D
ProteinaseK	AppliChem, D
Rnase	Sigma Aldrich, USA

2.4. Antibodies

Anti-Brachyury (C-19); polyclonal; sc-17745	Santa Cruz Biotechnologiy, USA
Anti-Desmin; polyclonal; D8281	Sigma Aldrich, USA
Anti Mesp1(T15); polyclonal; sc163078	Santa Cruz Biotechnology, USA
Anti-Nanog (M-149); polyclonal; sc-17745	Santa Cruz Biotechnology, USA
Anti-Nkx2.5 (H-114); polyclonal; sc-14033	Santa Cruz Biotechnology, USA
Donkey Anti-Goat IgG, AP-Conjugate	Promega, USA
Donkey Anti-Rabbit IgG, AP-Conjugate	Promega, USA
Protein A-Sepharose^{TM} CL-48	GE Healthcare, USA

2.5. Recombinant Proteins

Mouse LIF	Miltenyi Biotec GmbH, D

2.6. Size Markers

Gene Ruler 1 kb Plus DNA Ladder	Thermo Fisher Scientific, USA
Gene Ruler 100 bp DNA Ladder	Thermo Fisher Scientific, USA
Fast Ruler DNA Ladder Low Range	Thermo Fisher Scientific, USA
Unstained Protein MW Marker	Thermo Fisher Scientific, USA

2.7. PCR Primers for Chromatin Immunoprecipitation

The product length of primers for ChIP should be between 100 and 400 base pairs. Too short products are hard to distinguish from the primers themselves. Too distant primer pairs may not be able to both bind to the same sonicated DNA fragment, which should be between 400 and 600 base pairs long.

Table 2.1. Primer pairs used for PCR analysis of ChIP samples.

Gene	Primer Name	Sequence (5' to 3')	Product Size	Tm (°C)
Nanog	E1 f	TCACGTCGGACTGCTTCTTC	153 bp	65.6
	E1 r	GTCCTCTCTTTGGTCCCCT		62.1
	D0 f	CCTACATACCCTGGCTGGTG	167 bp	64.3
	D0 r	AGGGGCAGCTATTGTGTTAGG		64.1
	D1 f	ACTCCAAGGCTAGCGATTCA	196 bp	63.8
	D1 r	AATAGGGAGGAGGGCGTCTA		63.7
	D2 f	GTTTTGACTGCTAACCACCCAGAG	143 bp	66.7
	D2 r	GGCAGGCTTGCTACATTCCTTATC		67.5
	D3 f	CCGGCTTAGAGCTTGAACCA	117 bp	66.4
	D3 r	TCCCAAGGGCGACGTAATTT		67.4
	D4 f	GGAGATCAGGCTGGCTTTCA	147 bp	66.7
	D4 r	CATGCCTGAGGAAGTCAGAGG		66.2
	P0 f	AGATGTTACGAGAGGACGGC	168 bp	63.2
	P0 r	TGCTCTAGCTGGTCCCAAAC		64.3
	P1 f	AATGAGGTAAAGCCTCTTTTTGG	123 bp	63.2
	P1 r	ACCATGGACATTGTAATGCAAA		63.8
	I1 f	TGGTAGAACCAAGAGGCTGC	392 bp	64.3
	I1 r	CAGCACTTTTCCCACCTCCA		67.4
	I2 f	AATGAACGGGTGGGGATAGC	140 bp	66.9
	I2 r	TAGACAGCCCTGAAAGCAGC		64.5
	I3 f	TGTCTAAACAGGACTGTGCATT	274 bp	63.8
	I3 r	TTTGCATGGTGCTGGGTGAAT		69.6
	I4 f	CACCAGCACCATGCAAAA	105 bp	68.2
	I4 r	CTGCTCTCTGGGTCCTACCT		62.4
	I5 f	GAGGTAGGACCCAGAGAGCA	252 bp	63.4
	I5 r	AGAACCGTGACCGCATCTAC		64.1
	I6 f	TGAAGAGGGAGCAGGGAGAA	137 bp	66.3
	I6 r	ACAATTAGACCCTGGGAATCCT		63.5
	U1 f	TGGCTGCTGTGTCAGAGATG	282 bp	65.5
	U1 r	TGGGGAGGCCTCTTGAACTA		65.9
	U2 f	AAAGTGGGGAAAGACAGGCT	109 bp	63.8
	U2 r	AGGATTCATCCCTCCCTCCC		67.5
	U3 f	ACACAGGCTTTGGAGTGTCT	107 bp	61.4
	U3 r	ATCTCCCCCTAAGCCCCAG		66.3

	U0 f	GCCATCTGTATAGAGGCAACA	252 bp	61.7
	U0 r	CGCCCACTGAGAATGAGTGT		65.4
		Annealing temperature used for all Nanog primers: 52°C		
Brachyury	NK2 f	TCAGCCAAGCCCATCACTTT	132 bp	66.9
	NK2 r	GCTCCTTAGGCCCTCTCTCT		62.9
	D0 f	CCTTGCTTTCTTGCCCACAC	130 bp	66.9
	D0 r	AAGCACAGCCTGTCTCCTTG		64.6
	D1 f	TCTCCTCTGGGGTGTGTGAA	101 bp	65.9
	D1 r	GCACAGCCCTCACCATAGAA		65.2
	T Box f	GGGGACTGGAGAGGAGTTCA	899 bp	65.6
	T Box r	AAGCAGGGCAAAGAGACAAG		63.4
	D2 f	CTGTTCAGGTACCATCACACATAA	103 bp	62.5
	D2 r	GGATAGTTCAAAAGACATCAGCAGG		65.7
	D3 f	GGGCTGAAGGCTAATGGAGA	182 bp	65.4
	D3 r	CCCCTAGTTTCTGCCATGCT		65.2
	P0 f	TAAAACCCCGGCCACCAAAG	227 bp	69.6
	P0 r	CCCTGCTCCACTTCCTGTTT		65.4
	P1 f	AGGAAAGGGGGCTTTGTTGT	126 bp	65.7
	P1 r	AGGCAACATTCCCACCAAGA		66.6
	P2 f	TAGTGACCCAGCCTCCCTTG	146 bp	66.4
	P2 r	GCCAGTTTGTGTTGTGGGTT		64.8
	NK1 f	TGCTCTTTGTCACCTTCCCC	217 bp	66.4
	NK1 r	GTCCTCGCGTCCCACATAAA		67.1
	P3 f	GAGGTGCAAACATTTGGGGG	242 bp	68.5
	P3 r	AGAAAAATGGGTCCCCTCCC		67.1
		Annealing temperature used for all Brachyury primers: 54°C		
Mesp1	3'UTR f	TTCCTCCAGGTGCTGATTGG	382 bp	59.7
	3'UTR r	TCCCACCACTTCTGTCTCCA		60.1
	T1 f	ATTGGGTCCGCTGTAGCTTT	164 bp	59.7
	T1 r	AGCAGACTGGACTAAGGGGT		59.9
	P1 f	ACCTGTTAGGTTCATAGGCTGC	362 bp	60.9
	P1 r	GCCAGGCTCACAACTGC		58.6
	T2 f	CAGCAGACGAAGGATGGGAG	222 bp	60.2
	T2 r	ACCACCAGAAACCATCACCC		59.9
	T3 f	TTCACACTTTTGGGGCCTGT	149 bp	60.0
	T3 r	AGGCTTGACCTCTCACCTCT		59.9
	D1 f	CTCTTCCCCACCAACCTTCC	351 bp	59.9

	D1 r	TCAACCAGCATCAGACCCAC		59.9
	NK f	ACCAGTTGGTCTTGGTGGTG	301 bp	60.1
	NK r	AGCTTCCTTCAAACAGCGGA		59.9
	D2 f	CCAACAGCACCCTTTCACAC	176 bp	59.6
	D2 r	ATGTGCTCAAGCTGGGTAGG		59.7
	D3 f	CGTGTCAACAGGCAACCAGA	210 bp	60.6
	D3 r	GGTCTGGCTTCTCTGCCTTC		60.4
		Annealing temperature used for all Mesp1 primers: 58°C		
Nkx2.5	Prom1 f	CTCTGTTTGCTTTCTCGCCA	246 bp	58.8
	Prom1 r	ATTGGAGACAGGCAGCGTTA		59.4
	Prom2 f	GGCGTTGTTGAAGATAAAGC	258 bp	55.7
	Prom2 r	GCCTTTTAAAGACTTGGTGC		55.5
	E1 f	GAGATAAGATGACATACCAGA	352 bp	51.6
	E1 r	ATTAGTGTGAACACAACACT		52.7
	E2 f	TTATCTTCCCCGGAGGAAAT	330 bp	55.6
	E2 r	GCAGACACTCAAGCCCTTAA		57.8
	P1 f	GTCCCAGAACAACTGTCT	262 bp	53.7
	P1 r	CCAGCCAGACACACTCAA		56.8
	P2 f	CTCACTCTGTCCTGTTTCCT	373 bp	56.5
	P2 r	GAGTCCTGTTAAGTGAATCGG		56.1
	D1 f	AGATTCCCAGCAGTCCGTA	332 bp	57.7
	D1 r	CAAGAGTCACTGCCCGAA		56.9
	D2 f	GACATTGTGCATCTTAGCA	422 bp	53.4
	D2 r	TTCCACACTGCTGACCTG		56.8
	MCE f	ACACAATGTCACACGCCAAG	316 bp	59.3
	MCE r	AGATACAAGGCCACACAGCA		59.3
		Annealing temperature used for all Nkx2.5 primers: 51°C (except Nkx2.5-MCE: 55°C)		
Desmin	Prom f	TCTTCTGTCCTCTTGGGGCT	212 bp	65.2
	Prom r	CGGCTGGAGTGGATGTGAAG		68.1
	NK1 f	AGTCGTTGGAAACCTCAGCA	224 bp	64.8
	NK1 r	GGGTTTCACAGGGCAAGGAG		68.0
	E1 f	ACAACCCGTGGTGTGAGTG	202 bp	66.0
	E1 r	GGGAGGAATGTTGGCCTTCT		66.4
	E2 f	TGACTGAAGCTTGTCGCTGT	153 bp	63.9
	E2 r	TGCCAAGGCGAAATACCAGA		67.6
	NK2 f	TGCTATCGGAGAACTGGGGA	216 bp	66.9

	NK2 r	GAGAAAGCGGTGCCAAAAGG		67.9
	D2 f	TGGGATCCTTGCTTGCTCTG	342 bp	67.7
	D2 r	GCTGCCTCCCTTTGACTCTT		64.7
	NK3 f	ATGCTAAGGTGAGCGTGGG	200 bp	65.2
	NK3 r	GTGTGCCGATGAAGGGAGAG		67.1
	D1 f	TTGGTCTTGGTGGGCCTTTG	248 bp	69.2
	D1 r	CCTCAATGTCTTCCCTGGTGT		65.2
	D0 f	ACACGCACACACACTCACA	200 bp	62.9
	D0 r	CCCCAAACCAAAACTAAAGTGG		65.1
				Annealing temperature Desmin primers: 55°C

2.8. Cell Lines

2.8.1. Fibroblasts

The fibroblast cell line SNL 76/7, established by Alan Bradley, was obtained from the murine STO fibroblast cell line and has an additionally inserted LIF gene and a neomycin-resistance gene in the genome (McMahon and Bradley, 1990).

2.8.2. Embryonic Stem Cells

The W4 wild type cell line was isolated from the mouse strain C3H by Georg Weitzer (Lauss et al., 2005).

2.8.3. Cardiovascular Progenitor Cells

Of the twelve cardiovascular progenitor cell lines, which were isolated by Wolfgang Weber and Matthias Scheinast, only the A5 cell line was used here. The cells were isolated from the hearts of new-born mice, which carried a neomycin resistance gene in one allele of the hdac1 locus (Hoebaus et al., 2013).

2.9. Tissue culture

2.9.1. Buffer and Solutions for Cell Culture

10x PBS (Phosphate Buffered Saline) stock solution

- 80 g NaCl
- 2 g KCl
- 10.72 g $Na_2HPO_4x7H_2O$
- 2 g KH_2PO_4
- The ingredients are dissolved in 800 ml MilliQ water and the pH is adjusted to 7.2 with saturated Na_2HPO_4. The solution is filled to 1 liter and sterile filtered (Nalagene Filter, 0.22 µm pore width).

100x GPS (Glutamine-Penicillin-Streptomycin)

- 4.25 g NaCl
- 1.5 g Penicillin
- 2.5 g Streptomycin
- 14.6 g L-(+)-Glutamine
- MilliQ water is added to a final volume of 500 ml and the solution is aliquoted into 50 ml falcons, stored at -20°C and then kept at 4°C after thawing.

100x β-Mercaptoethanol (10^{-2}M)

- 200 mL 1x PBS
- 144 µL β-Mercaptoethanol (14 M)
- The solution is sterile filtered and aliquoted into 50 ml falcons, stored at -20°C and then kept at 4°C after thawing.

Trypsin

- 3.5 g NaCl
- 0.5 g D-Glucose
- 0.09 g $Na_2HPO_4.7H_2O$
- 0.185 g KCl
- 0.12 g KH_2PO_4
- 0.2 g EDTA
- 1.25 g Trypsin (Gibco)
- 1.5 g Tris Base
- MilliQ water is added to a final volume of 500 ml and the final pH of 7.6 is adjusted with concentrated HCl. The solution is aliquoted into 50 ml falcons, stored at -20°C and then kept at 4°C after thawing.

1% Gelatine Stock Solution

- 10 g Gelatine (Difco)
- MilliQ water is added to a final volume of 1 litre and the solution is sterile filtered.

DMEM (Dulbecco's Modified Eagle's Medium)

- Half a can of DMEM (Gibco, +4500 mg/l Glucose, -$NaHCO_3$, -Pyruvate #52100-039) is dissolved in 4.5 l MilliQ water in a 5 l Erlenmayer flask. Next, 18.5 g $NaHCO_3$ are added. The pH is adjusted to 7.4 with concentrated HCl and the medium is filled up to a final volume of 5 l with MilliQ water.
- The medium is sterile filtered into cell culture flasks. Before usage all bottles should be checked for possible contamination. Therefore, a small aliquot of each bottle is incubated at 37°C for at least 24 hours. If a contamination can be detected under the microscope the affected bottle has to be sterile filtered again.

Freezing Medium (2x stock)

- 60% DMEM
- 20% Fetal Bovine Serum (brand depends on cell line)
- 20% Dimethylsulfoxide (DMSO) – add last and drop wise!

2.9.2. Growth Media

M15Hy - Medium for ESCs and CVPCs

- 83 % DMEM
- 15% Fetal Bovine Serum (HyClone)
- 1% GPS (Glutamine-Penicillin-Streptomycin)
- 1% β-Mercaptoethanol

M15Si - Medium for Embryoid and Cardiac Bodies

- 83% DMEM
- 15% Fetal Bovine Serum (Sigma)
- 1% GPS (Glutamine-Penicillin-Streptomycin)
- 1% β-Mercaptoethanol

M10Gi - Medium for Fibroblasts

- 89% DMEM
- 10% Fetal Bovine Serum (Gibco)
- 1% GPS

3. Methods

3.1. Tissue Culture

Soap is fatal for embryonic stem cells (ESCs) and cardiovascular progenitor cells (CVPCs). Therefore, all glassware has to be cleaned separately from normal tissue culture materials.

3.1.1. Cleaning of Pipettes

First, the pipettes are left in water with hypochlorite for at least 10 min to kill all cells left on them, then they are rinsed with tab water for about three hours. After that, the pipettes are transferred into a container with MilliQ water and left to soak over night. The next day, they are dried at 60°C for at least 4 hours or until they are completely dry. Now the pipettes can be stuffed with cotton wool and be autoclaved to make them sterile once again.

3.1.2. Cleaning of Bottles

Glass bottles are treated similarly to pipettes. They are filled with water and a shot of hypochlorite, and are left for about 10 to 20 min before being rinsed with tab water for about 10 to 15 min. After that, they are rinsed with MilliQ water once and are filled up with MilliQ and left over night. They are air dried and autoclaved before further usage.

3.1.3. Cultivation of ESCs and CVPCs

ESCs and CVPCs are cultivated in 24-well plates on a layer of feeder cells and fed with 1.5 ml of M15Hy medium daily. After sucking the old medium off gently, the new medium has to be applied drop by drop so as not to detach the cells from the plate. The feeder cells are mitotically inactivated fibroblasts excreting leukemia inhibitory factor (LIF), which is necessary for the cells to remain in an undifferentiated state. The plates are incubated at 37°C and at a CO_2 concentration of 5%.

3.1.4. Splitting of ESCs and CVPCs

Ideally the cells should always be split 1:3. Two hours before the actual splitting, the cells are pre fed with 0.5 ml and the new feeder wells are pre fed with 2 ml M15Hy medium. After two hours, the medium of the ESCs or CVPCs has to be completely aspirated and the cells are washed once with 1xPBS to get rid of the remaining medium. If the cells are

cultivated on a 24- well plate, 200 µl trypsin are added and they are incubated for 20 min at 37°C. This time frame is ideal to detach the ESCs and CVPCs without putting them through too much stress. The trypsin is then inactivated with 1 ml medium from the pre fed feeder wells and the suspension is resuspended 18 to 20 times with a 1000 µl pipette to separate the cells. When splitting 1:2 600 µl of the resuspended cells are transferred to the pre fed feeder well, when splitting 1:3 400 µl are transferred. After the procedure, the cells need at least 24 hours to reattach to the dish before they can be fed the next day.

3.1.5. Splitting of Fibroblasts

In contrast to ECSs and CVPCs, fibroblasts do not have to be pre fed before splitting. The medium is aspirated completely and the 10 cm plate is washed with 1xPBS before adding 1 ml trypsin. The plates are incubated for only 5 to 7 min, because fibroblasts are more sensitive towards prolonged exposure to trypsin and need less time to detach from the plate. While waiting, new plates are prepared and enough medium is added so that the end volume will be 8 ml. The trypsin is then inactivated with at least twice the amount of M10Gi medium as previously added trypsin. If the cells are for example split 1:6 on 6 new plates, it is best to add 5 ml medium. The cells are then resuspended thoroughly and 1 ml cell suspension is added to each new plate. Fibroblasts only need a couple hours to attach to the new cell culture dish.

3.1.6. Production of Feeder Cells

First, the fibroblasts have to be mitotically inactivated through addition of 80 µl Mitomycin C per 4 ml medium (the medium is lifted off the cells with a pipette and everything but 4 ml is discarded). Before continuing with the cells after 4 hours, 24-well plates have to be coated with 0.1 % gelatine solution for at least 2 hours.

After 4 hours the cells are trypsinized (see 3.1.5) and 4 ml M10Gi medium are added. All of the cells are then transferred into a 50 ml falcon tube and centrifuged at 1000 rpm for 7 min. The supernatant is discarded and the pellet is resuspended thoroughly before the cell number can be counted. The cell concentration is then brought to $35x10^4$ cells/ml medium and 0.5 ml suspension are transferred onto each well of a 24-well plate.

3.1.7. Production of Embryoid and Cardiac Bodies

Embryoid bodies (EBs) are generated from ESCs, while cardiac bodies (CBs) consist of CVPCs, but the general procedure is the same.

The day before the production of EBs or CBs the cells have to be split 1:2 for the cells to be in optimal condition the next day, because the procedure is rather stressful for the cells. For about 20 dishes of EBs, 4 wells of a 24-well plate are needed.

24 hours after splitting, the cells can be pre fed with 1 ml M15Si medium and are incubated for 2 hours. Next, the cells are trypsinized and transferred onto a gelatinized 10 cm cell culture dish, where the feeders are supposed to attach to the plate while the other cells remain in suspension. In the meantime the plates for the actual EBs can be prepared: because a smooth surface is needed, the lids of microbiology dishes are used. To prevent the EBs or CBs from drying out, the bottom of each plate is covered in sterile MilliQ water. After one hour the cell suspension is transferred into a 50 ml falcon tube and the plate is rinsed once with 5 ml medium. The falcon tube is centrifuged at 1000 rpm for 7 min to get rid of the trypsin and other unwanted factors in the medium. The supernatant is then discarded and the cells are resuspended thoroughly in M15Si medium (amount depends on the expected cell number). After pipetting the suspension 20 to 30 times, the cells should be resuspended thoroughly enough to be counted in a cell counter (10 µl cell suspension per chamber). The cells are then brought to a final concentration of $9x10^4$ cells/ml and counted again.

On each lid about 80 to 100 20 µl drops are placed in regular intervals and the lid is carefully put back on the dish. The finished plates should be left in the hood for a little while before being transferred into the incubator, as the vibrations of the hood are essential for the aggregation process of the cells. The EBs and CBs should be aggregated and ready to be transferred to gelatinized cell culture dishes on day 4.5 of differentiation.

3.1.8. Freezing of Cells

Each well of a 24-well plate which is to be frozen is treated with 200 µl trypsin for 20 min. During the incubation period, the freezing medium is prepared (see 2.8.1). The trypsin is inactivated with 800 µl M15Hy medium and the cells are resuspended 18 to 20 times with a 1000 µl pipette. First, 200 µl freezing medium are added drop wise and the plate is moved crosswise about 5 times after each drop to avoid too much damage to the cells. Next, 800 µl freezing medium are added in the same fashion. This procedure is optional as the 2x freezing medium can be simply mixed with the cell-type specific growth medium (1:1) and added to the cells. In each well there are now 2 ml that are distributed to 2 sterile cryo-tubes. The cells have to be frozen slowly to maximize the

survival rate: therefore, they are put in a tightly closed styrofoam box in the -80°C freezer for 24 to 48 hours before being transferred to liquid nitrogen for long term storage.

3.1.9. Thawing of Cells

The cells should be thawed quickly in a 37°C water bath after being taken out of the liquid nitrogen tank. The tubes can also be thawed in the hand but caution is advised as they can explode. The 1 ml cell suspension is transferred into a 15 ml falcon tube. First, 200 µl M15Hy medium are added drop wise and the tube is shaken gently for about 5 seconds after each drop. Next, 800 µl are added in the same fashion, then 1 ml and finally 5 ml. The suspension is centrifuged at 1000 rpm for 7 min and the supernatant is discarded. 1 ml M15Hy from pre fed feeder wells is used to resuspend the pellet and the cell suspension is put in one well of a 24-well plate with a feeder layer.

3.2. Molecular Biology Methods

3.2.1. Chromatin Immunoprecipitation (ChIP)

3.2.1.1. ChIP Stock Solutions and Buffer

ChIP Stock Solutions	Amount	pH adjustment	Solvent	Volume
• 1M Tris pH 8.1	12.1 g	HCl	ddH_2O	100 ml
• 1M Tris pH 6.5	12.1 g	HCl	ddH_2O	100 ml
• 1M Tris pH 7.5	12.1 g	HCl	ddH_2O	100 ml
• 1M HEPES pH 7.9	23.8 g	2M NaOH	ddH_2O	100 ml
• 0.5M EGTA pH 7.5	19.0 g	2M NaOH	ddH_2O	100 ml
• 0.5M EDTA pH 7.2	18.6 g	2M NaOH	ddH_2O	100 ml
• 3M NaOAc pH 5.2	24.6 g	HCl	ddH_2O	100 ml
• 10% TritonX-100	20 ml		ddH_2O	200 ml
• 10% NP-40	20 ml		ddH_2O	200 ml
• 10% SDS	20 ml		ddH_2O	200 ml
• 1M LiCl	16.96 g		ddH_2O	400 ml
• 1M Tris pH 8.1	48.4 g		ddH_2O	400 ml
• 4M NaCl	23.36 g		ddH_2O	400 ml
• 10% NaDoc	10 ml		ddH_2O	100 ml

Wash- Buffer I	Wash Buffer II
• 10 ml 10% TritonX-100 • 8 ml 0.5M EDTA pH 7.5 • 400 µl 0.5M EGTA pH 7.5 • 4 ml 1M HEPES pH 7.9 • ddH_2O ad 400 ml	• 20 ml 4M NaCl • 800 µl 0.5M EDTA pH 7.5 • 400 ml 0.5M EGTA pH 7.5 • 4 ml 1M HEPES pH 7.9 • ddH_2O ad 400 ml
Dilution Buffer	**RIPA Buffer**
• 400 µl 10% SDS • 44 ml 10% TritonX-100 • 960 µl 0.5M EDTA pH 7.5 • 6.68 ml 1M Tris pH 8.1 • 16.7 ml 4M NaCl • ddH_2O ad 400ml	• 15 ml 4M NaCl • 20 ml 1M Tris pH 8.1 • 4 ml 10% SDS • 20 ml 10% NaDoc • 40 ml 10% NP-40 • ddH_2O ad 400 ml
Hi-Salt Buffer	**LiCl Buffer**
• 50 ml 4M NaCl • 20 ml 1M Tris pH 8 • 4 ml 10% SDS • 40 ml 10% NP-40 • ddH_2O ad 400 ml	• 100 ml 1M LiCl • 20 ml 1M Tris pH 8 • 20 ml 10% NaDoc • 40 ml 10% NP-40 • ddH_2O ad 400 ml
Elution Buffer – make fresh	**TE Buffer**
• 1 ml 10% SDS • 0.5 ml 1M $NaHCO_3$ • 50 µl 1M DTT • ddH_2O ad 5 ml	• 4 ml 1M Tris pH 8 • 800 µl 0.5M EDTA • ddH_2O ad 400 ml

Lysis Buffer (desired vol.) – make fresh

- 10% SDS
- 10mM EDTA pH 7.5
- 50mM Tris pH 8
- Proteinase Inhibitor (PI); 1 tablet in 1 ml MilliQ= 50x stock
- 0.1 mM PMSF; use 1000x stock; solved in isopropanol

3.2.1.2. ChIP Procedure

Best results are expected when about 7 million cells are used per ChIP, but the procedure can also be done with varying cell numbers. For the following procedure all cells should be in one 10 cm cell culture dish with 10 ml medium.

Day 1

In order to crosslink proteins and DNA, 270 µl formaldehyde per 10 ml medium are added to the cells in the fume hood and the cells are incubated for 20 min at room temperature (RT). To stop the crosslinking, 100 µl 2,5M Glycine are added and the cell culture dish is shaken for 5 min horizontally.

The medium is then aspirated and the cells are washed twice with ice-cold 1xPBS, before harvesting them in ice-cold 1xPBS with a cell scraper. The suspension is transferred into a 15 ml falcon tube and centrifuged at 1200 rpm at 4°C for 5 min.

The supernatant is discarded and the pellet is resuspended in 1 ml WASH I buffer (+ 200 µl PI 50x stock and 10 µl PMSF 1000x stock per 10 ml). Another 4 ml WASH I are added and the cells are incubated for 10 min on ice. The suspension is centrifuged again at 1200 rpm at 4°C for 5 min. The supernatant is again discarded and the washing step is repeated in the same way as before with WASH II buffer (+ 200 µl PI and 10 µl PMSF per 10 ml). After another centrifugation step, the supernatant is discarded and the pellet is resuspended in about 600 µl fresh lysis buffer (depending on the cell number). The suspension is stored at 4°C and can be left there for a while.

Day 2

Before usage, the sonicater has to be thoroughly cleaned to avoid contamination of the samples. Best use soap, 70% Ethanol and NaOH repeatedly before rinsing with MilliQ water. The standard conditions for shearing the chromatin of the cells are as follows: 90% Dot Cycle, 45% output and 30 seconds on ice (not longer as SDS will fall out), which is repeated 5 times. It should be considered that DNA from undifferentiated ESCs and CVPCs is more easily and efficiently sonicated than DNA from embryoid or cardiac bodies. The lysate is centrifuged at full speed (14 krpm) at 4°C for 10 min to get rid of unwanted membrane proteins. The cleared supernatant is then transferred into a new eppi.

Input sample preparation

The protein concentration in each input sample should amount to approximately 50 µg per 100 µl. The OD_{280} of a diluted lysate sample (1:50) is measured with a NanoDrop photometer to determine the amount of protein in the lysate. If the measured OD_{280} is 0.1

mg/ml, 100 µl of sonicated lysate are used for the input sample. If the OD varies, the amount of used lysate has to be adjusted, but must not exceed 140 µl. After the different input samples are brought to an equal volume with lysis buffer, the final volume is expanded to 400 µl with freshly prepared elution buffer. To reverse the crosslinking in the input samples, 20 µl 4M NaCl are added and the samples are incubated shaking at 300 rpm at 65°C on the heat block over night.

Day 3

The samples for the "input" are taken from the heat block and the proteins are degraded by Proteinase K. Therefore 8 µl 0.5M EDTA, 16 µl 1M Tris pH 6.5 and 2 µl Proteinase K (20 mg/ml) are added to the input samples and the samples are shaken at 500 rpm at 55°C for one hour. Next, the DNA is retrieved by phenol-chloroform-isopropanol (PCI 25:24:1) extraction. The three components are mixed in the fume hood and 600 µl are added to each sample, which is then vortexed and centrifuged at full speed (14 krpm) at 4°C for 10 min. After the centrifugation, the upper aqueous phase containing the DNA can be transferred into a new eppi. For further DNA precipitation 800 µl -20°C 96% Ethanol, 12 µl 3M NaOAc and 1 µl glycogen (20 mg/ml) are added and the samples are incubated at -20°C for 30 min. The samples are then centrifuged at 14 krpm at 4°C for 10 min. The pellet is washed with 500 µl -20°C 70% Ethanol and is again centrifuged at 14 krpm at 4°C for 10 min. The supernatant is discarded and the pellets are air dried at RT. When there are no more liquid drops left in the eppi, the pellet is dissolved in 200 µl autoclaved MilliQ water. 1 µl RNase is added and the eppis are left on the heat block shaking at 1200 rpm at 37°C for 30 min. The input samples can then be stored at -20°C until needed.

Before the actual immunoprecipitation the length of the sonicated chromatin fragments has to be checked by loading 20 µl of the input samples on a 0.8% agarose gel. Additionally 20 µl of the lysate sample can be loaded, but it has to be considered that the DNA fragments in the lysate samples are still cross-linked to proteins and, therefore, produce signals at a different position than the input samples. Ideally the majority of fragments should be between 400 and 600 bp long. If they are too long it is likely get false positive signals later on, if they are too short it is hard to get any signal at all with PCR.

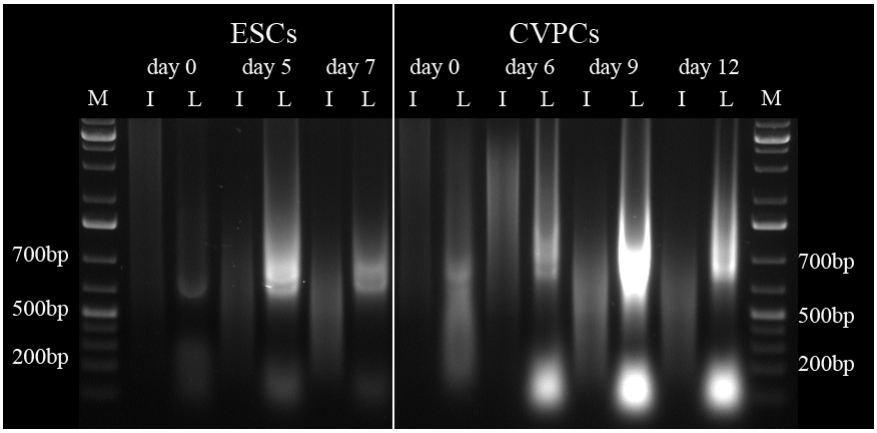


Figure 3.1. Size of the DNA fragments, which were separated on an agarose gel, in the input and the lysate samples after sonication (derived from ESCs on days 0, 5 and 7 of differentiation, and from CVPCs on days 0, 6, 9 and 12 of differentiation). I, input sample; L, sonicated lysate sample before immunoprecipitation; M, size marker

Actual immunoprecipitation

For each immunoprecipitation the same amount of sonicated sample as for the input sample is used. If the before measured OD_{280} of the 1:50 diluted lysate is 0.1 mg/ml, 100 µl are used. Again the samples are brought to equal volumes with lysis buffer (max. volume: 140 µl). Then the sample is diluted 1:10 with dilution buffer (+PI, +PMSF). To pre-clear the chromatin, about 20 µl of blocked protein-A bead suspension is added and the samples are left shaking at 4°C for 1 hour.

To remove the unspecifically loaded beads, the samples are centrifuged at 1200 rpm at 4°C for 5 min. The supernatant is then transferred into a new eppi and 4 µg of the specific antibody per immunoprecipitation are added. The samples are incubated shaking at 4°C over night.

Day 4

To harvest the antibody-protein-DNA complexes, 30 µl blocked protein A beads are added and the samples are shaken for at least 1 hour at 4°C (max. 4 hours). The samples are then centrifuged at 1200 rpm at 4°C for 5 min and the supernatant is discarded, as the protein of interest connected to the specific antibody is now bound to the beads. The beads are now washed extensively in several washing steps in the 4°C laboratory to get rid of unspecific binding: during each washing step 900 µl buffer are added to the beads and they are left shaking for 10 min before centrifugation at 1200 rpm at 4°C for 5 min to discard the supernatant and wash the beads with the next buffer. Caution is advised as the

beads do not stick to the eppi like a pellet and get disturbed easily. These washing steps are performed with the following buffers: once with RIPA buffer and once with HI Salt buffer to get rid of unspecific binding, once with LiCl buffer to get rid of crosslinked RNA and twice with TE buffer to neutralize and to get rid of remaining salts.

To elute the desired DNA-protein-antibody complexes from the beads, 400 µl elution buffer are added and the samples are shaken at 1400 rpm on the heat block at RT for 30 min. The samples are then centrifuged at 3000 rpm at RT for 5 min, before the supernatant is transferred into a new eppi and the crosslinking is reversed by addition of 20 µl 4M NaCl and putting the samples on the heat block shaking at 300 rpm at 65°C over night.

Day 5

In order to get rid of the proteins and antibody in the sample, a proteinase K digestion is set up. Per immunoprecipitation 8 µl 0.5M EDTA, 16 µl 1M Tris pH 6.5 and 2 µl Proteinase K (20 mg/ml) are added and left shaking for 1 hour at 55°C. To retrieve the DNA, a phenol-chloroform-isopropanol (PCI) extraction is performed in the fume hood. Like before, 600 µl PCI are added per sample, the eppis are vortexed thoroughly and then centrifuged at 14 krpm at 4°C for 10 min. The upper aqueous phase containing the DNA is transferred into a new eppi and the lower phenol phase containing the unwanted proteins is discarded properly.

To further purify the DNA, 600 µl -20°C 96% ethanol, 12 µl 3M NaOAc and 1 µl glycogen (20 mg/ml) are added to the samples and they are incubated at -20°C for 30 to 60 min. Next, the samples are centrifuged at 14 krpm at 4°C for 20 min before the pellet is washed once with 500 µl -20°C 70% to rehydrate the DNA. To make sure the pellet is not lost, the samples are once more centrifuged at 14 krpm at 4°C for 10 min before the supernatant is discarded, the pellet is air dried and then dissolved in 200 µl autoclaved MilliQ water. In order to degrade unwanted RNA, 1 µl RNase is added and the immunoprecipitation samples are incubated on the heat block at 37°C for 30 min. Now the samples can be stored at -20°C until needed.

Blocking of protein-A or -G coupled sepharose beads

In general it can be expected that Protein-A sepharose beads are more suitable for polyclonal antibodies and Protein-G sepharose beads achieve better results with monoclonal antibodies.

The following procedure is dimensioned for about 0.1 g beads (about three times the tip of a small spatula). The beads are put into a 2 ml eppi and overlaid with 1 ml TE buffer. They are left to soak at RT for one hour before being centrifuged at 1000 rpm for 3 min. The supernatant is cautiously discarded and the beads are washed 3 times with 1 ml TE buffer in the same way.

In order to later avoid unspecific binding to the beads they have to be blocked. For 100 µl beads the following blocking solution is used: 2 µl herring or salmon sperm (10 mg/ml TE), 10 µl BSA (10 mg/ml), 5 µl 2% sodium azide solution and 84 µl TE buffer are mixed in an eppi. The beads are overlaid with one volume of blocking solution and left shaking at 300 rpm at 4°C for 1 hour. Before usage the beads are washed once more and overlaid with one volume TE buffer.

Storage of sepharose beads

After chromatin immunoprecipitation the beads have to be treated for storage. They are centrifuged at 1000 rpm for 3 min and the supernatant is cautiously discarded. The beads are washed at least once in the same way with TE buffer before being overlaid with one volume TE and left at 4°C until further usage.

3.2.1.3. ChIP PCRs

Due to the enormous amount of polymerase chain reactions (PCRs) necessary to get a statistically relevant result, PCR set ups of 20 µl instead of 50 µl are advisable. The DNA solution from the input and from the immunoprecipitation sample can be used in several dilutions (undiluted, 1:10 and 1:20) since the higher the sample is diluted, the more specific the signal gets. For each PCR reaction two master mix set-ups are prepared: one with the DNA of the immunoprecipitation and one with the DNA from the input sample. Both samples should always be used in the same dilution.

Set up for 20µl PCR reaction

- 5,6 µl ddH_2O
- 10 µl Peqlab Master Mix Gold S
- 0.4 µl DNA

For each 20 µl PCR reaction, 16 µl of the Master Mix with ddH_2O and DNA are transferred into a 0.5 ml eppi and 4 µl of a 1:10 diluted primer mix, containing forward and reverse primer, are added.

The following conditions are used for the PCR reactions:

Table 3.1. Conditions for PCR reactions

Step	Temperature (°C)	Time (seconds)
1	95	300
2	95	60
3	Primer specific (see Table 2.1)	60
4	72	60
5	72	300
6	4	pause

Steps 2 to 4 are repeated 39 times (altogether 40 cycles). The lid temperature should be 10°C higher than the highest reaction step, which means 105°C in this case. If needed, the PCRs can then be stored at 4°C or at -20°C.

To analyse the samples, 4 µl 6x loading dye are added to the 20 µl PCR reaction and the samples are separated on a 2% agarose gel, which is run at 80 V for about an hour (depending on the size of the gel). The gel is examined under UV light and a picture is taken.

6x Loading Dye

- 505 µl 88% Glycerin
- 600 µl 0.5 M EDTA
- 0.0022 g Bromphenolblue Fill up to 5 ml with dH_2O

3.2.1.4. Agarose Gels

As ethidium bromide (EtBr) is suspected to be a mutagen and therefore a carcinogen, everything in the labeled working area has to be handled with great care and nitril gloves have to be worn at all times.

To make a gel, first the appropriate amount of agarose has to be weighed out in an Erlenmeyer flask (Table 3.2). The agarose is mixed with 1xTAE buffer and the Erlenmeyer flask is heated in the microwave at full power until the agarose is completely dissolved. The gel is cooled down to about 60°C while constantly pivoting the flask to avoid unequal hardening. To make the DNA later visible under UV light, ethidium bromide (1:20) is used to stain the DNA. The gel is poured into the clean apparatus and all bubbles have to be removed with a clean pipette tip before putting the combs in. The

agarose has to harden completely before the combs can be taken out gently and the gel can be loaded. All equipment is washed with hot tab water and then rinsed with ddH_2O.

Table 3.2. Separation range of agarose gels.

agarose gel (%)	range of separation (bp)
0.5	1,000 - 30,000
0.7	800 - 12,000
1.0	500 - 10,000
1.2	400 - 7,000
1.4	200 - 4,000
2.0	50 - 2,000

3.2.2. Protein Extraction with NP-40 Lysis Buffer

Total volume of NP-40 Lysis Buffer = 50 ml

- 2 ml of 1M HEPES pH 7.5 (40mM)
- 1.2 ml of 5M NaCl (120mM)
- 0.1 ml of 0.5M EDTA pH 8.0 (1mM)
- 1 ml of 0.5M 2-glycerolphosphate (10mM)
- 0.5 ml of 0.05M Na_3VSO_4 (0.5mM)
- 5 ml of 0.5M NaF (50mM)
- 0.5 ml of NP-40 (1%)
- Mix buffer on a roller mixer, store at 4°C

5 µl of proteinase inhibitor (PI) and 5µl of PMSF are added to 490 µl of NP-40 lysis buffer. The cell pellet is resuspended in 2x the volume with NP-40 lysis buffer. The sample is incubated for 20 min on ice, afterwards it is centrifuged at 15000 rpm for 20 min at 4°C. Finally, the supernatant containing the proteins is transferred into a new eppi and stored at -80°C.

3.2.3. Western Blot

3.2.3.1. Solutions and Buffer for Western Blot

3x Sample Buffer	Acrylamide solution
• 3 ml Glycerin • 0,9 g SDS • 3,75 ml of Buffer B (pH 8,8) • 1,75 ml H2O • 6 mg Bromphenolblue	• 30 % Acrylamide • 0.8 % Bis N,N'-Bisacrylamid, fill up to 100 ml with dH_2O
Buffer B (pH 8,8)	**Buffer C (pH 6,8)**
• 36,3 g Tris (1,5M), adjust pH • 4 ml 20% SDS (0,4%), fill up to 200 ml	• 12,1 g Tris (0,5M), adjust pH • 4 ml 20% SDS (0,4%), fill up to 200 ml
10x Electrophoresisbuffer	**Harlow Buffer (Transferbuffer)**
• 60,6 g Tris • 292 g Glycine • 100 ml 20% SDS, fill up to 2 l	• 29 g Tris-Base • 145 g Glycine • 25 ml 20% SDS • 4 l dH_2O • 1 l Methanol • Needs to be stored at 4°C
10xTBS	**1xTBS/T**
• 175,32 g NaCl (1,5M) • 121,14 g Tris pH 7,4 (0,5M) • Adjust pH with 37% HCl; fill up to 2 l	• 100 ml 10xTBS • 1 ml Triton-X100 • Fill up to 1 l with dH_2O
Ponceau-S	**AP Buffer**
• 100 mg Ponceau-S • 100 ml 1% glacial acetic acid	• 12,1 g Tris (100mM) • 5,8 g NaCl (100mM) • 10,2 g MgCl2 Hexahydrat (50mM) • Adjust pH to 9,5 and fill up to 1 l
BCIP	**NBT**
• 50 mg BCIP • 1 ml Dimethylformamid	• 75 mg NBT • 700 µl Dimethylformamid • 300 µl H2O

Coomassie brilliant blue	Destain
• 45% (v/v) Methanol • 10% (v/v) Acetic acid • 0,4% (w/v) Coomassie brilliant blue	• 4,5% Methanol • 10% Acetic acid

3.2.3.2. SDS Page Gels

Table 3.3. Preparation of SDS Page gels.

Running gel						**Stack gel**	
	5%	**7%**	**10%**	**10,5%**	**11%**		**5%**
Total volume	12 ml	12 ml	12 ml	12 ml	12 ml	**Total volume**	3 ml
Solutions						**Solutions**	
Acryl Amide	2 ml	2,8 ml	4 ml	4,2 ml	4,4 ml	Acryl Amide	0,5 ml
Buffer B	3 ml	3 ml	3 ml	3 ml	3 ml	Buffer C	0,75 ml
H_2O	7 ml	6,2 ml	5 ml	4,8 ml	4,6 ml	H_2O	1,75 ml
10% APS	120 µl	120 µl	120 µl	120 µl	120 µl	10% APS	30 µl
TEMED	10 µl	10 µl	10 µl	10 µl	10 µl	TEMED	2,5 µl

3.2.3.3. Western Blotting

The running gel is prepared at a percentage dependent on protein size and it is covered with approximately 1 ml of isopropanol. After at least 20 min of polymerization the alcohol is washed away with dH_2O, the stack gel is added and the comb is put between the spacers.

In the meantime the samples are prepared for loading. The protein concentration is measured with NanoDrop and the amount of protein to be loaded is decided by keeping the abundance of the protein of interest in the cells in mind (minimum of 10 µg protein per sample). The same volume of 2x sample buffer with ß-mercaptoethanol (850 µl 2x sample buffer plus 150 µl ß-ME) is added. To denature all proteins, the samples and the marker have to be boiled at 95°C for 5 to 10 min.

Before the samples can be loaded onto the gel, the slots have to be washed with electrophoresis buffer in the apparatus to remove residual acrylamid. It is also helpful to make the slots visible with highly diluted sample buffer.

The SDS-Page is run at 60 V until a clear straight line is visible, afterwards 90 to 100 V can be applied. When the loading dye front has reached the last ¼ of the gel, the

nitrocellulose membrane needs to be incubated in Harlow buffer. A gel can be stained with coomassie blue to make the separated proteins visible, but the proteins of this gel can no longer be transferred onto a membrane. The gel is incubated in the coomassie brilliant blue staining solution for about 10 min on the shaker. When the gel is sufficiently stained, destain solution is added several times until the protein bands are clearly visible and the background of the gel is no longer blue. The coomassie brilliant blue solution can be used several times.

For blotting the proteins onto a nitrocellulose membrane the following assembly has to be made: 1x sponge, 2x Whatman Paper, SDS-Page gel, nitrocellulose membrane, 2x Whatman Paper, 1x sponge.

The blotting apparatus is filled up with Harlow buffer (4°C) and the cool pack is added. The transfer has to be run for one hour at 350 mA at either RT or 4°C.

Afterwards the membrane has to be shortly rinsed with dH_2O and is stained with Ponceau-S for at least 5-10 min to evaluate if equal amounts of protein were loaded and if proteins of all sizes were transferred. Then the membrane is washed 1-2 times for 5 min with TBS/T. The next step is to block the membrane for one hour at RT with a 5% milk-TBS/T blocking buffer to avoid unspecific binding. To remove residual blocking buffer, the membrane has to be washed 3 times with TBS/T for 10 min.

The primary antibody is diluted in a 5% BSA-TBS/T solution (dilution dependent on antibody; always check datasheet) and the membrane is incubated with it overnight at 4°C on a roller mixer in a 50 ml falcon tube (3 ml of antibody-dilution is sufficient and can be used several times).

The next day, the membrane is washed 3 times with TBS/T for 10 min and is incubated with the secondary antibody in a 5% milk-TBS/T buffer for approximately one hour. Again the membrane needs to be washed 3 times for about 10 min. The membrane has to be incubated in AP buffer for 5 min before 10 ml fresh AP buffer are added. First, 33µl NBT, then 66µl BCIP are added and the blot is incubated in the dark until bands are visible (at least 10 min). Removing the AP buffer and washing the membrane with water stop the reaction.

Secondary antibody	Dilution used
• Anti-goat AP-conjugated (Promega)	1:5000
• Anti-rabbit AP conjugated (Promega)	1:10000

4. Results

The goal of this thesis was to further elucidate certain aspects of the transcriptional network regulating early events in cardiomyogenesis.

In our experiments we used the A5 cell line (murine CVPCs) as well as the W4 cell line (murine ESCs) to generate cardiac and embryoid bodies as in vitro models for cardiac development. With chromatin immunoprecipitation we investigated the physical interactions of the transcription factors (TFs) Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with their respective genes at different time points. Nanog is known to be essential in maintaining pluripotency in stem cells, whereas Mesp1, Brachyury, Nkx2.5 and Desmin are believed to play an important role in early cardiomyogenesis.

In order to determine the binding pattern of these TFs to the respective genes, we created primer pairs covering DNA segments with an accumulation of putative TF binding sites for the five genes of interest. We assumed that these DNA segments were most likely to show strong physical interaction between the protein of interest and the DNA and therefore contain regulatory elements. Primer pairs were also generated for DNA segments where no TF binding was expected. These negative controls were meant to demonstrate the specificity of the ChIP experiments.

4.1. Protein Synthesis in ESCs and CVPCs During Differentiation/ Testing of Antibodies for ChIP

The specificity of the antibodies that were used in the ChIP experiments was tested by western blotting. This also presented an opportunity to investigate the occurrence of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin protein in ESCs and CVPCs during the differentiation process. Embryoid and cardiac bodies were generated and cells were harvested at specific time points: day 0, day 5 and day 18 in ESCs and day 0, day 6, day 9, day 12 and day 18 in CVPCs.

4.1.1. Nanog

Western blots with the anti-Nanog antibody produced very distinct signals at the correct position of 40 kDa, with only minimal variation in signal intensity up to and including day 18. As Nanog protein has a half life of only about two hours (MacArthur et al., 2012), these results suggest a continued Nanog expression throughout differentiation.

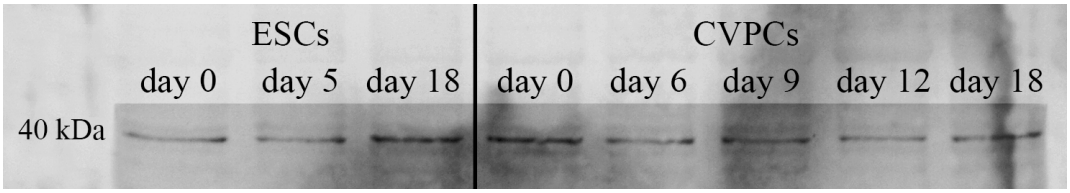


Figure 4.1. Western blot with anti-Nanog antibody (1:600 dilution of primary antibody). Cells were differentiated the depicted amount of days before lysis.

So far Nanog has been mostly known for its role in pluripotency. However, more and more evidence is being found that Nanog also plays an important role throughout differentiation (MacArthur et al., 2012).

4.1.2. Brachyury

The anti-Brachyury antibody also produced only one clear signal at the expected position of 49 kDa with no unspecific binding.

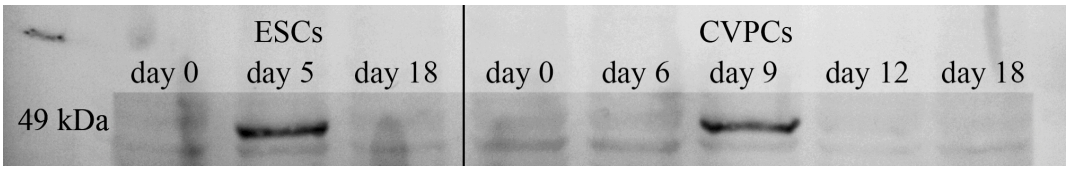


Figure 4.2. Western blot with anti-Brachyury antibody (1:600 dilution of primary antibody). Cells were differentiated the depicted amount of days before lysis.

It was not surprising that Brachyury protein gave a signal on day 5 of ESC differentiation as several other groups have found the same (Evans et al., 2012), but there were no data available concerning CVPCs. Surprisingly we found a distinct signal on day 9 of CVPC differentiation, with the signal intensity resembling that of ESCs on day 5 of differentiation. As CVPCs are believed to be several days ahead in development when compared to ESCs, it is possible that this peak of Brachyury protein expression on day 9 has a different purpose in developmental processes than the one found in ESCs.

4.1.3. Mesp1

The anti-Mesp1 antibody did not produce a signal on any of the investigated days in neither ESCs nor CVPCs. It is known that in ESCs, Mesp1 is expressed shortly after Brachyury with a peak that is limited to a short period between day 6.5 and 7.5 of differentiation (Bondue and Blanpain, 2010). Therefore, it is very likely that a western

blot with samples harvested on day 6 of ESC and day 10 of CVPC differentiation would produce a signal and that the samples we used simply did not contain enough protein to be detected.

4.1.4. Nkx2.5

The anti-Nkx2.5 antibody was expected to yield several signals between 30 and 50 kDa, but produced a completely different pattern above 116 kDa with very distinct and sharp signals.

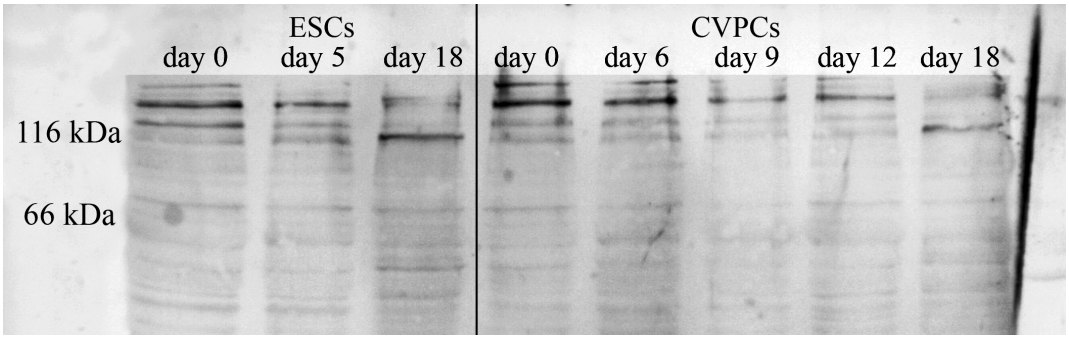


Figure 4.3. Western blot with anti-Nkx2.5 antibody (1:600 dilution of primary antibody). Cells were differentiated the depicted amount of days before lysis.

Nkx2.5 has been described to function as a dimer, as well as being modified posttranslationally by O-linked glycosylation that decreases during the course of differentiation (Kim et al., 2009). What we observed could be this change of glycosylation over time. In the Ponceau staining the area above 116 kDa hardly showed any protein on the nitrocellulose membrane, therefore it can be eliminated that areas with protein abundance were stained unspecifically. Additionally, it is possible that the lysis conditions were not strong enough to break up the Nkx2.5 dimer. The biggest discrepancy however is the strong signal in undifferentiated ESCs, as Nkx2.5 should not be present at this time, especially not in such abundance.

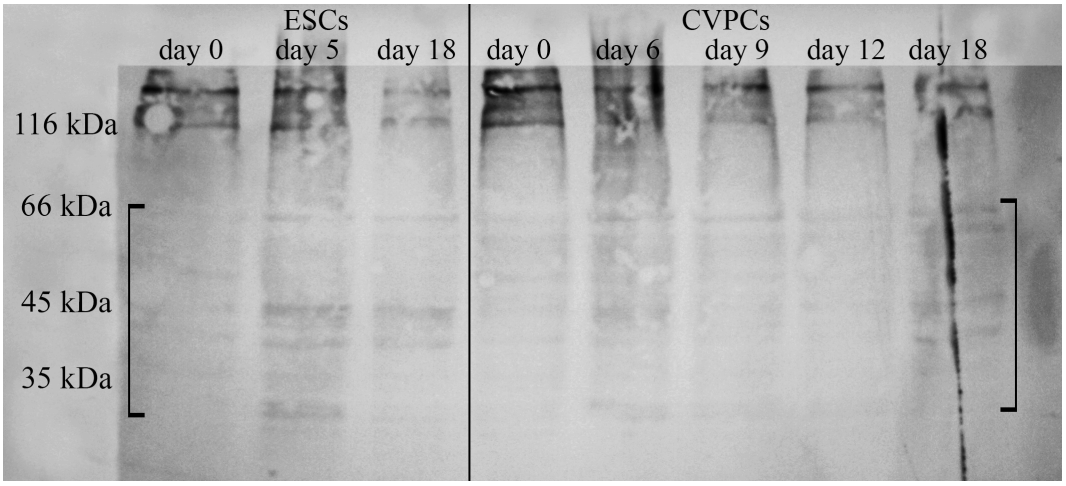


Figure 4.4. Western blot with anti-Nkx2.5 antibody (1:1000 dilution of primary antibody). Cells were differentiated the depicted amount of days before lysis. Black rectangles mark the position of the possible Nkx2.5 signals.

Repeating the experiment with a higher antibody dilution (1:1000) yielded similar results in respect to the signals above 116 kDa. There were, however, several signals between 30 and 70 kDa on day 5 of ESC differentiation that were not visible in undifferentiated ESCs. On day 18 of ESC differentiation and throughout CVPC differentiation, signals between 30 and 70 kDa could be observed, which showed varying intensity and distribution. It could be that those signals represent the unmodified Nkx2.5 protein.
Further experiments with different antibodies and lysis conditions would be necessary to draw a clearer conclusion.

4.1.5. Desmin

The anti-Desmin antibody produced a signal at the correct position of 55 kDa. In ESCs this signal was only visible on day 18 of differentiation and in CVPCs its intensity increased steadily over the course of differentiation.

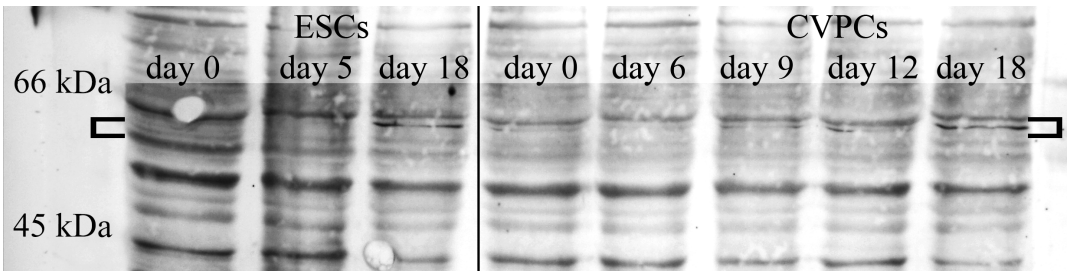


Figure 4.5. Western blot with anti-Desmin antibody (1:200 dilution of primary antibody). Cells were differentiated the depicted amount of days before lysis. Black rectangles mark the height of the possible Desmin signal.

Unfortunately western blots with the anti-Desmin antibody showed a high background that correlated with proteins visible during the Ponceau staining. A control experiment eliminated the possibility that the secondary antibody produced these unspecific signals. Therefore it can be assumed that the primary antibody stained every region of the membrane with large amounts of proteins. Nevertheless the signals at the correct height were much sharper than the background signals and not visible in the Ponceau staining. Furthermore, the same antibody produced nice immunofluorescence staining of Desmin filaments in immunofluorescence microscopy (diploma thesis of Christiane Fuchs and Philipp Heher). All this suggests that Desmin is present in differentiating ESCs and CVPCs and that the antibody detected Desmin in ChIP experiments, but that this antibody may not be ideally suited for western blot experiments.

4.2. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin Genes by Chromatin Immunoprecipitation

4.2.1. Generation and Analysis of the ChIP Data

For ChIP, undifferentiated cells were used, as well as cells from carefully selected time points in differentiation. Embryoid bodies derived from ESCs were harvested on days 5 and 7 of differentiation. Cardiac bodies derived from CVPCs were harvested on days 6, 9 and 12 of differentiation. The experiments with undifferentiated cells and cells from day 5 of ESC and day 6 of CVPC differentiation could be performed twice, but due to time limits ChIP with cells from the other time points could be done only once. Apart from small differences in protein binding pattern and ChIP signal intensity, the two independent experiments produced results that were comparable enough to summarize them in one illustration.

Semi-quantitative PCR was performed with the obtained ChIP samples of a specific time point. Each primer pair was used with an input sample as control (sample taken after sonication, but before immunoprecipitation) and with the antibody specific pull-down.

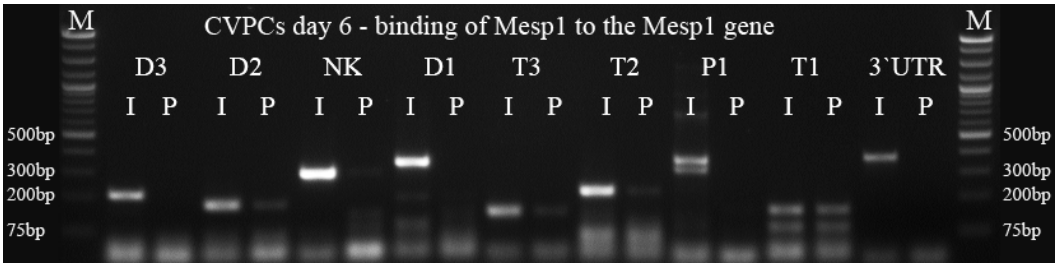


Figure 4.6. Typical PCR analysis of the binding of Mesp1 to the Mesp1 gene at the DNA segments D3, D2, NK, D1, T3, T2, P1, T1 and 3'UTR on day 6 of CVPC differentiation. I= Input sample; P= Pull-down sample; M= size marker

The PCR samples were then separated on an agarose gel and the intensity of the signals was analyzed with Adobe Photoshop CS5 by measuring the integrated density of the input (I) and the pull-down (P) signal, as well as a background (B) signal. Using the background subtraction method, the background signal was then subtracted from the input and from the pull-down signal, respectively. This method takes the variation of the gel background into consideration. Next, the intensity of the pull-down signal in relation to the input signal was calculated by dividing (P-B) through (I-B) (Haring et al., 2007).

Percentage of input = (P-B)/(I-B)

For each interaction of interest PCRs were performed and the mean percentage of input was calculated, not incorporating significantly aberrant values. In cases where at least half of the performed PCRs produced a signal with a certain primer pair, the PCRs where no signal was visible were considered significantly aberrant. This reasoning was based on the fact that we worked very close to the limit of detection. When less than half of the performed PCRs produced a signal with a certain primer, it was assumed that the obtained signals were significantly aberrant and the result of always present contamination of the ChIP samples with unbound DNA (Furey, 2013). These signals were then not incorporated when calculating the mean percentage of input. The standard deviation was determined for every interaction. Additionally, the Limit of Detection (LOD) was calculated for each interaction in order to assess statistically significant results:

LOD= yB+3sB *yB: mean score of the blank value*
sB: standard deviation of the blank value

To determine the blank value for each gel the area next to the input signal of negative control PCRs were used.

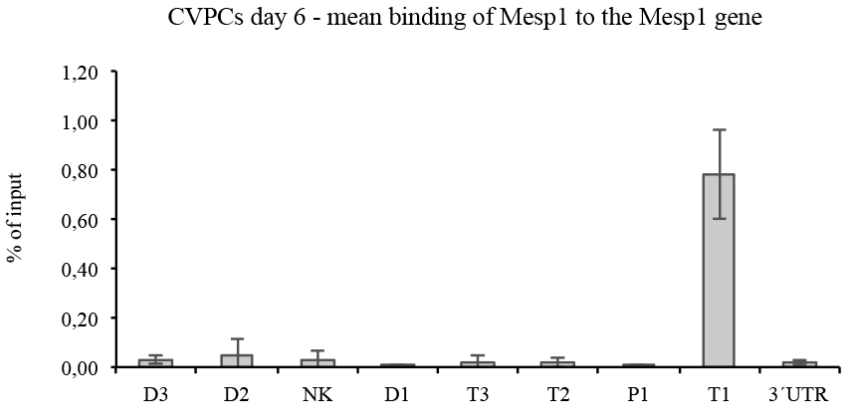


Figure 4.7. Typical set of data demonstrating the binding of Mesp1 to the Mesp1 gene at the DNA segments D3, D2, NK, D1, T3, T2, P1, T1 and 3'UTR on day 6 of CVPC differentiation. Error bars depict the standard deviation. n=4, LOD= 0,043

In Figure 4.7 an example of the evaluation of several agarose gels analyzing the binding of Mesp1 to its own gene on day 6 of CVPC differentiation is given. It can be clearly seen that the T1 primer pair produced a strong signal whereas the other investigated sites did not give a strong or constant signal. Therefore, it can be assumed that Mesp1 binds primarily the T1 DNA segment at this point in differentiation.

As it is not possible to show all results in such detail, from here on the binding signal intensities of all five investigated TFs to a certain gene at a specific time point were illustrated in one gene map. Binding signal intensity refers to the affinity of a protein to a certain DNA segment. Binding frequency refers to the number of DNA segments within a gene which were occupied by one or all investigated TFs.

PCR data investigating the interaction of Brachyury with the Nanog gene were generated by Georg Weitzer. PCR data investigating the interaction of Nanog with the Nanog, Brachyury, Nkx2.5 and Desmin genes were generated by Florian Tabery under my supervision.

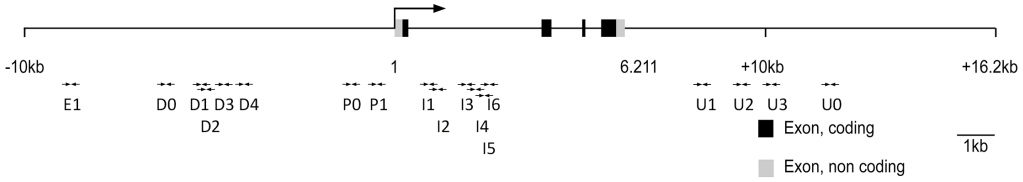
4.2.2. The Physical Interaction of Nanog and Brachyury with the Nanog Gene

4.2.2.1. In Silico Analysis of Putative Transcription Factor Binding Sites and Primer Design

Primer pairs for the Nanog gene were designed by Georg Weitzer and amplified DNA segments of the Nanog gene are listed in Table 4.1. They comprise a putative enhancer element (E1), a distal promoter (D0-D4), a proximal promoter (P0, P1) as well as certain DNA segments in intron 1 (I1-I6) and the 3'UTR (U0-U3) of the Nanog gene (Kuroda et al., 2005; Rodda et al., 2005; Suzuki et al., 2006).

In silico analysis identified numerous binding sites for both Nanog and Brachyury along the Nanog gene. Primer pairs were designed to amplify the DNA segment including the putative TF binding sites. However, due to high redundancy of the sequence some primer pairs were designed to amplify DNA segments closely adjacent to the putative TF binding site.

Table 4.1. Location of putative transcription factor binding sites containing DNA segments of the Nanog gene. Arrows indicate the position of the primer pairs.



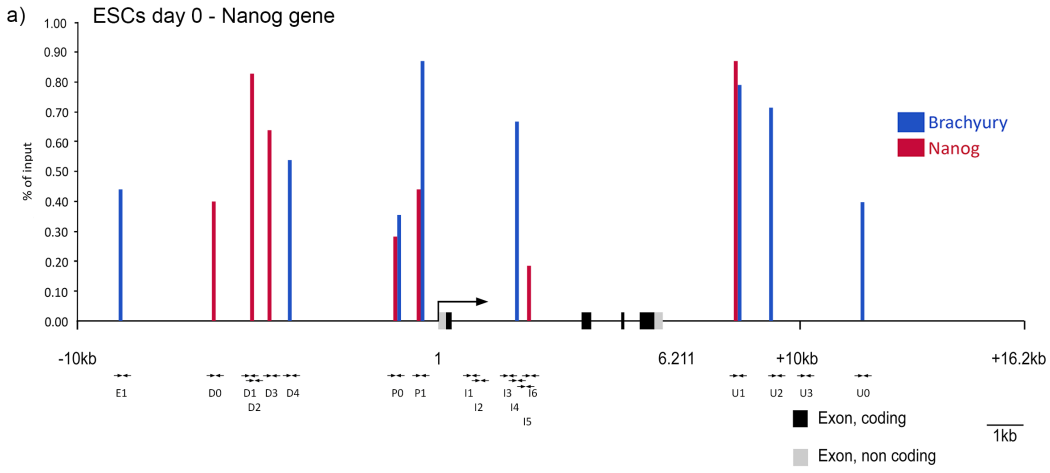
Putative binding sites in DNA segments amplified or in close proximity for the Nanog gene		
Name/ Extend of the amplified DNA segment [1]	Nanog	Brachyury
E1 (-8656 to -8809)	TAATTT (-8520)[2]	$(AC)_8$ (-8542)
D0 (-6115 to -6282)	TAATGG (-5980)	$(GT)_3$ (-5955), GGTGT (-5956)
D1 (-5085 to -5281)	TAATTT (-4793)	
D2 (-5046 to -5189)	TAATTT (-4793)	
D3 (-4657 to -4774)	TAATTT (-4793), AAATTA (**-4680**)[3]	$(AC)_3$ (**-4661**)
D4 (-4021 to -4168)	AAATTA (-3718)	$(GT)_7$ (-4212), AGGTGTGA (**-4119**)
P0 (-1079 to -1247)	TAATTT (-886)	Several $(GT)_n$ sites upstream
P1 (-347 to -470)	TAATTT (-115)	GGTGT (-171)

I1 (+738 to +1130)	TAATGT (**+916**)	Several $(GT)_n$ sites (between **+765 and +884**)
I2 (+1137 to +1277)	TAATTT (**+1241**)	Several $(GT)_n$ sites (between **+1171 and +1205**)
I3 (+1849 to +2123)	TAATGG (**+1925**), TAATGT (**+1981**)	$(GT)_{12}$ (**+1985**)
I4 (+2105 to +2210)	TAATGT (+1981)	$(GT)_{12}$ (+1985)
I5 (+2189 to +2441)	TAATGT (+1981)	$(GT)_{12}$ (+1985)
I6 (+2443 to +2580)	TAATTG (**+2574**), TAATTT (+2582)	$(GT)_3$ (+2745)
U1 (+7918 to +8200)	TAATTT (+8710)	GGTGT (**+8107**)
U2 (+9110 to +9219)	AAATTA (+9359), TAATTT (+9263)	TCACACCT (+9230)
U3 (+9880 to +9987)	AAATTA (**+9841**), TAATTG (+9824)	$(AC)_{27}$ (+9763)
U0 (+11396 to +11648)	TAATTT (+11213)	$(GT)_4$ (+11378)

[(1)] Primers are located at the very beginning and end of the denoted DNA segment (for primer sequence see Table 2.1). Numbers depict base pairs in relation to the transcription initiation site.

[(2)] regular letter = binding site lies adjacent to the amplified DNA segment; [(3)] bold letter = binding site lies within the amplified DNA segment

4.2.2.2. The Physical Interaction of Nanog and Brachyury with the Nanog Gene in ESCs



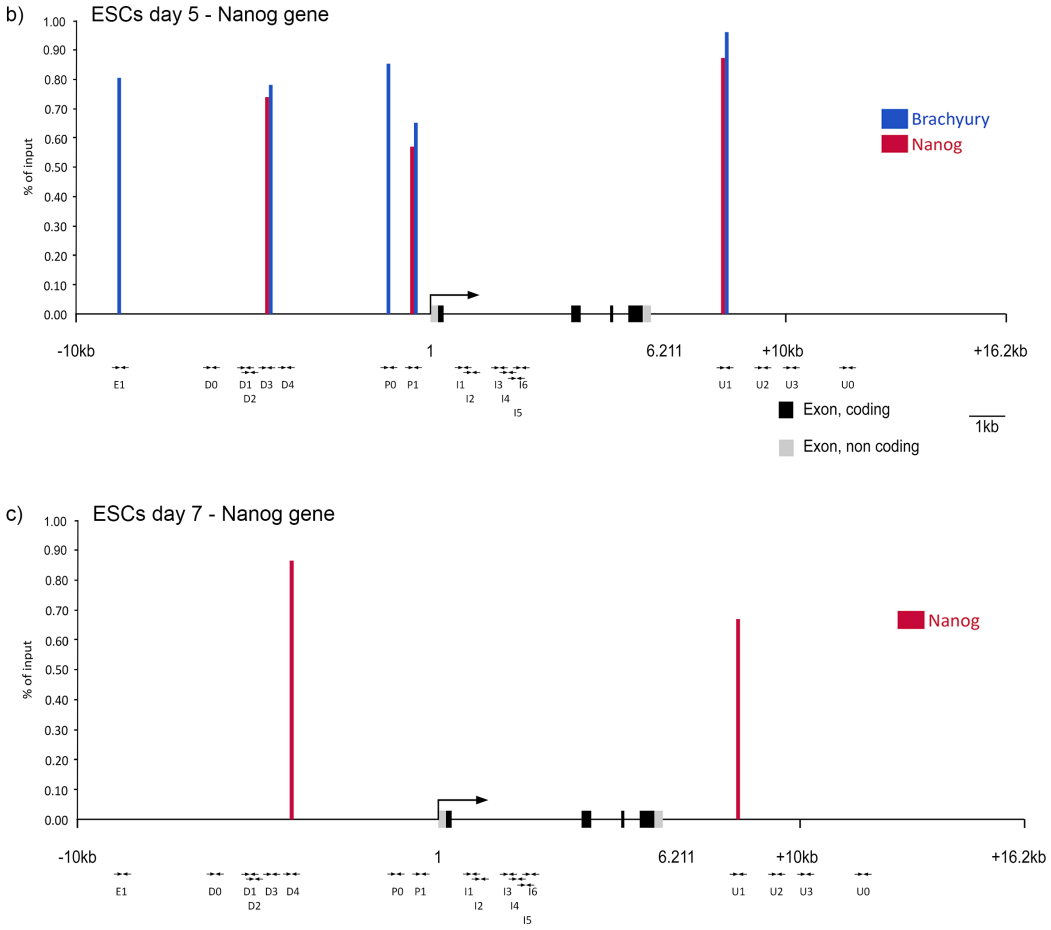


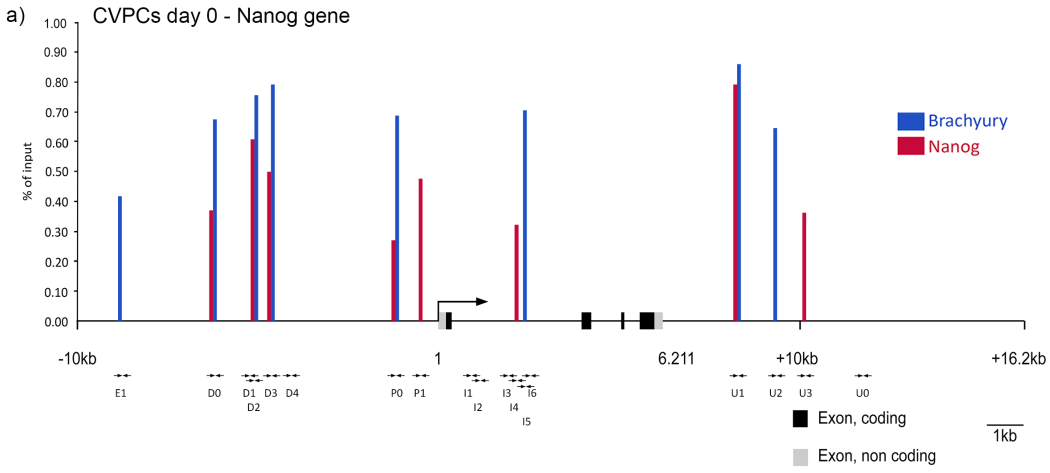
Figure 4.8. The interaction of Nanog and Brachyury with the Nanog gene. In a) undifferentiated ESCs, b) on day 5 of ESC differentiation and c) the binding of Nanog to the Nanog gene on day 7 of ESC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated ESCs Nanog is known to regulate its own expression (Mitsui et al., 2003). It was not surprising that there were several strong binding signals from Nanog to its own gene. Especially the DNA segments from D0 to D3 were occupied quite frequently by Nanog in undifferentiated ESCs. The first intron was hardly occupied at all. Brachyury bound intensely to several DNA segments distributed over the whole gene. So far it had been shown that Brachyury has an activating effect on Nanog gene expression, by binding an imperfect T-box Palindrome in the distal promoter (Suzuki et al., 2006). Our data strongly suggest that Brachyury is able to bind several DNA segments of the Nanog gene.

On day 5 of ESC differentiation the binding frequency (number of DNA segments occupied on a given gene) of Nanog to its own gene declined greatly, with only three signals remaining. Interestingly those three signals were found in DNA segments that had already been occupied in undifferentiated ESCs. The binding frequency of Brachyury to the Nanog gene also decreased, but the signal intensity at the remaining sites increased significantly when compared to undifferentiated cells.

On day 7 of ESC differentiation only the interaction of Nanog with the Nanog gene was investigated. Only the U1 and the D4 DNA segment were occupied at this time point. On the previously investigated days the D4 DNA segment was not occupied, but the D3 DNA segment was occupied strongly by Nanog. Since those two primer pairs are in very close proximity to each other, it might very well be that the same binding event was detected as before. The binding of Nanog to its own proximal promoter (P0 & P1) disappeared completely. Self-regulation of the Nanog gene appears to become less important as differentiation progresses, and possibly other yet unknown factors take over.

4.2.2.3. The Physical Interaction of Nanog and Brachyury with the Nanog Gene in CVPCs



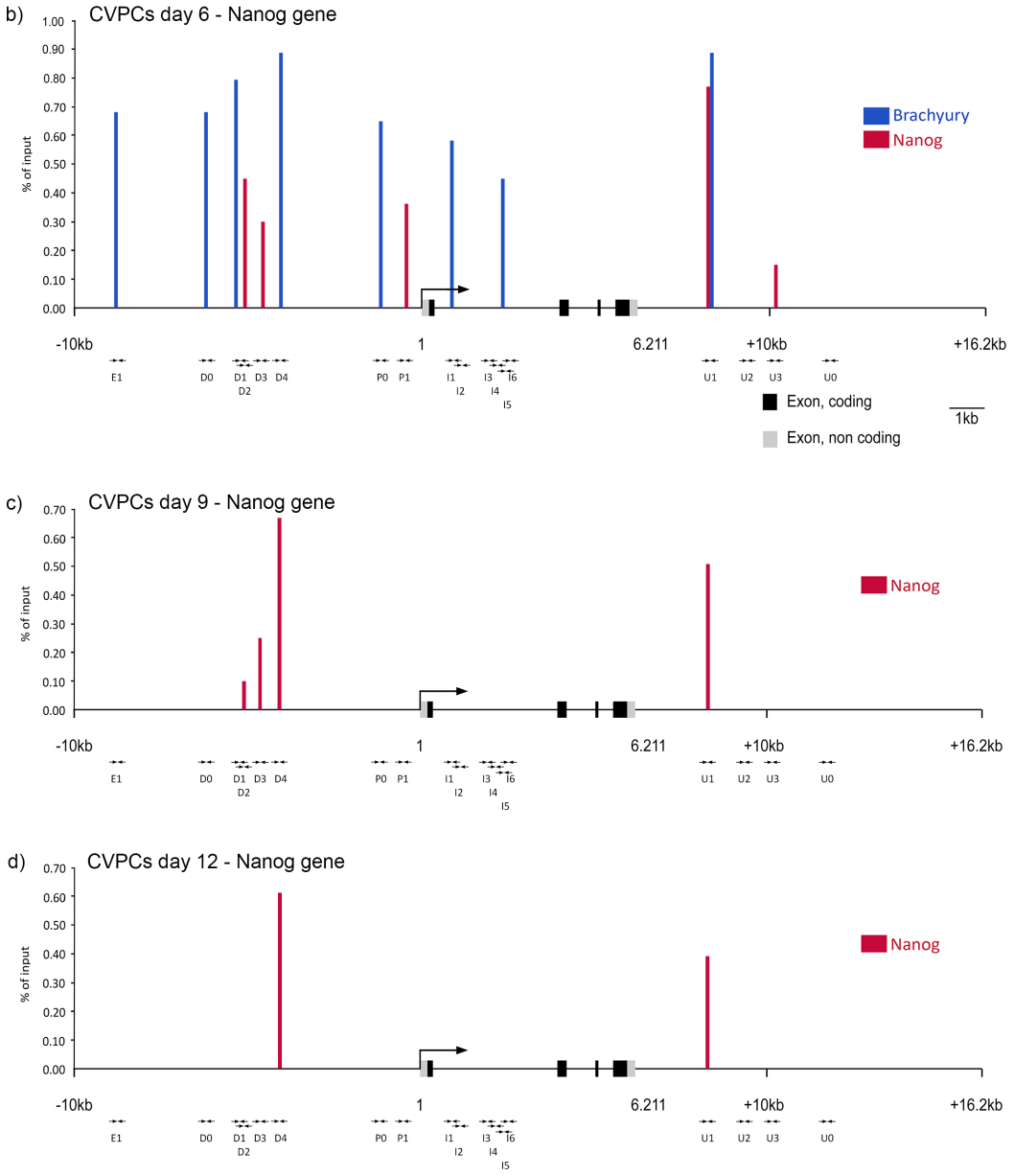


Figure 4.9. The interaction of Nanog and Brachyury with the Nanog gene. In a) undifferentiated CVPCs, b) on day 6 of CVPC differentiation and c) the binding of Nanog to the Nanog gene on day 9 and d) day 12 of CVPC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the used primer pairs for PCR analysis.

In undifferentiated CVPCs Nanog showed pretty much the same binding pattern as in undifferentiated ESCs, with only one additional signal at the U3 DNA segment. This suggests that self-regulation of Nanog functions in a very similar way in undifferentiated

ESCs and CVPCs. Brachyury on the other hand showed a different behavior in CVPCs and ESCs, with visibly stronger signals in CVPCs.

On day 6 of CVPC differentiation the binding of Nanog to its own gene declined visibly. Nanog occupied the same DNA segments as in undifferentiated CVPCs, but some signals disappeared completely.

The frequency of Brachyury binding signals was the same on day 6 of differentiation and in undifferentiated cells, but slightly different DNA segments were occupied. There was no longer a signal at the beginning of the first intron, but instead the U2 DNA segment was occupied. Noticeable is also the fact that Brachyury interacted with the Nanog gene much more strongly than Nanog itself.

On day 9 of CVPC differentiation the interaction of Nanog with the Nanog gene decreased further. Like on day 7 of ESC differentiation, there was no longer any binding to the promoter, whereas the binding to the D1 to D4 DNA segments remained comparably unchanged. The interaction of Brachyury with the Nanog gene was not investigated beyond day 6 of CVPC differentiation.

On day 12 of CVPC differentiation Nanog only interacted with the D4 and U1 DNA segments, the latter was occupied in both cell lines on every investigated day.

4.2.2.4. Résumé of the Protein Interaction with the Nanog Gene

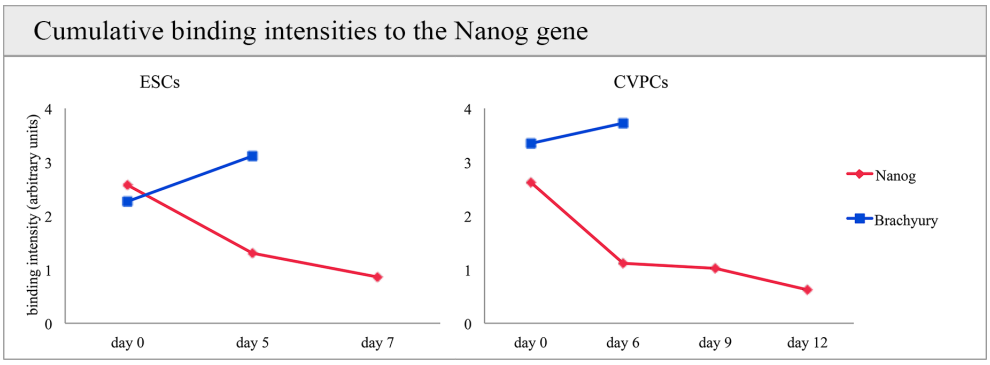


Figure 4.10. Cumulative binding intensities of Nanog and Brachyury to the Nanog gene over time. Only signals obtained from the 5'UTR of the Nanog gene were considered in order to make the results comparable to the Brachyury gene.

To compare the cumulative binding intensities of Nanog and Brachyury to the Nanog gene over time both in ESCs and CVPCs, the intensity of the protein interaction with all DNA segments of the gene where the protein bound was summarized (Figure 4.10). Only

relative changes can be depicted because the affinities of the Nanog and Brachyury antibodies used for ChIP are not known.

In ESCs, the Nanog gene showed a very clear trend: the binding intensity of Brachyury increased from day 0 to day 5 of differentiation, whereas the presence of the Nanog protein on the gene clearly decreased over time. Linear regression analysis predicted a linear decline (R= -1) for the binding intensity of Nanog to the Nanog gene over the course of ESC differentiation. For details see chapter 4.2.7.

In CVPCs, the same trend could be observed as in ESCs: an increase in Brachyury binding and a decrease in Nanog binding. The most significant difference between the two cells lines was the observation that Brachyury was bound more strongly to the Nanog gene in CVPCs than in ESCs. This is not surprising as CVPCs only generate cells of the mesoderm, which are more prone to expressing Brachyury, than ESCs, which give rise to cells of all three germ layers, which should lead to fewer Brachyury positive cells.

In general it can be assumed that Nanog gene regulation becomes more dependent on Brachyury as differentiation proceeds, while the self-regulating aspect of Nanog is most important in undifferentiated stem cells.

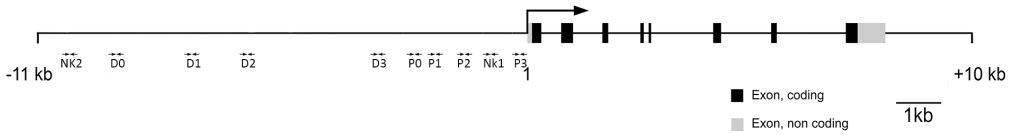
4.2.3. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Brachyury Gene

4.2.3.1. In Silico Analysis of Putative Transcription Factor Binding Sites and Primer Design

The Brachyury gene is not very well described yet, therefore, primer pairs were chosen in DNA segments with an accumulation of putative binding sites for various TFs as identified by in silico analysis of the sequence, assuming that these DNA segments were most likely to contain regulatory elements. Amplified DNA segments of the Brachyury gene are listed in Table 4.2. Primer pairs P0 to P3 are located in the proximal DNA segment, D0 to D3 in the distal DNA segment of the gene. The primer pairs labeled with NK were meant to act as a negative control, where no binding was expected.

Especially in the DNA segment containing the primer pairs P1 and P0, as well as in proximity of D0 and NK2, a large number of Sox2, Oct4, Klf4 and Nanog binding sites were found. It can be speculated that these DNA segments might indicate the presence of something resembling the recently described super enhancers (Whyte et al., 2013).

Table 4.2. Location of putative transcription factor binding sites containing DNA segments of the Brachyury gene. Arrows indicate the position of the primer pairs.



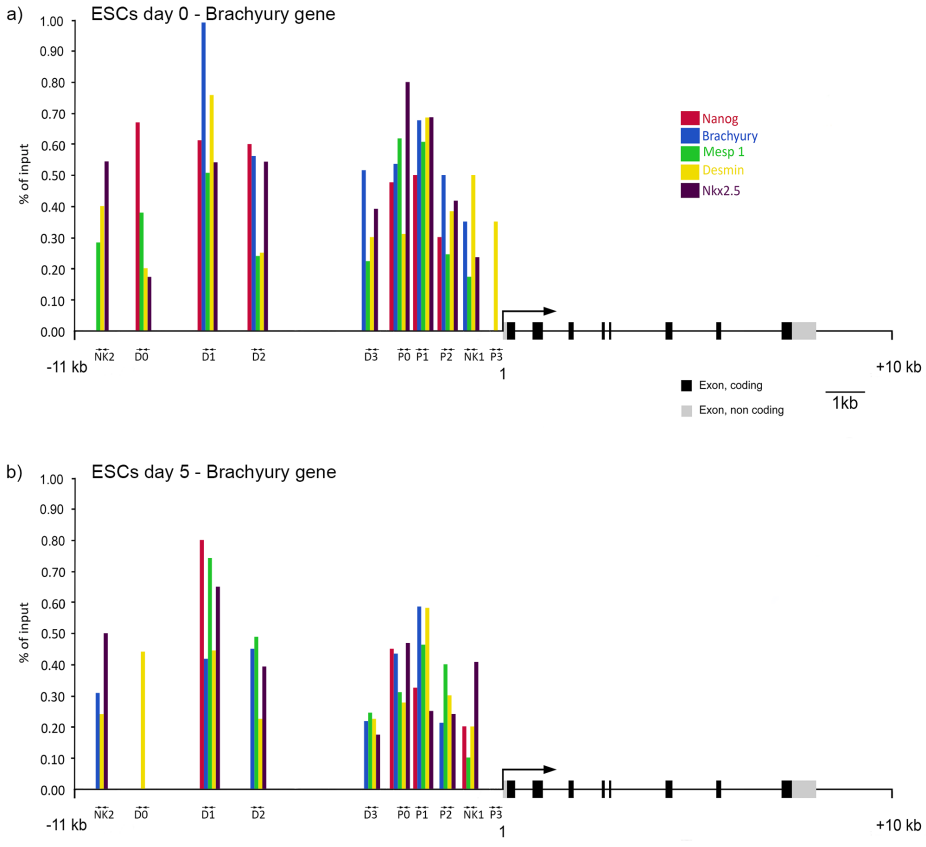
Putative binding sites in DNA segments amplified or in close proximity for the Brachyury gene				
Name/ Extend of the amplified DNA segment[1]	Nanog	Brachyury	Mesp1	Nkx2.5
P3 (-109 to -356)	GGACCCA-TTT (**-113**)[2]	GGTGT (**-184; -322**)	CAGCTG (**-200**)	TCAAGTG (-63)[3]
Nk1 (-771 to -986)	CCCATTGA (**-979**)	ACACC (-754; **-897**)	CAGGTG (-672) CAACTG (**-909**)	
P2 (-1480 to -1625)	GCTCATTG (-1450)	GGTGT (-1412) $(AC)_2$ (-1470)	CAAATG (-1428)	AAGTG (-1393)
P1 (-2144 to -2269)	AAATTA (-1904)	ACACC (-1768) $(GT)_2$ (-2400)	CAGCTG (-1982) CAACTG (-2068)	TAAT (-1885)
P0 (-2468 to -2694)	TAATGT (**-2662**)	$(GT)_2$ (**-2489**) ACACC (-2828)	CAACTG (-2389)	AAGTG (-2458) TAAT (**-2647**)
D3 (-3353 to -3534)	TAATGT (**-3517**)	ACACC (-3214) GGTGT (-3235)	CAGCTG (**-3445**)	TAAT (**-3502**)
D2 (-6288 to -6389)	TAATTG (-6163) AAATTA (**-6352**)	$(AC)_3$ (**-6372**) multiple (AC)n sites between D2 and D1	CACTTG (-6268) CATGTTG (-6425)	TAAT (**-6348**)
D1 (-7567 to -7667)	TAATTT (**-7597**)	$(GT)_3$ (**-7628**) GGTGT (**-7634**)	CATATG (7516) CATGTG (-7695)	TAAT (**-7577**)
D0 (-9209 to -9338)	CAATTA (-9142)	ACACC (-9299)	CAGCTG (**-9255**)	TAAT (-9062)
NK2 (-10270 to -10401)	TAATTT (-9962)	$(AC)_2$ (**-10320**) ACACC (-10473)	CAGATG (**-10311**)	AAGTG (-10462)

[1] Primers are located at the very beginning and end of the denoted DNA segment (for primer sequence see Table 2.1). Numbers depict base pairs in relation to the transcription initiation site.

[2] Bold = binding site lies within the amplified DNA segment; [3] Regular = binding site lies adjacent to the amplified DNA segment

Between the primer pairs D2 and D1 a highly repetitive (AC)n DNA segment was found that could be bound by Brachyury itself. As it was not possible to design primer pairs binding directly to this DNA segment, flanking segments were amplified. For ChIP the chromatin is sheared into ideally 400 to 600 bp long fragments. Most primer pairs amplify a DNA segment of approximately 100 to 300 bp, therefore, it can be assumed that a signal can be obtained even if the protein-DNA interaction takes place outside the amplified DNA segment. This fact makes it possible to study highly repetitive DNA segments, but also makes it harder to pinpoint the exact location of the interaction in the case of a signal. The amplified DNA segment of the P3 primer pair comprises an extremely high C/G content (161 bp out of 242 bp), making PCR amplification less efficient. This should however not affect binding signal intensity, because the percentage of input is calculated by dividing the pull-down signal through the input signal.

4.2.3.2. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Brachyury Gene in ESCs



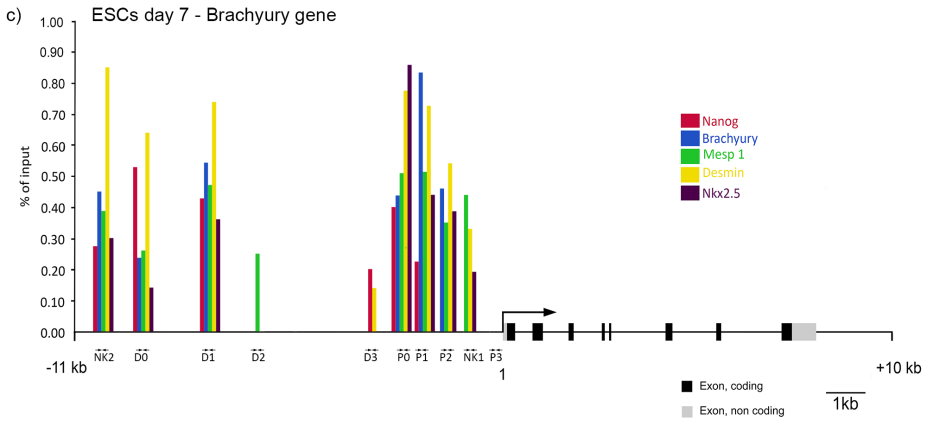


Figure 4.11. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the 5'UTR of the Brachyury gene. In a) undifferentiated ESCs, b) on day 5 and c) on day 7 of ESC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated ESCs all five investigated proteins showed very strong binding at multiple sites. This was not expected as in undifferentiated ESCs, of the five investigated proteins only Nanog should be present, but could be explained by the fact that in standard serum/LIF culture conditions there are always a certain percentage of cells that are already differentiated to one degree or another. Addition of 2000 Units of LIF per ml medium supplementary to cultivation of the ESCs on a LIF-secreting feeder cell layer did not reduce the intensity or the amount of obtained signals. Furthermore, expression of proteins during differentiation processes is mostly determined by western blot or mRNA expression analysis. Both techniques require a relatively large amount of protein and mRNA copies, respectively, in the majority of investigated cells to obtain a signal, whereas ChIP can generate a signal even if only one protein binds to a certain DNA segment in the majority of investigated cells. Therefore it cannot be eliminated that the TFs are present in the cells in numbers too low to detect with western blot or mRNA expression analysis, but nevertheless bind certain genes strongly and frequently enough to produce signals with high intensity in ChIP experiments.

The strongest and most frequent signals could be observed at the D1, as well as at the P0 and P1 DNA segments. These DNA segments were occupied by all investigated TFs with high signal intensity. The other DNA segments were also occupied, but the obtained signals were markedly weaker. Interestingly, Nanog and Brachyury occupied the Brachyury gene the least frequently. Mesp1, and Nkx2.5 occupied every DNA segment,

except P3, closest to the transcription initiation site and Desmin was present at every investigated DNA segment. The strong occupation of the NK2 and NK1 DNA segments was at first surprising, but further analysis of the DNA sequence identified several at least partial TF binding sites in these segments.

On day 5 of ESC differentiation there were visibly less and weaker binding signals than in undifferentiated cells. Our western blot data as well as previous work from our group suggest that the Brachyury gene is expressed on day 5 of differentiation. Especially the D0 and NK2 DNA segments produced significantly less signals on day 5 when compared to day 0 of ESC differentiation, possibly suggesting that this DNA segment might have an inhibitory effect on Brachyury expression. It can however be excluded that a putative regulatory element more than 5 kb upstream of the transcription initiation site is sufficient for Brachyury inhibition, as FACS experiments with cells that contain only a 5 kb Brachyury promoter in front of an EGFP reporter gene showed a clear peak on day 5 of differentiation (results from the diploma thesis of Lucia Leitner). Another notable change is that Nanog, Mesp1 and Nkx2.5 signal intensity at D1 clearly increased when compared to day 0, while the majority of all other signals decreased.

On day 7 of ESC differentiation the Brachyury protein could no longer be detected by western blot, suggesting a very short time frame of high levels of protein expression, probably limited to day 5.

Interestingly, the binding frequency and intensity to the D0 and NK2 DNA segments increased again, while the binding of Nanog, Mesp1 and Nkx2.5 to D1 declined. The binding of Desmin to the Brachyury gene decreased on day 5 when compared to undifferentiated cells, but strongly increased on day 7, especially in the proximal and very distal DNA segments. A similar but not as distinct observation could be made for Brachyury.

4.2.3.3. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Brachyury Gene in CVPCs

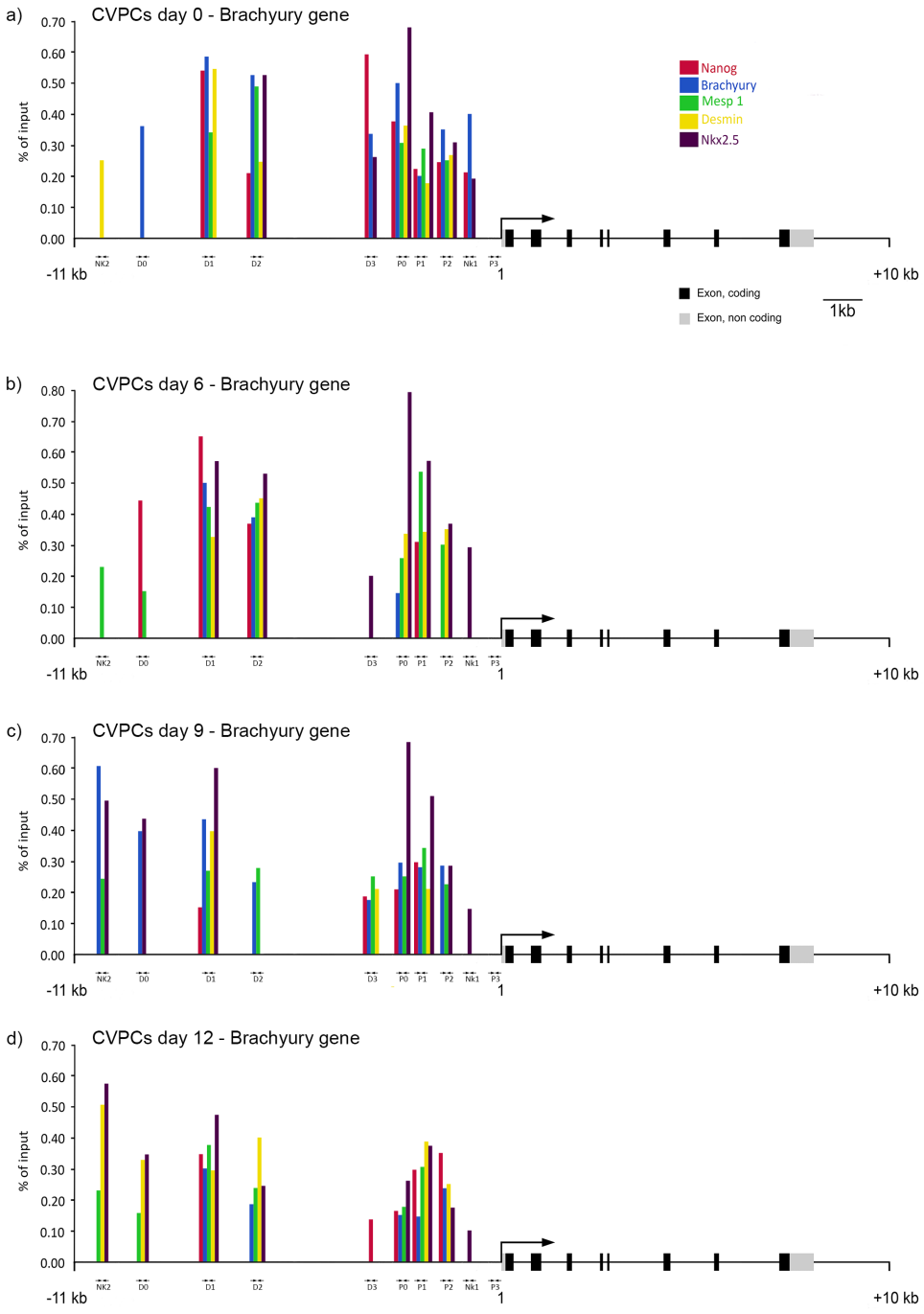


Figure 4.12. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the 5'UTR of the Brachyury gene. In a) undifferentiated CVPCs, b) on day 6, c) day 7 and d) day 12 of CVPC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated CVPCs all five tested proteins bound strongly to the Brachyury gene, but the overall binding frequency and intensity was slightly smaller than in undifferentiated ESCs. Especially the very distal DNA segments were occupied quite infrequently: the NK2 DNA segment was occupied only by Desmin and the D0 DNA segment only by Brachyury. Possibly these very distal DNA segments have a different regulatory function in CVPCs than in ESCs.

On day 6 of CVPC differentiation the proximal DNA segment of the 5'UTR of the Brachyury gene was occupied distinctly less frequently than in undifferentiated CVPCs, while binding to the D1 and D2 DNA segments was slightly stronger than in undifferentiated CVPCs. The very distal DNA segment (NK2 and D0) was still occupied only sparsely. Especially the binding frequency of Brachyury dropped drastically when compared to undifferentiated CVPCs, with only three signals in the P0, D2 and D1 DNA segments.

On day 9 of CVPC differentiation western blot experiments showed high Brachyury protein expression (Figure 4.2). Interestingly the same FACS experiments as mentioned before (4.2.3.2) did not show a peak of reporter gene expression in CVPC differentiation (data from the diploma thesis of Lucia Leitner), perhaps indicating that the reporter lacks an essential regulatory element.

In the ChIP experiments there were several similarities between day 9 of CVPC differentiation and day 5 of ESC differentiation in regard to the occupation frequency of the Brachyury gene, but there were also some clear differences: in CVPCs the very distal DNA segment (NK2 and D0) showed strong TF binding on the day of maximal Brachyury protein expression, but in ESCs this distal DNA segment produced the weakest binding signals on the day of Brachyury protein expression. Especially Brachyury and Nkx2.5 bound strongly to this distal DNA segment on day 9 of CVPC differentiation. The binding to the D1 and D2 DNA segments was visibly weaker on day 9 than on the previously investigated days of CVPC differentiation. These data from the ChIP experiments together with the before mentioned FACS data might indicate a different regulation of Brachyury in CVPCs and ESCs.

On day 12 of CVPC differentiation the binding signal intensity to the Brachyury gene decreased when compared to the other time points. Occupation frequency did not change significantly, but different DNA segments showed high occupation frequency on day 12 when compared to day 9 of CVPC differentiation: the D3 DNA segment was occupied only by Nanog on day 12 of CVPC differentiation, but the D2 DNA segment showed

more frequent occupation on day 12 than on day 9 of CVPC differentiation. Nkx2.5 still occupied the Brachyury gene very frequently on day 12 of CVPC differentiation and produced binding signals with a high intensity, especially in the more distal DNA segments. Desmin occupied the distal DNA segments of the Brachyury gene (NK2 to D2) much more frequently on day 12 than on day 9 of CVPC differentiation. At the same time Brachyury produced only few signals with relatively low intensity when compared to day 9 of CVPC differentiation.

4.2.3.4. Résumé of the Protein Interaction with the Brachyury Gene

To compare the cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the Brachyury gene over time in both ESCs and CVPCs, the intensity of the protein interaction with all DNA segments of the gene where the protein bound was summarized (Figure 4.13). Only relative changes can be depicted because the affinities of the Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin antibodies used for ChIP are not known.

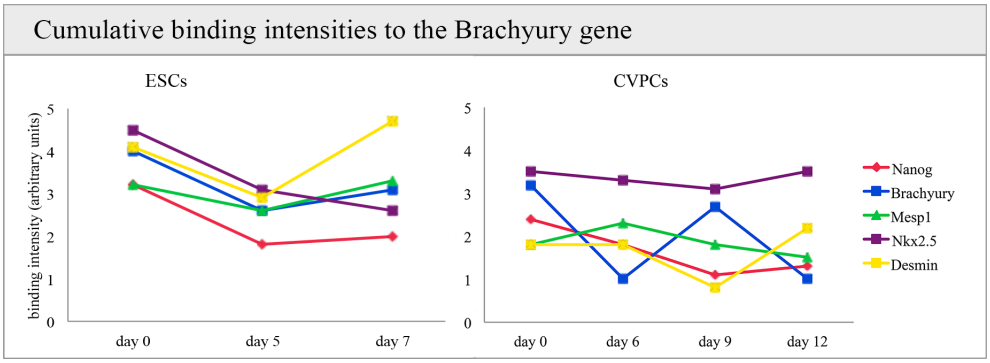


Figure 4.13. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the Brachyury gene in a) ESCs and b) CVPCs over time.

In ESCs, all investigated TFs showed a similar trend: the cumulative binding intensity of the five proteins decreased on day 5 of ESC differentiation when compared to undifferentiated cells. Only the binding intensity of Desmin rose again sharply on day 7 of ESC differentiation; The other TFs showed only a slight increase or decrease in binding intensity on day 7 when compared to day 5 of ESC differentiation. Nanog, Brachyury and Nkx2.5 had the strongest cumulative binding intensity in undifferentiated ESCs, but Mesp1 and Desmin on day 7. It was not surprising that Nanog bound the Brachyury gene strongest in undifferentiated cells, as Nanog has been shown to at least indirectly suppress Brachyury gene expression (Suzuki et al., 2006). Mesp1 was shown to

suppress Brachyury expression through a negative feedback loop starting around E6.5 (Bondue and Blanpain, 2010), therefore the slight increase in Mesp1 binding intensity on day 7 might be the remnants of this described interaction. Linear regression analysis predicted a linear decline (R= -1) for the binding intensity of Nkx2.5 to the Brachyury gene over the course of ESC differentiation.

In CVPCs, the strong binding of Nkx2.5 to the Brachyury gene throughout differentiation was most noticeable. Even though no regulation of Brachyury by Nkx2.5 has been described yet, such an interaction would not be surprising as Nkx2.5 is one of the widest used markers defining CVPCs (Tanaka et al., 1999) and should therefore be abundantly present throughout CVPC differentiation. Brachyury showed the strongest fluctuation in cumulative binding intensity, with drastic drops in binding signal intensity on days 6 and 12 of CVPC differentiation. The overall trend for Nanog and Mesp1 binding intensity was more or less a steady decline. Desmin binding signal intensity rose sharply on day 12 after a drop on day 9 of CVPC differentiation.

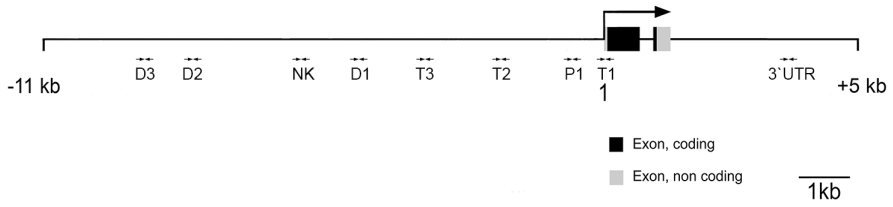
4.2.4. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Mesp1 Gene

4.2.4.1. In Silico Analysis of Putative Transcription Factor Binding Sites and Primer Design

For the Mesp1 gene three regulatory DNA segments have been described: T1, T2 and T3 (van den Ameele et al., 2012). Primer pairs were designed for these described regulatory segments, as well as for DNA segments in which putative TF binding sites could be identified in silico. Amplified DNA segments of the Mesp1 gene are listed in Table 4.3.

A canonical Brachyury binding site was found in the 3'UTR and several non-canonical Brachyury binding sites were identified in the 5'UTR of the Mesp1 gene. The P1 primer pair is located in the proximal segment of the gene, while D1 to D3 are in the more distal segment. The NK primer pair was designed to act as a negative control, because only one putative E-box could be found in the amplified DNA segment.

Table 4.3. Location of putative transcription factor binding sites containing DNA segments of the Mesp1 gene. Arrows indicate the position of the primer pairs.



Putative binding sites in DNA segments amplified or in close proximity for the Mesp1 gene				
Name/ Extend of the amplified DNA segment[(1)]	Nanog	Brachyury	Mesp1	Nkx2.5
3'UTR (+3832 to +4201)		GGAGTTCCACTTA -TCTGTGAAAAG (**+3980**)[(2)]	CAGGTG (**+4073**)	CCACTTA-TCTGTGA (**+3984**)
T1 (-19 to -169)	CCATTG (-17)[(3)] TCATTTA (-297)	AGGTGTGA (**-51**)	CACCTG (-184)	
P1 (-903 to -1254)		(GT)n (**-943**)	CAGCTG (**-1207**)	
T2 (-2329 to -2550)	CATTGC (**-2360**)	TCACACCT[(4)] (**-2392**)	CACCT (**-2395**)	TAAT (-2294)
T3 (-3835 to -3983)	CTCATTTCC (-4100)	GGTGT (-3478)	ACAAGT (**-3861**)	AAGTG (**-3863**; **-3933**)
D1 (-5087 to -5437)	TCATTTA (-5500)	TCACACCT (**-5187**)		TAAT (**-5255**; **-5438**)
NK (-6164 to -6464)		ACACC (-6524)	CAACTG (**-6241**) CACCTG (-6630)	
D2 (-8283 to -8459)	TAATTG (-8628)	TCACACCT (**-8297**)	CAGGTG (-8113) CATATG (**-8314**)	TAAT (-8628)
D3 (-9225 to -9434)	CCATTA (-9499)	AGGTGTGA (**-9245**) (AC)n (-9550)	CACCTG (**-9343**)	TAAT (**-9308**)

[(1)] Primers are located at the very beginning and end of the denoted DNA segment (for primer sequence see Table 2.1). Numbers depict base pairs in relation to the transcription initiation site.

[(2)] Bold letter = binding site lies within the amplified DNA segment; [(3)] Regular letter = binding site lies adjacent to the amplified DNA segment

[(4)] published Brachyury responsive element (David et al., 2011)

4.2.4.2. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Mesp1 Gene in ESCs

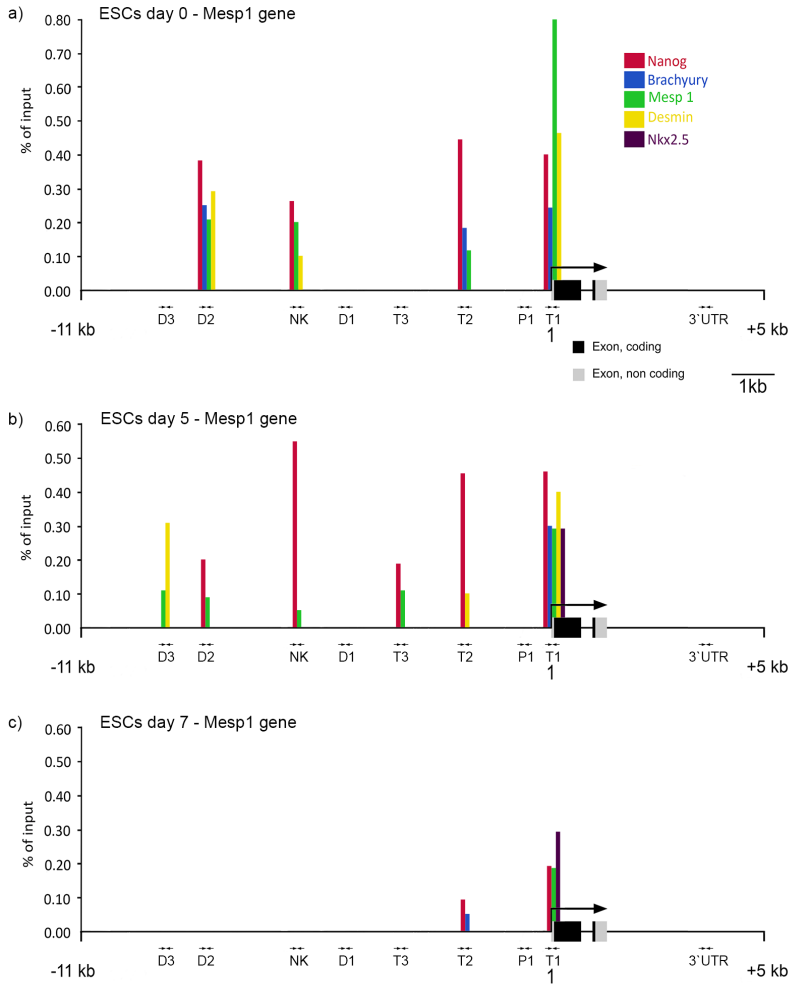


Figure 4.14. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Mesp1 gene. In a) undifferentiated ESCs, b) on day 5 and c) on day 7 of ESC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

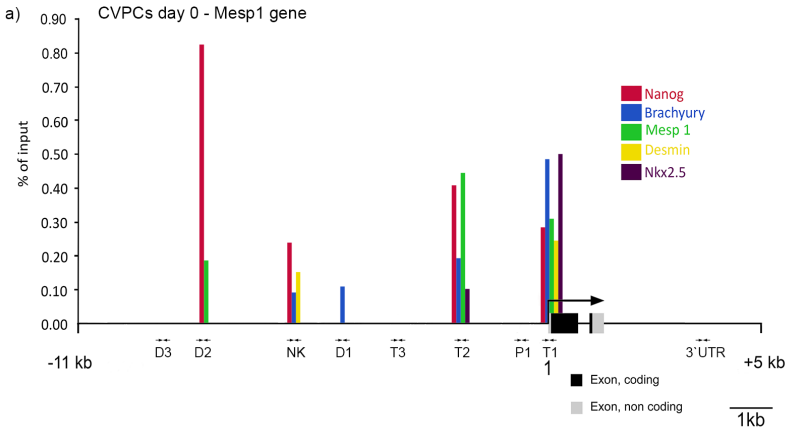
In undifferentiated ESCs there were more interactions of TFs with the Mesp1 gene than expected, but by far less than on the Brachyury gene. Several DNA segments were not occupied by any of the investigated TFs (3'UTR, P1, T3, D1 and D3), while others were bound by multiple TFs. Especially the described T1 DNA segment was bound rather strongly in undifferentiated ESCs. Nanog, Brachyury, Mesp1 and Desmin all occupied several DNA segments, but Nkx2.5 did not interact with the Mesp1 gene at this time

point. Interestingly, the NK DNA segment was occupied by Nanog, Mesp1 and Desmin, even though only putative binding sites for Brachyury and Mesp1 were found in close proximity. The simplest explanation would be that there are probably binding sites for the investigated TFs that are not known yet. Another explanation could be that our proteins of interest did not interact directly with DNA, but rather with other DNA binding factors, which could mediate their interaction with DNA.

On day 5 of ESC differentiation we found a similar amount of TF binding signals, but more DNA segments were occupied than in undifferentiated ESCs. Only the 3´UTR, P1 and D1 DNA segments were not occupied by any TF. Nanog bound to the Mesp1 gene more strongly on day 5 of ESC differentiation than on day 0. In contrast to undifferentiated ESCs, Nkx2.5 occupied the Mesp1 gene.

On day 7 of ESC differentiation the Mesp1 gene was hardly occupied at all by the investigated TFs. Only the described T1 and T2 DNA segments were occupied, but even those binding signals were not very strong. According to literature, Mesp1 should be expressed from E6.5 to E7.5 in ESCs (Bondue and Blanpain, 2010). It could be that especially the distal DNA segments have a negative regulatory function and are no longer occupied when the gene is transcribed. Especially Desmin showed rather strong DNA binding on days 0 and 5, but did not occupy the Mesp1 gene at all on day 7 of ESC differentiation. It could be speculated that Desmin might have an inhibitory effect on Mesp1 expression. Surprisingly, Brachyury never bound to the canonical binding site in the 3'UTR of the Mesp1 gene.

4.2.4.3. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Mesp1 Gene in CVPCs



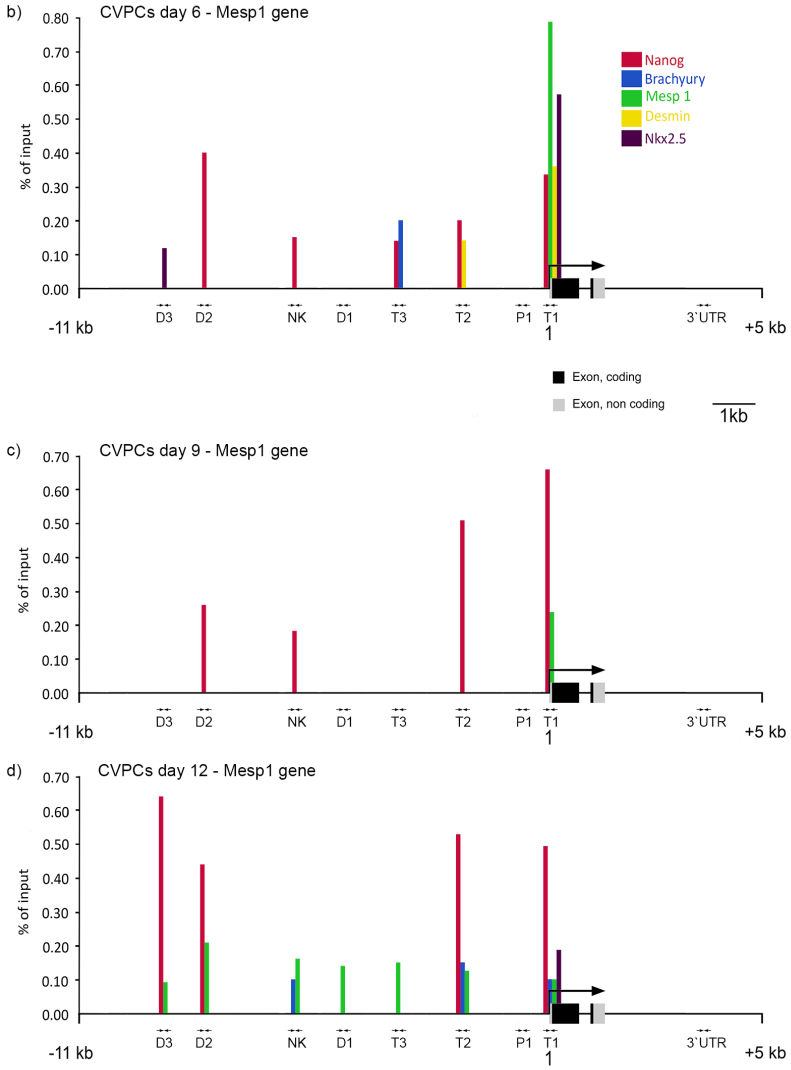


Figure 4.15. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Mesp1 gene. In a) undifferentiated CVPCs, b) on day 6, c) day 7 and d) day 12 of CVPC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated CVPCs pretty much the same DNA segments of the Mesp1 gene were occupied as in undifferentiated ESCs. There were however some differences in the distribution of the individual TFs: the D2 DNA segment was occupied by fewer TFs in undifferentiated CVPCs than in undifferentiated ESCs. The T1 and T2 DNA segments were still occupied by the same TFs as in undifferentiated ESCs, but here in CVPCs Nkx2.5 also bound to these DNA segments.

On day 6 of CVPC differentiation only Nanog occupied the Mesp1 gene at more than two of the investigated DNA segments and the highest concentration of binding signals was observed at the T1 site, similar to the situation on day 0.

On day 9 of CVPC differentiation only Nanog occupied the Mesp1 gene frequently. It is also noteworthy that Nanog occupied the same DNA segments (T1, T2, NK and D2) on days 0, 6 and 9 of CVPC differentiation, however with changing signal intensity. The time frame of Mesp1 protein expression in CVPCs is not known yet. We did however find that Brachyury peaks on day 9 of CVPC differentiation (Figure 4.2) and it is known that in ESCs Mesp1 follows 1.5 to 2 days after Brachyury expression (Saga et al., 1996, 1999). Therefore it could be speculated that Mesp1 might be expressed around day 10.5 in CVPC differentiation. Further experiments would be needed to clarify that issue. If the situation is similar in CVPCs and ESCs, where there were hardly any signals on day 7 of ESC differentiation, the decline in signal frequency on day 9 of CVPC differentiation could be a sign that the gene will soon be expressed.

On day 12 of CVPC differentiation binding signal frequency increased again. Except for Desmin, all investigated TFs occupied the Mesp1 gene at this time point, however with very different frequency and intensity. Nanog occupied slightly different DNA segments and signal intensity increased when compared to the other time points. Mesp1 occupied more DNA segments than before but with rather weak binding signal intensity. Brachyury was found almost exclusively at the same DNA segments as in undifferentiated CVPCs and Nkx2.5 solely bound to T1.

4.2.4.4. Résumé of the Protein Interaction with the Mesp1 Gene

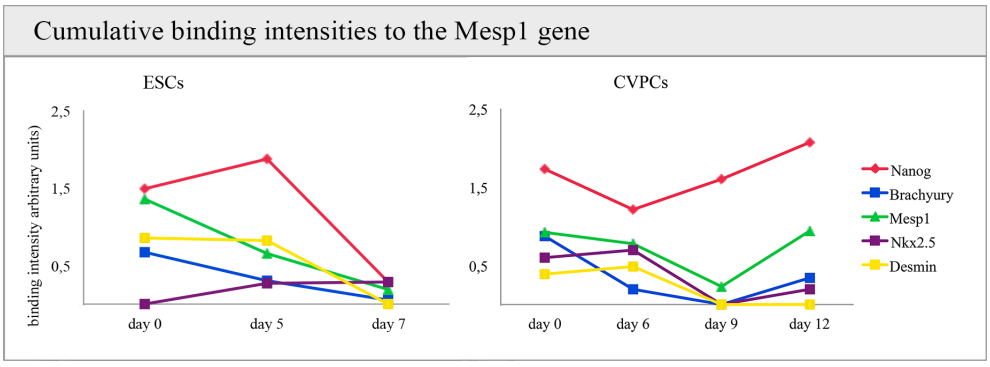


Figure 4.16. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the Mesp1 gene in a) ESCs and b) CVPCs over time. For criteria and method applied see Figure 4.13.

In ESCs, the cumulative binding intensities of Brachyury, Mesp1 and Desmin declined over the course of differentiation. Nkx2.5 did not occupy the Mesp1 gene in undifferentiated ESCs, but its binding intensity increased slightly on days 5 and 7 of ESC differentiation. In ESCs Nanog occupied the Mesp1 gene with the highest cumulative binding intensity on day 5 of differentiation and almost disappeared from the Mesp1 gene on day 7 of ESC differentiation. Linear regression analysis predicted a linear decline for the binding intensities of Brachyury (R= -0,99) and Mesp1 (R= -0,99) to the Mesp1 gene over the course of ESC differentiation, and a linear increase for the binding intensity of Nkx2.5 (R= 0,98).

In CVPCs, Brachyury, Mesp1, Nkx2.5 and Desmin showed a similar trend, with a drop in cumulative binding intensity on day 9 of differentiation. Brachyury, Mesp1 and Nkx2.5 binding intensity increased again on day 12, while Desmin did not seem to bind the Mesp1 gene in CVPCs beyond somewhere between days 6 and 9. Nanog on the other hand was present much more strongly than the other TFs and its cumulative binding intensity increased steadily after a drop on day 6. It can be excluded that the stronger Nanog binding was only a product of different antibody efficiencies, because this phenomenon could not be observed on any other gene.

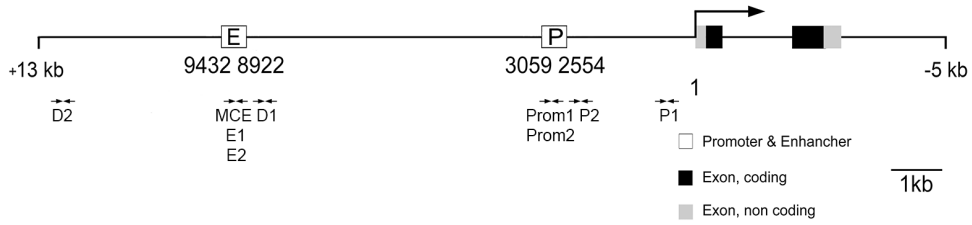
4.2.5. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nkx2.5 Gene

4.2.5.1. In Silico Analysis of Putative Transcription Factor Binding Sites and Primer Design

Primer pairs for the 5´UTR of the Nkx2.5 gene were designed by Christiane Fuchs and Sonja Gawlas in the course of their diploma thesis projects. The primers were designed to cover the described promoter and minimal cardiac enhancer (MCE) segment of the 5'UTR of the Nkx2.5 gene (Lien et al., 1999; Searcy et al., 1998), as well as several controls adjacent to the described regulatory DNA segments, where no TF binding was expected. Amplified DNA segments of the Nkx2.5 gene are listed in Table 4.4.

For the promoter and MCE segment two partly overlapping primer pairs each (E1 & E2 and Prom1 & Prom2, respectively) were used, but in the following illustrations only the stronger binding signal is shown for reasons of clarity and comprehensibility. The MCE primer pair was designed subsequently to make sure that the E1 and E2 primer pairs, which showed relatively low efficiency in PCR reactions, did not miss any binding signal.

Table 4.4. Location of putative transcription factor binding sites containing DNA segments of the Nkx2.5 gene. Arrows indicate the position of the primer pairs.



Putative binding sites in DNA segments amplified or in close proximity for the Nkx2.5 gene				
Name/ Extend of the amplified DNA segment [1]	Nanog	Brachyury	Mesp1	Nkx2.5
P1 (-374 to -635)		$(AC)_3$ (**-382**)[2] GGTG (**-438; -498**)	CATTTG (-317)[3] CACCTG (-792)	TAAT (-254; -687) TGAAGTG (-762)
P2 (-2103 to -2475)	TAATTG (**-2191**) TAATTT (-2528)	GGTG (**-2437**; -2570)	CAGGTG (-2497)	ATTA (**-2302**; **-2327**), TAAT (**-2191; -2383**) CACTT (**-2450**)
Prom2[4] (-2933 to -3187) Prom1 (-3143 to -3387)	AAATTA (-2857) TAATTG (**-2936**) ACATTA (**-3166**) TAATTG (**-3219**)	$(GT)_3$ (-2747; -2757; -2767) GGTG (-2896) ACACCT (-3508)	CAGATG (**-3070**) CACCTG (-3444)	ATTA (-2814; -2859; **-3168**) ATTAAT (**-2936**; **-2975**) CACTT (**-3173**) TAAT (**-3219**)
D1 (-8792 to -9123)	TAATGT (**-8949**)	GGTG (-8668; -8732)	CAGCTG (-8572)	AAGTG (-8753) TAAT (**-8849**)
E2[4] (-9280 to -9609) E1 (-9356 to -9707) MCE (-9258 to -9574)	CAATTA (-9209)	GGTGT (-8980) GGTG (-9001; -9062; **-9283**) AGGTG (-9118) CACACC (-9204)	CAGCTG (-8572)	AAGTG (-8934) ATTA (-9211) TAAT (**-9467**)
D2 (-12620 to -13041)	TAATGT (**-12753**; **-12779**)	ACACCT (-12598) GGTG (**-12657**)		TAAT (**-12642**; **-12692**; **-12753**; **-12779**)

[1] Primers are located at the very beginning and end of the denoted DNA segment (for primer sequence see Table 2.1). Numbers depict base pairs in relation to the transcription initiation site.

[2] Bold = binding site lies within the amplified DNA segment; [3] Regular = binding site lies adjacent to the amplified DNA segment

[4] Putative binding sites in the promoter (Prom1 and Prom2) and MCE (E1, E2 and MCE) segments, respectively, were summarized, because the amplified segments of the primer pairs overlap.

4.2.5.2. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nkx2.5 Gene in ESCs

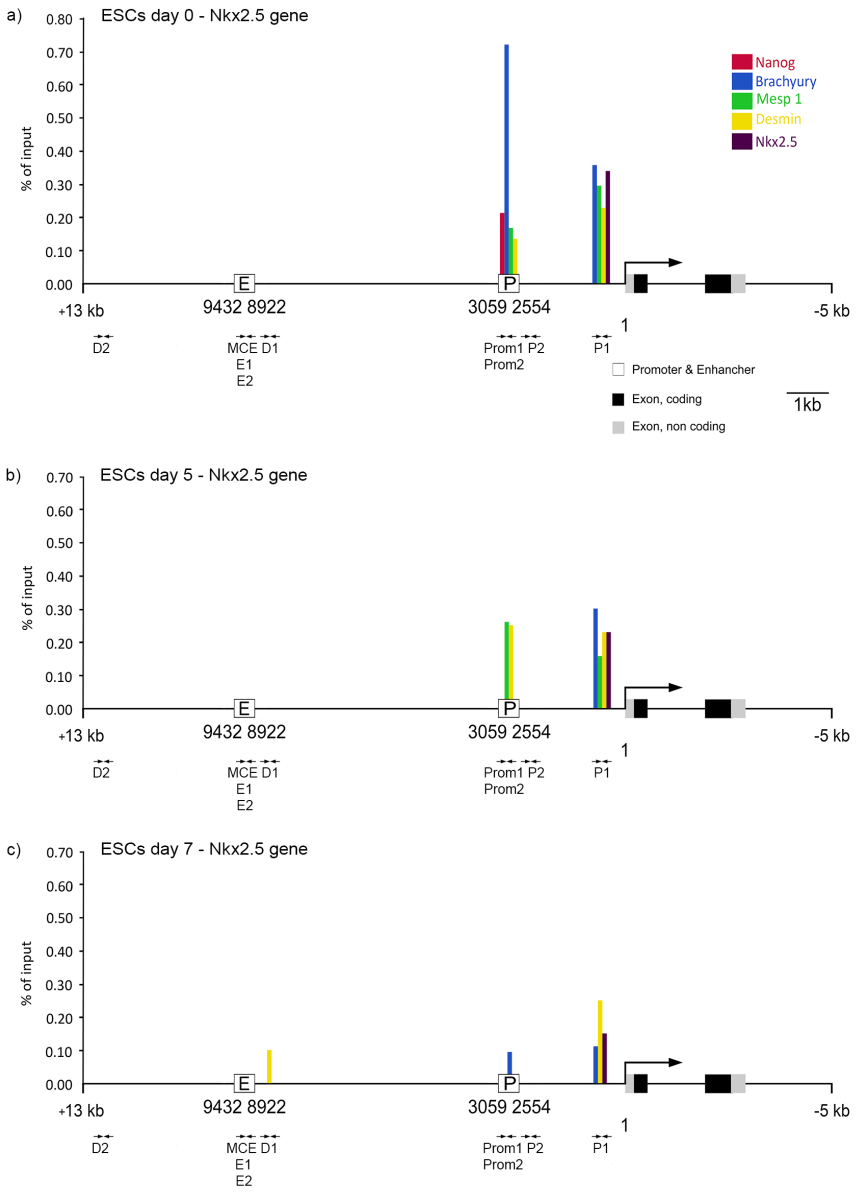


Figure 4.17. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nkx2.5 gene. In a) undifferentiated ESCs, b) on day 5 and c) on day 7 of ESC differentiation, illustrated in

a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated ESCs only the described promoter (Prom1 and Prom2) and the P1 DNA segment were occupied by the investigated TFs. Notable is the very strong binding signal of Brachyury to the promoter. Nkx2.5 should not be expressed until at least day 5 of EB differentiation (Kim et al., 2009), hence, it could be speculated that strong Brachyury binding inhibits activation of the Nkx2.5 gene. The P1 DNA segment lies in proximity to the transcription initiation site; therefore it would not be surprising to find a regulatory element in this area. The control experiments done with the P1 primer pair showed no contamination with DNA and it can be assumed that the binding signal was specific. The other control DNA segments P2, D1 and D2 were not occupied.

On day 5 of ESC differentiation binding signal frequency and intensity of the investigated TFs to the Nkx2.5 gene generally decreased when compared to undifferentiated ESCs. Only the binding signal intensities of Mesp1 and Desmin to the described promoter (Prom1 and Prom2) increased. Nanog and Brachyury no longer occupied the promoter (Prom1 and Prom2) segment. All four binding signals at the P1 DNA segment were weaker than in undifferentiated cells, but essentially did not change. Again no binding to the MCE (E1, E2 and MCE) could be observed. Possibly other not investigated DNA binding factors occupy the MCE segment during the early stages of development. It was shown that Gata factors bind this MCE and thus regulate Nkx2.5 gene expression at least until activity of the MCE becomes restricted to the right ventricle at E11.5 and other factors are suspected to be needed to guarantee tissue specific Nkx2.5 expression (Lien et al., 1999). NFAT was also shown to bind the MCE (Chen & Cao referred to the DNA segment as AR1) and activate Nkx2.5 gene expression, possibly together with Gata (Chen and Cao, 2009).

On day 7 of ESC differentiation the binding of almost all investigated TFs diminished. Only Desmin was still bound as strongly to the P1 DNA segment on day 7 as on day 5 of ESC differentiation. These data suggest that other factors might be more important for Nkx2.5 regulation at this point in development.

4.2.5.3. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nkx2.5 Gene in CVPCs

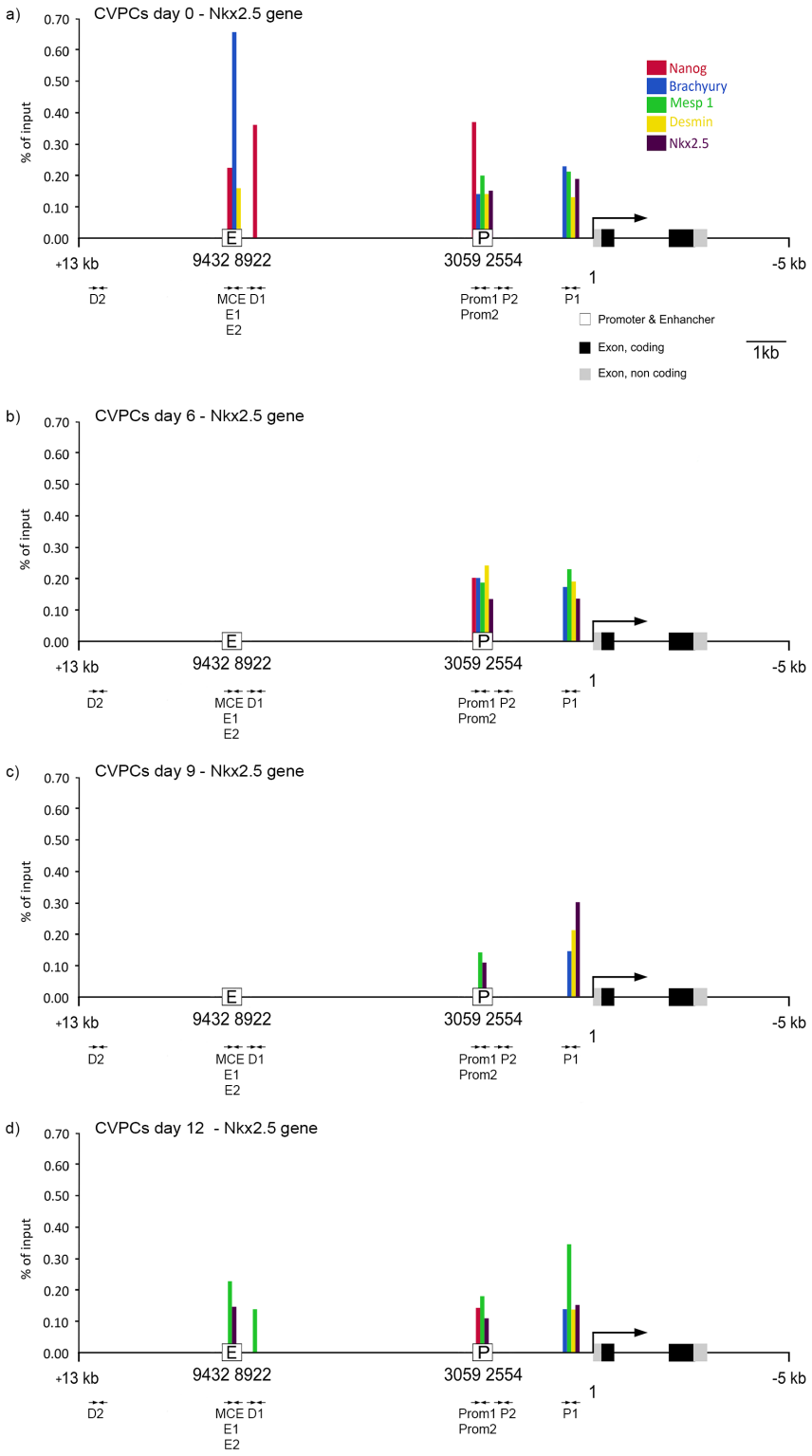


Figure 4.18. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nkx2.5 gene. In a) undifferentiated CVPCs, b) on day 6, c) day 7 and d) day 12 of CVPC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated CVPCs all five investigated TFs occupied the 5´UTR of the Nkx2.5 gene. The strongest binding signals could be observed at the MCE (E1, E2 and MCE), with Nanog, Brachyury and Desmin occupying this DNA segment. The described promoter (Prom1 and Prom2) was occupied by all five TFs, but the obtained signal intensity was not very high. Like in ESCs, Nanog did not occupy the P1 DNA segment. The same four TFs occupied the P1 DNA segment in undifferentiated CVPCs and ESCs, but the binding signal intensity was noticeably weaker in CVPCs.

On day 6 of CVPC differentiation the described promoter (Prom1 and Prom2) and the P1 DNA segment were occupied by the same TFs as in undifferentiated CVPCs, but the intensity of Nanog's binding signal to the promoter was markedly weaker than in undifferentiated CVPCs. The signals at the MCE disappeared completely. Again other not investigated DNA binding factors might occupy the MCE segment during this time in development (see 4.2.5.2).

On day 9 of CVPC differentiation binding signal frequency and intensity decreased greatly. Especially the described promoter (Prom1 and Prom2) was occupied a lot less frequently than before, with only Mesp1 and Nkx2.5 occupying this DNA segment. Only the binding signal intensity of Nkx2.5 to the P1 DNA segment increased when compared to day 6 of CVPC differentiation. This trend was comparable to the decline in binding signal frequency and intensity in ESCs on day 7 of differentiation.

On day 12 of CVPC differentiation there was again an increase in binding signal frequency and intensity when compared to day 9 of CVPC differentiation. Notable was also the dominant presence of Mesp1 at the MCE (E1, E2 and MCE) and promoter (Prom1 and Prom2), as well as at the P1 DNA segment. Nkx2.5 is a known target of Mesp1 (Bondue and Blanpain, 2010), therefore Mesp1 might become more and more important for Nkx2.5 regulation as development progresses. Besides day 0 of CVPC differentiation, day 12 was the only investigated time point when the Nkx2.5 MCE (E1, E2 and MCE) was occupied by any of the analyzed TFs. Interestingly it was no longer Nanog, Brachyury and Desmin that occupied this DNA segment, but Mesp1 and Nkx2.5.

Nanog generally seemed to bind the Nkx2.5 gene less frequently as differentiation progressed.

4.2.5.4. Résumé of the Protein Interaction with the Nkx2.5 Gene

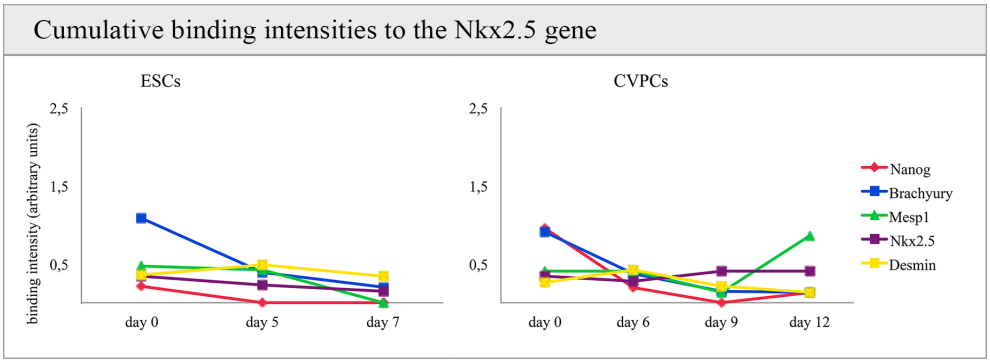


Figure 4.19. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the Nkx2.5 gene in a) ESCs and b) CVPCs over time. For criteria and method applied see Figure 4.13.

In ESCs, the cumulative binding intensities of Nkx2.5 and Desmin to the Nkx2.5 gene stayed more or less the same. The binding intensities of Nanog, Brachyury and Mesp1 appeared to decline, but only Brachyury displayed a clear change in binding intensity. Linear regression analysis predicted a linear decline for the binding intensity of Brachyury (R= -0,99) to the Nkx2.5 gene over the course of ESC differentiation.

In CVPCs, the cumulative binding intensities of Nanog, Brachyury and Mesp1 also decreased over time, but unlike in ESCs, the binding of Mesp1 increased again after day 9. We can only speculate that the same would be true for ESCs, because we did not investigate ESCs beyond day 7. The more or less constant cumulative binding intensity of Nkx2.5 to its own gene in both ESCs and CVPCs might be evidence of the described negative feedback loop, which ensures tight control of Nkx2.5 gene expression during development (Tanaka et al., 1999).

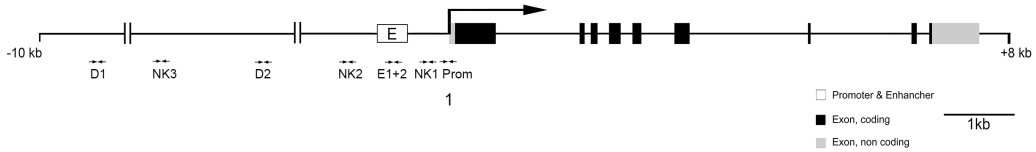
4.2.6. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Desmin Gene

4.2.6.1. In Silico Analysis of Putative Transcription Factor Binding Sites and Primer Design

An element in close proximity to the transcription initiation site is considered to act as the Desmin promoter and was shown to be sufficient for muscle specific Desmin expression,

and a classical enhancer further upstream was shown to be necessary for high levels of Desmin expression (Li and Capetanaki, 1993; Li and Capetanakil, 1994). Primer pairs were designed for these described regulatory DNA segments (Prom, E1 and E2), as well as other DNA segments with TF binding site accumulation as identified by in silico analysis of the 5'UTR of the Desmin gene. Amplified DNA segments of the Desmin gene are listed in Table 4.5.

Table 4.5. Location of putative transcription factor binding sites containing DNA segments of the Desmin gene. Arrows indicate the position of the primer pairs.



Putative binding sites in DNA segments amplified or in close proximity for the Desmin gene				
Name/ Extend of the amplified DNA segment[1]	Nanog	Brachyury	Mesp1	Nkx2.5
Prom (+27 to -185)			CAGCTG[2] (**-79**)[3]	CACTT (-380)[4]
NK1 (-203 to -427)			CATGTG (-534)	CACTT (**-380**)
E1 (-577 to -779)		GGTGTGA (**-753**; **-772**) GGTGT (**-762**)	CATGTG (-534)	CACTT (-380)
E2 (-833 to -986)		ACACCT (**-954**) CACACC (-1023)	CAGCTG[5] (**-832**) CAGGTG[6] (**-936**)	TAAT (-1162)
NK2 (-1233 to -1449)	TAATGT (-2000)	GGTGT (**-1370; -1418**)		TAAT (-1202) ATTA (**-1385**)
D2 (-4713 to -5055)	TAATTG (-5057)	GGTGT (-5154)	CACATG (-5204)	CACTT (-4657) TAAT (**-4986**; -5057)
NK3 (-6331 to -6531)	TAATG (-6135) TAATTT (-6998)	ACACACC (**-6335**) GGTGT (**-6395**)	CACTTG (**-6407**) CACCTG (**-6456**)	CACTT (**-6407**) ATTA (**-6478**)
D1 (-9102 to -9350)	TAATTT (**-9264**)	ACACC (**-9123**) GGTGTGT (**-9237**)	CACTTG (-9476)	TAAT (**-9198**; **-9264**; **-9292**)

[1] Primers are located at the very beginning and end of the denoted DNA segment (for primer sequence see Table 2.1). Numbers depict base pairs in relation to the transcription initiation site.

[2], [5], [6] published potential cis-acting E boxes (Li and Capetanaki, 1993)

[3] Bold letter = binding site lies within the amplified DNA segment; [4] Regular letter = binding site lies adjacent to the amplified DNA segment

For the described enhancer, two primer pairs (E1 & E2) were designed to cover the whole DNA segment, but in the following illustrations only the stronger signal is shown for reasons of clarity and comprehensibility. The primer pairs labeled with NK were meant to act as negative controls, where no TF binding was expected. Like with the Brachyury gene, the DNA segment in close proximity to the transcription initiation site contains a high percentage of C and G nucleotides, making it difficult to design primers that function as efficiently as those located in other DNA segments.

4.2.6.2. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Desmin Gene in ESCs

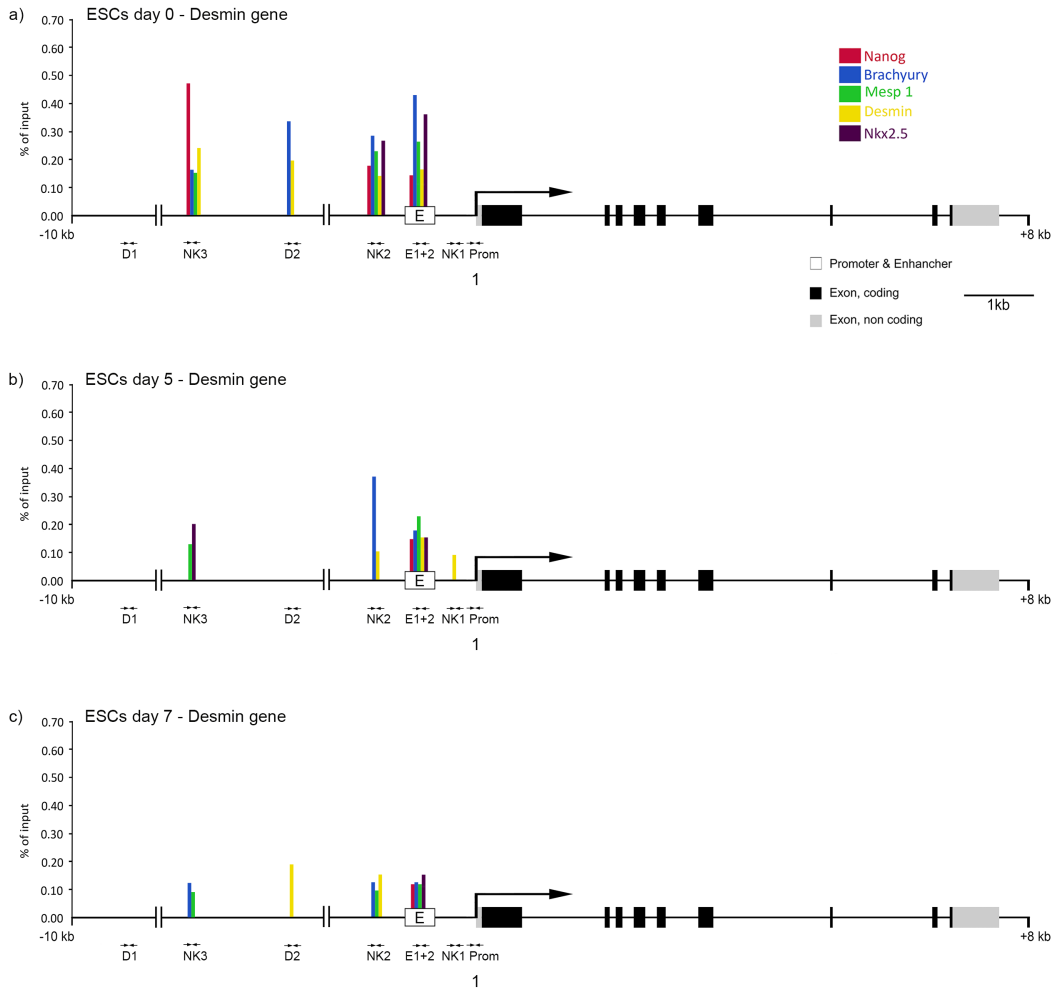


Figure 4.20. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Desmin gene. In a) undifferentiated ESCs, b) on day 5 and c) on day 7 of ESC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated ESCs the Desmin gene was bound quite strongly and frequently by all five investigated TFs. Especially the NK3, NK2, E1 and E2 DNA segments had high signal intensities of TFs. The strong occupation of the NK3 and NK2 DNA segments was at first surprising, but further analysis of the DNA sequence identified several at least partial DNA binding sites in these segments. The D2 DNA segment was also occupied, but only by Brachyury and Desmin. The Prom, NK1 and D1 DNA segments were not occupied in undifferentiated ESCs. The lack of signals in the Prom DNA segment could be evidence that the Desmin gene is not transcribed until approximately day 8.5 of differentiation (Li and Capetanaki, 1993). Another explanation could of course be that the investigated TFs simply do not occupy the Desmin Prom DNA segment.

On day 5 of ESC differentiation not only the binding signal frequency of almost all investigated TFs, but also the intensity of the obtained binding signals decreased. Only the enhancer (E1 and E2) was still bound by all five investigated TFs, but also with decreasing intensity. The binding to the D2 DNA segment vanished completely and the also rather distal NK3 DNA segment was occupied a lot less frequently than in undifferentiated ESCs.

On day 7 of ESC differentiation the number of detected binding signals was similar to that of day 5 of ESC differentiation. The binding signal intensity on the other hand generally decreased when compared to day 5. The enhancer segment (E1 and E2) was again occupied most frequently, with four out of five TFs occupying this DNA segment on day 7 of ESC differentiation.

4.2.6.3. The Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Desmin Gene in CVPCs

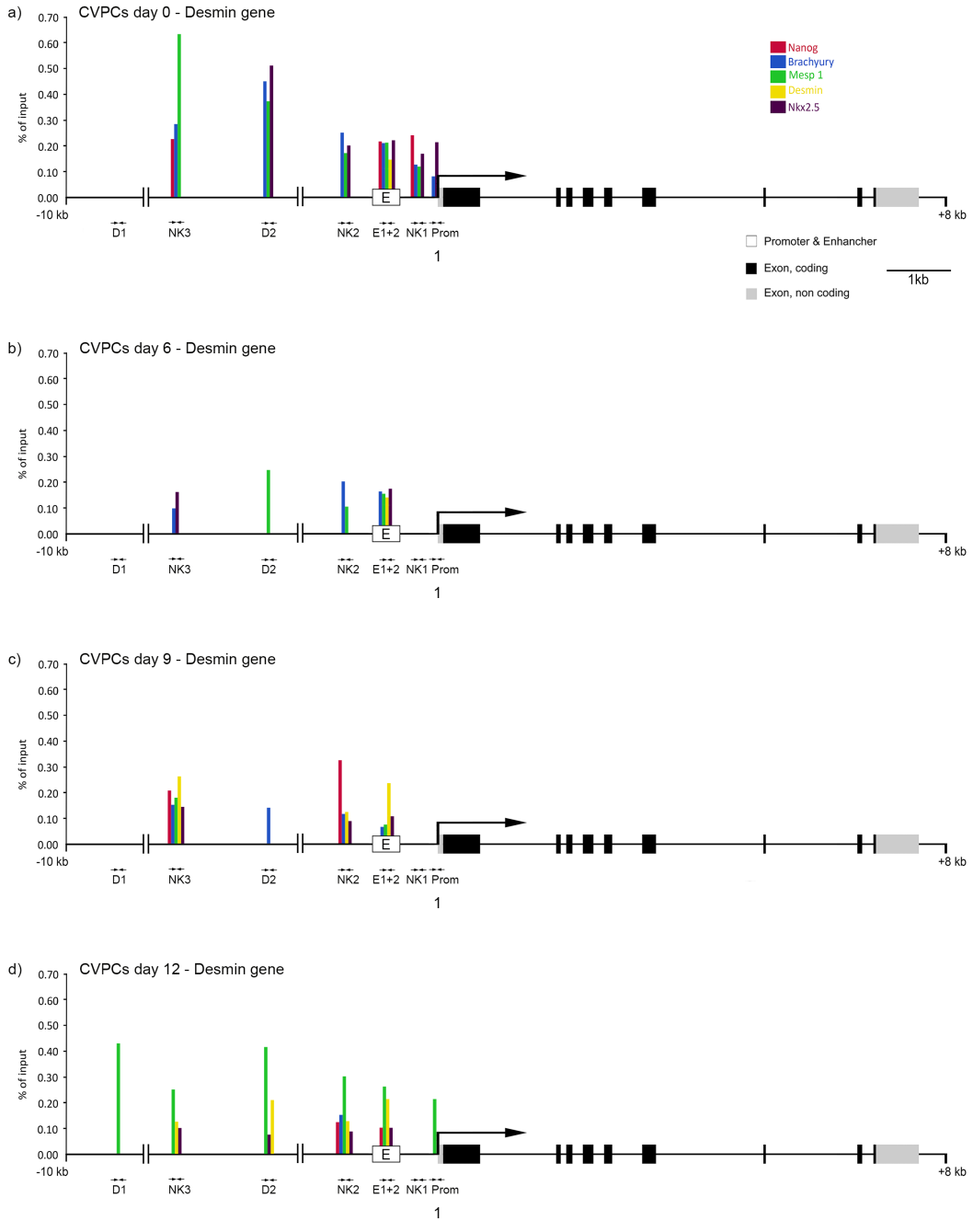


Figure 4.21. The interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with the Nkx2.5 gene. In a) undifferentiated CVPCs, b) on day 6, c) day 7 and d) day 12 of CVPC differentiation, illustrated in a true to scale gene map. Bar height represents the mean percentage of input from all experiments. Arrows mark the position of the primer pairs used for PCR analysis.

In undifferentiated CVPCs the Desmin gene was bound significantly more strongly and frequently than in ESCs, supporting the assumption that Desmin plays an important role in cardiomyogenesis. CVPCs are already specialized and only give rise to cells of the mesoderm. Therefore, it would stand to reason that Desmin is more important in cells derived from CVPCs than from ESCs, which give rise to cells of all three germ layers. In contrast to ESCs, even undifferentiated CVPCs were shown to express Desmin (Hoebaus et al., 2013), which is supported by the fact that the described Desmin promoter (Prom) is occupied by Brachyury and Nkx2.5. Surprisingly the NK1 DNA segment was also occupied quite frequently and the D2 DNA segment showed three of the strongest signals. Desmin only occupied the enhancer (E1 and E2) segment with rather weak binding signal intensity. Mesp1 occupied the Desmin gene very frequently in undifferentiated CVPCs and most of the obtained binding signals displayed high binding signal intensity.

On day 6 of CVPC differentiation the binding signal frequency and intensity to the Desmin gene decreased. The Prom and NK1 DNA segments were no longer occupied and the binding to more distal DNA segments also decreased visibly. Nanog did not occupy the Desmin gene on day 6 of CVPC differentiation. Like in undifferentiated CVPCs, Desmin only occupied the enhancer (E1 and E2) segment. The binding signal intensities of Brachyury and Mesp1 to the Desmin gene also decreased on day 6 when compared to undifferentiated CVPCs, but they still each occupied several DNA segments.

On day 9 of CVPC differentiation more proteins bound to the Desmin gene than on day 6 of differentiation, but binding signal intensity was not higher on day 9 than on day 6. Especially Nanog and Desmin occupied several DNA segments on day 9, which they had not occupied on day 6. The NK3 and NK2 DNA segments were occupied even more frequently than in undifferentiated cells. The enhancer (E1 and E2) segment was still occupied by Brachyury, Mesp1, Nkx2.5 and Desmin, but with rather low binding signal intensity.

On day 12 of CVPC differentiation binding signal frequency and intensity of Mesp1 to the Desmin gene increased drastically. Mesp1 was bound to all investigated DNA segments of the Desmin gene, except for the NK1 DNA segment. Desmin also occupied more DNA segments than on the previously investigated days of CVPC differentiation. Brachyury on the other hand occupied only the NK2 DNA segment on day 12 of CVPC differentiation and the binding signal intensity also decreased when compared to day 9 of CVPC differentiation. Binding signal frequency to the E1, E2 and NK2 DNA segments

remained pretty much constant, but more TFs interacted with the D2 DNA segment on day 12 than on day 9 of CVPC differentiation.

4.2.6.4. Résumé of the Protein Interaction with the Desmin Gene

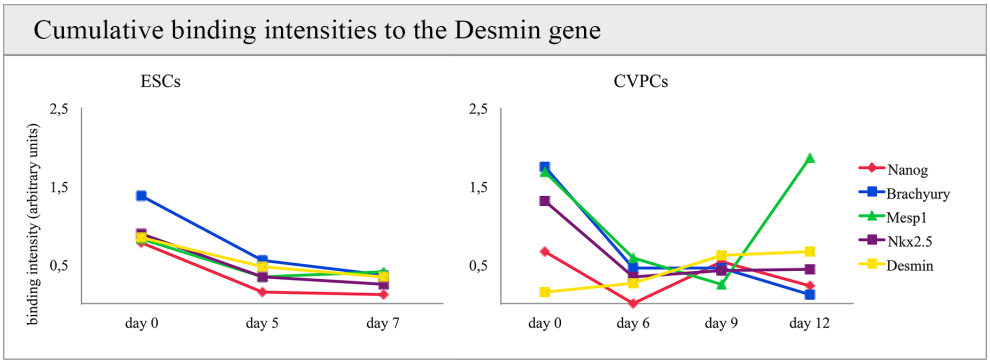


Figure 4.22. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the Desmin gene in a) ESCs and b) CVPCs over time. For criteria and method applied see Figure 4.13.

In ESCs, the cumulative binding intensities of all five investigated TFs declined steadily over the course of differentiation. Only Mesp1 showed a very slight increase in cumulative binding signal intensity on day 7 of ESC differentiation. It could be that Mesp1 binding signal intensity to the Desmin gene would increase strongly after day 7, but further investigations would be necessary to make a clearer statement. Desmin should not be expressed strongly until day 8.5 of differentiation (Li and Capetanaki, 1993). Therefore, it is possible that greater changes in the occupation of the Desmin gene by Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin could be observed beyond day 7 of ESC differentiation. Linear regression analysis predicted a linear decline for the binding intensities of Brachyury (R= -0,99) and Nkx2.5 (R= -0,99) to the Desmin gene over the course of ESC differentiation.

In CVPCs, the change in cumulative binding signal intensity of the investigated TFs over the course of differentiation was much more complex than in ESCs. On the first two investigated time points Nanog, Brachyury, Mesp1 and Nkx2.5 showed a trend similar to the one observed in ESCs: a clear decrease in cumulative binding signal intensity. Only the binding signal intensity of Desmin increased from day 0 to day 6 of CVPC differentiation, a trend very different from the one observed in ESCs. On day 9 the cumulative binding signal intensity of Mesp1 decreased further, but all other investigated TFs showed at least a slight increase. On day 12 of CVPC differentiation the cumulative

binding signal intensities of Nanog and Brachyury decreased again. Only the binding signal intensity of Mesp1 increased strongly on day 12, suggesting that Mesp1 might regulate the Desmin gene during this stage in differentiation.

4.2.7. Linear Regression Analysis of the Cumulative Binding Intensities of the Five Proteins to the Individual Genes Over Time

We did a linear regression analysis to determine if a linear relationship exists between the cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the five genes and the investigated time points of ESC and CVPC differentiation. For this we calculated the Pearson's correlation coefficient (R). For R, values between -1 and +1 are possible, with -1 signifying a clear linear decline and +1 signifying a clear linear increase, whereas a value around 0 indicates that there is no linear correlation between the two variables. Because merely 3 time points were investigated in ESCs and 4 in CVPCs, only values higher than 0.975 or lower than -0.975 were considered statistically conclusive. Values between 0.890 and 0.975 as well as between -0.890 and – 0.975 were interpreted as a weak linear correlation.

Table 4.6. Pearson coefficient (R) investigating the linear relationship between the cumulative binding intensities of the five proteins to each individual gene and the time of differentiation.

Gene	Protein	Pearson coefficient (R)	
		ESCs	CVPCs
Nanog	Na[(1)]	**-1.000[(6)]**	**-0.962**
Brachyury	Na	**-0.916**	**-0.925**
	T[(2)]	-0.804	-0.627
	M[(3)]	-0.110	-0.383
	N[(4)]	**-1.000**	-0.255
	D[(5)]	0.091	-0.016
Mesp1	Na	-0.542	0.341
	T	**-0.992**	-0.734
	M	**-0.991**	-0.267
	N	**0.979**	-0.701
	D	-0.749	-0.763
Nkx2.5	Na	**-0.961**	**-0.895**
	T	**-0.997**	**-0.965**
	M	-0.781	0.381
	N	**-0,989**	0.580
	D	0.108	-0.457
Desmin	Na	**-0.961**	0.341
	T	**-0.997**	-0.734
	M	-0.781	-0.267
	N	**-0.989**	-0.701
	D	0.108	-0.763

[(1)] Na, Nanog; [(2)]) T, Brachyury; [(3)] M, Mesp1; [(4)] N, Nkx2.5; [(5)] D, Desmin;
[(6)] bold letter = linear correlation

At the **Nanog gene**, linear regression analysis predicted a clear linear decline of Nanog binding intensity in ESCs (R= -1) over time, whereas in CVPCs, a not as pronounced linear decline could be observed (R= -0,962).

At the **Brachyury gene**, Nanog showed a weak linear correlation between cumulative binding signal intensity and time of differentiation in both ESCs (R=-0,916) and CVPCs (R= -0,925). For the binding intensity of Nkx2.5 a clear linear decline could be predicted over the course of ESC differentiation (R= -1), but not in CVPCs. All other proteins did not show a linear decrease or increase of binding signal intensity to the Brachyury gene over the course of differentiation.

At the **Mesp1 gene**, linear regression analysis predicted a clear linear decline for the binding intensities of Brachyury (R= -0,992) and Mesp1 (R= -0,991) and a clear linear increase for the binding intensity of Nkx2.5 (R= 0,979) to the Mesp1 gene over the course of ESC differentiation. For all other proteins no linear correlation between cumulative binding intensity to the Mesp1 gene and time of differentiation could be predicted.

At the **Nkx2.5 gene**, Brachyury had a clear linear decline of binding signal intensity over time in ESCs (R= -0,997) and a not as pronounced linear decline in CVPCs (-0,965). For Nanog a weak linear correlation between signal intensity and time of differentiation was found in ESCs (R= -0,961) and CVPCs (R= -0,895). For Nkx2.5 a clear linear decrease of binding signal intensity over time was found in ESCs (R= -0,989), but not in CVPCs. For all other proteins no linear correlation between cumulative binding intensity to the Nkx2.5 gene and time of differentiation could be predicted.

At the **Desmin gene**, linear regression analysis predicted a clear linear decline for the binding intensities of Brachyury (R= -0,997) and Nkx2.5 (R= -0,989) and a not as pronounced linear decline for the binding intensity of Nanog (R= -0,961) over the course of ESC differentiation. For all other proteins no linear correlation between cumulative binding intensity to the Desmin gene and time of differentiation could be predicted.

In summary it can be said that the cumulative binding intensities of Nanog, Brachyury and Nkx2.5 often showed a linear decrease or increase over the course of ESC differentiation, but only seldom over the course of CVPC differentiation. A linear decline of Mesp1 binding intensity over time could be predicted only once in ESCs. We never found a linear relationship between the cumulative binding intensity of Desmin and the time of differentiation.

5. Discussion

5.1. General Aspects of the Investigation of the Physical Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with Their Respective Genes

5.1.1. Applicability of EBs and CBs as in Vitro Differentiation Models for ChIP Analysis

All of the ChIP experiments were conducted under conditions as similar as possible. The EBs and CBs were generated with low-passage ESCs and CVPCs at approximately the same time of day and were rinsed off also maintaining the same time frame for all cells. Nevertheless, it must be assumed that there are always some slight variations in the differentiation speed and efficiency of EBs and CBs due to variations in exterior conditions, like room temperature and different batches of culture medium. Those differences are probably not detectable when looking at more general processes of differentiation, like the peaks of protein expression, but might be visible in ChIP experiments, which investigate dynamic and fast changing processes of transcriptional regulation. Furthermore, it should be considered that CBs and especially EBs do not generate a homogeneous cell population. EBs give rise to cells of all three germ layers, and even though CBs only generate cells of the mesoderm, they can still differentiate into either cardiomyocytes, endothelial cells or smooth muscle cells. Therefore, differentiation experiments with EBs and CBs will always reflect the situation of all cells in a population of several hundred EBs or CBs.

Brachyury and Mesp1 are markers for early mesoderm (Bondue and Blanpain, 2010; Evans et al., 2012) and Nkx2.5 is a marker for the cardiac lineage, although recent studies suggest that Nkx2.5 positive cells can also give rise to endothelial and smooth muscle lineages (Behrens et al., 2013). Desmin is a marker for muscle-derived cells and, thus, also for mesoderm (Meng et al., 2010; Torres et al., 2012). Therefore, it is presumable that all four proteins are only expressed in cells of the mesoderm and are silenced in cells which differentiate into other cell lineages. This long-term silencing is most likely achieved through the presence of heterochromatin at the loci of the silenced genes (Efroni et al., 2009; Golob et al., 2011); hence, no or very little TF binding should take place at

these tightly packaged loci. Moreover, no or very little Brachyury, Mesp1, Nkx2.5 and Desmin protein should be present in cells of other lineages than the mesoderm. Thus, we may assume that these proteins cannot be pulled down with ChIP. Hence, the other cell lineages present in EBs and CBs should not influence the data obtained with anti-Brachyury, anti-Mesp1, anti-Nkx2.5 and anti-Desmin antibodies in the ChIP experiments. Nanog should be expressed in all undifferentiated stem cells, but it was shown that Nanog levels fluctuate greatly on a single cell level (MacArthur et al., 2012). Therefore, not only cells with high Nanog levels but also cells with a short-lived Nanog downregulation will be cross-linked with formaldehyde during the ChIP procedure. Cells with Nanog downregulation are in a state primed for differentiation without definite commitment (MacArthur et al., 2012). Those cells probably exhibit different TF occupation patterns at genes like Brachyury or Mesp1, which are expressed in early mesodermal cells shortly after the differentiation process starts. Furthermore, undifferentiated ESCs always harbor small side populations of already differentiated cells, which are, for example, already Brachyury positive. These at least partially differentiated cells also exhibit TF binding events, which are not typical of undifferentiated ESCs or CVPCs. Nevertheless, the percentage of differentiated cells should be small enough for results obtained with ChIP to mostly reflect the transcriptional network present in undifferentiated stem cells. Not much is known about Nanog's role during differentiation, which is why it is impossible to say for sure in which cell types Nanog is expressed after differentiation starts. We could, however, show with western blot that the overall Nanog protein levels stay more or less the same up until and including day 18 of both ESC and CVPC differentiation (Figure 4.1). This could indicate that all cell lineages keep expressing Nanog throughout early differentiation. Therefore it cannot be completely excluded that in ESCs some of the observed binding of Nanog to the Brachyury, Mesp1, Nkx2.5 and Desmin gene might reflect silencing of the genes by Nanog in endodermal and ectodermal cell lineages. This is, however, unlikely, as long-term silencing in differentiating stem cells is mostly achieved through heterochromatin (Efroni et al., 2009; Golob et al., 2011).

5.1.2. Significance of ChIP Data

ChIP is a valuable and well-established method for investigating the physical interaction of a TF with genomic DNA. Nevertheless, there are still several aspects which need to be considered when analyzing the generated data.

The large amount of cells necessary for a successful ChIP experiment inevitably results in a very heterogeneous population of cells in each conducted ChIP experiment (see 5.1.1). Especially undifferentiated mESCs are highly heterogeneous regarding their overall expression profile (MacArthur et al., 2012). They also proliferate rapidly and, therefore, cells of all stages of the cell cycle are present in one cell population. Data obtained by ChIP can only produce a snap shot of the dynamic and complex interactions taking place and, therefore, reflect the momentarily most common state. Signals with the highest intensity most likely mirror a binding event which is most common in the overall cell population because binding signal intensity is dependent on the number of DNA segments pulled down in the ChIP experiments. The more DNA segments are pulled down, the more DNA templates are available for the primer pairs to amplify in a PCR set-up.

Crosslinking proteins and DNA with formaldehyde presents another issue: it can never be excluded that the investigated TF is actually part of a larger DNA binding protein complex and does not directly bind the investigated DNA segment. It is not possible to only crosslink proteins which are directly bound to the DNA double helix, because formaldehyde crosslinks primary amino groups in proteins nonselectively with close-by nitrogen atoms in either other proteins or DNA via a $-CH_2-$ linkage (Fraenkel-Conrat and Olcott, 1948; Orlando, 2000).

Furthermore, it is not possible to determine the exact interaction site of a TF on the DNA sequence due to the limited resolution of ChIP experiments (Carey et al., 2009). The resolution is highly dependent on the size of the DNA fragments produced by sonication, but even if the chromatin is sheared into fragments of a more or less homogeneous and ideally suited size, the exact location of an interaction can only be narrowed down to an area of several hundred base pairs.

It is also important to be aware of the fact that ChIP is not a functional assay. ChIP alone cannot demonstrate the functional significance of a protein-DNA interaction. In recent years evidence has been found that proteins occupy genomic DNA segments at which they have no obvious function (Carey et al., 2009). It was even proposed that there might not be a positive selection pressure to maintain binding sites and that many TF binding sites might occur at random without having a physiologically relevant function. Therefore, many TF binding events detected with ChIP might reflect “noise” in regulation due to the random distribution of TF binding sites (Li and Johnston, 2001). It was also suggested that at least in Drosophila only highly bound genes are actual targets

of the investigated TF, whereas genes bound at comparably lower levels might not be regulated by the bound TF (Li et al., 2008). This, however, is a relatively new theory and does not reflect the general belief.

Nevertheless, ChIP is a valuable tool, which can provide the foundation for further analysis of the interaction of TFs with specific DNA segments.

5.1.3. The Influence of the Status of Chromatin on ChIP

In eukaryotic cells two types of chromatin have been characterized: euchromatin, which is accessible and permissive for gene activation, and heterochromatin, which is more tightly packaged and for the most part silenced. Most stem cells were shown to have an open chromatin mostly comprised of euchromatin (Gaspar-Maia et al., 2009). Especially in mESCs the chromatin is transcriptionally globally hyperactive and becomes increasingly heterochromatic as cells start to differentiate. In the undifferentiated state, tissue-specific genes were reported to be expressed sporadically at low levels (Efroni et al., 2009). The Brachyury gene was found to have an intermediate level of both euchromatin- and heterochromatin-associated modifications in undifferentiated ESCs. When directed differentiation to pre-cardiac mesoderm is induced, the Brachyury locus becomes active and shows high levels of tri-methyl lysine 4, which is a sign for euchromatin. After gene silencing, the Brachyury locus was found to be in a heterochromatic state. These findings indicate that during ESC differentiation a genome-wide commitment to euchromatin or heterochromatin takes place (Golob et al., 2011).

These findings may also explain the great number of TF binding signals we obtained from undifferentiated ESCs and CVPCs. If the locus of a tissue-specific gene has intermediate levels of both euchromatin- and heterochromatin-associated modifications in undifferentiated stem cells, this locus may be occupied more strongly in ChIP experiments than the same locus during differentiation. In differentiating cells two scenarios are likely: (1.) The locus is inactivated through heterochromatin and, therefore, not easily accessible for TFs. Hence, fewer signals would be expected in ChIP experiments. (2.) The locus is activated through euchromatin and, therefore, probably more efficiently sheared by sonication. It stands to reason that euchromatic segments of the chromatin, which are not as compact and tightly packaged as heterochromatic segments, break more easily when submitted to sonication. This more efficient sonication of euchromatin may lead to chromatin fragments, which are too short to be amplified with PCR.

In summary it can therefore be said that the transcriptionally hyperactive state and the intermediate levels of both euchromatin- and heterochromatin-associated modifications of undifferentiated mESCs might lead to signals in ChIP which are even stronger than the ones obtained on days when the gene is actively transcribed.

Sonication of the chromatin is a key step in classical ChIP experiments. Primer pairs analyzing the interaction of a TF with a certain gene typically amplify DNA segments which are between 100 and 400 bp long. If the chromatin fragments, after sonication, are too short, no signals can be obtained with PCR analysis. If the chromatin fragments are too long, false positive signals are likely.

The fact that certain DNA segments might break more easily than others might also present a problem. If such a hot spot for sonication-induced DNA double strand breaks is in the amplified DNA segment of a primer pair, it will prove difficult to generate a signal with PCR analysis, because there will be far less DNA segments which can be amplified with PCR than at other sites. A weak PCR signal of the input sample generated with a primer pair, which produces only one specific signal of the correct size, might be an indication for such a hot spot for sonication-induced DNA double strand breaks. Therefore, all primer pairs used for the analysis of ChIP samples should also be tested with unsheared genomic DNA.

Different sonication efficiencies of undifferentiated and differentiating stem cells present another issue. We observed much higher sonication efficiency in undifferentiated stem cells, resulting in distinctly shorter DNA fragments after sonication when the same sonication protocol was applied to undifferentiated stem cells and differentiating EBs or CBs. As discussed above, the reason for this difference might lie in the different organization and compaction of the genome. In order to obtain DNA fragments of similar size in both undifferentiated and differentiating stem cells, the sonication protocol had to be optimized (Table 5.1).

Table 5.1. Optimized sonication protocol for ESCs and CVPCs.

Differentiation state and cell line	Dot cycle	Output	Time (s)	Number of cycles
Undifferentiated ESCs and CVPCs	90 %	45 %	15”	4-5
EBs and CBs	90 %	45 %	15”	6

5.1.4. Reliability of the PCR Reactions Used for the Analysis of the ChIP Samples

Ideally all primer pairs investigating the interaction of proteins with a certain gene should generate signals with the input sample of the same intensity. This proved difficult due to the immense number of primer pairs necessary for investigating the interaction of the proteins with five different genes (9 to 18 primer pairs per gene) and the fact that the presence of highly repetitive DNA segments in the tested genes complicated DNA amplification with PCR. Unfortunately, the detection of TF binding to highly repetitive DNA segments still remains a challenge with all ChIP protocols (Chung et al., 2011; Furey, 2013). The majority of primer pairs, however, did generate signals with the input samples of a very similar intensity. In some cases, when PCR amplification of individual DNA segments proved difficult, we chose two or more overlapping primer pairs to minimize the chances of missing binding signals of proteins to this DNA segment. The signal intensity of the individual primer pairs with the input samples, which they displayed in the majority of the performed PCR reactions, was categorized into high intensity (++), average intensity (+) and low intensity (+/-) signals. The specificity of the individual primer pairs with the input samples, which they displayed in the majority of the performed PCR reactions, was categorized into high specificity (one signal = ++), average specificity (two or three signals = +) and low specificity (four or more signals = +/-). Very weak unspecific signals were not considered, especially if the signal at the correct position had high intensity. Primer pairs with their determined signal intensity and specificity are listed in Table 5.2. Primer pairs with a high intensity of the input sample signal are most likely to generate reproducible signals because even small amounts of template DNA can be amplified reliably and efficiently. High specificity of the input sample signal, which contains DNA templates representing the whole genome, lets us assume that a signal obtained with the pull-down sample, which contains a small percentage of the DNA templates present in the input sample, is also highly specific.

Table 5.2. Signal intensity and specificity of the primer pairs used for PCR analysis of the ChIP samples.

Gene	Name/ Extend of the amplified DNA segment[(1)]	Intensity of the input sample signal[(2)]	Specificity of the input sample signal[(3)]
Nanog	E1 (-8656 to -8809)	+	+
	D0 (-6115 to -6282)	+	+/-

	D1 (-5085 to -5281)	++	+
	D2 (-5046 to -5189)	+	+
	D3 (-4657 to -4774)	+	++
	D4 (-4021 to -4168)	+	+
	P0 (-1079 to -1247)	+	+
	P1 (-347 to -470)	+	++
	I1 (+738 to +1130)	+	+/-
	I2 (+1137 to +1277)	+/-	+
	I3 (+1849 to +2123)	+/-	+
	I4 (+2105 to +2210)	+	+/-
	I5 (+2189 to +2441)	+	++
	I6 (+2443 to +2580)	+	+/-
	U1 (+7918 to +8200)	++	+
	U2 (+9110 to +9219)	+	+
	U3 (+9880 to +9987)	+/-	+
	U0 (+11396 to +11648)	++	++
Brachyury	P3 (-109 to -356)	+/-	+
	Nk1 (-771 to -986)	+	++
	P2 (-1480 to -1625)	+	+
	P1 (-2144 to -2269)	+	+
	P0 (-2468 to -2694)	++	++
	D3 (-3353 to -3534)	+	++
	D2 (-6288 to -6389)	+	++
	D1 (-7567 to -7667)	+	+
	D0 (-9209 to -9338)	+	+
	NK2 (-10270 to -10401)	+	++
Mesp1	3’UTR (+3832 to +4201)	++	+
	T1 (-19 to -169)	++	+
	P1 (-903 to -1254)	+	+
	T2 (-2329 to -2550)	++	++
	T3 (-3835 to -3983)	++	+
	D1 (-5087 to -5437)	++	++
	NK (-6164 to -6464)	+	++
	D2 (-8283 to -8459)	+	+
	D3 (-9225 to -9434)	++	+
Nkx2.5	P1 (-374 to -635)	+	+
	P2 (-2103 to -2475)	+	++
	Prom2 (-2933 to -3187)	+	++
	Prom1 (-3143 to -3387)	+	+
	D1 (-8792 to -9123)	++	++
	E2 (-9280 to -9609)	+/-	++

	E1 (-9356 to -9707)	+/-	++
	MCE (-9258 to -9574)	+	+
	D2 (-12620 to -13041)	+/-	+/-
Desmin	Prom (+27 to -185)	+/-	++
	NK1 (-203 to -427)	+	+
	E1 (-577 to -779)	++	+
	E2 (-833 to -986)	+	++
	NK2 (-1233 to -1449)	++	++
	D2 (-4713 to -5055)	+	+
	NK3 (-6331 to -6531)	+	+
	D1 (-9102 to -9350)	+	++

(1) Primers are located at the very beginning and end of the denoted DNA segment (For primer sequence see Table 2.1). Numbers depict base pairs in relation to the transcription initiation site.

(2) ++ = high intensity; + = average intensity; +/- = low intensity

(3) ++ = one signal; + = 2 or 3 signals; +/- = 4 or more signals

Furthermore, we made sure to reduce the variations of the PCR analysis to a minimum. We used the Peqlab PCR Master-Mix Gold S, which focuses on specificity of the performed PCR reactions instead of quantity, to eliminate one source of possible variations in the PCR reactions. This master mix eliminated the necessity of pipetting small amounts of DNA polymerase, buffer, dNTPs and Mg^{2+} solution. Furthermore, we always made a master mix containing Peqlab PCR Master-Mix Gold S, dH_20 and DNA for the investigation of one protein with a certain gene, distributed equal amounts of this master mix to the individual eppis and then added a 1:10 dilution of primer pairs to each eppi. We also made sure to always use the same equipment to eliminate possible influences of different pipettes or eppis on the efficiency of the PCR reactions. The agarose gels used to separate the PCR samples were always of the same percentage and the gels were run with the same voltage and for a constant amount of time.

5.1.5. Evidence of the Specificity of ChIP Data

There are several aspects which support the claim that the conducted ChIP experiments yielded highly specific and significant data.

There were several primer pairs that never produced a significant signal in either one or even both cell lines. Most of those worked very efficiently with the input sample and produced only one specific signal at the correct position.

The **Brachyury gene** was generally occupied extremely frequently; therefore it was not surprising that almost all investigated DNA segments were occupied at one time or another. Only the P3 DNA segment was never occupied in CVPCs.
At the **Mesp1 gene** the P1 and 3'UTR DNA segments were never occupied by any of the investigated TFs on any investigated day, even though both primer pairs produced a strong and specific signal with the input sample. This is a strong indication that the obtained binding signals are highly specific, especially when considering that the P1 DNA segment is in relatively close proximity to the T1 DNA segment (733 bp apart), which was occupied very frequently and with high binding signal intensity by all proteins. Furthermore, the D1 DNA segment was never occupied in ESCs and only twice in CVPCs.
At the **Nkx2.5 gene** the D2 and P2 DNA segments were never occupied by any of the five TFs on any investigated day. The P2 primer pair produced a very strong signal at the correct position with the input sample, but never with the pull-down sample. The D2 primer pair was not very specific and produced several signals with the input sample, but there were, nevertheless, no signals of any size with the pull-down samples.
At the **Desmin gene** the D1 DNA segment was never occupied in ESCs and only once in CVPCs. The Prom DNA segment was never occupied in ESCs.

The primer pairs labeled with NK were originally designed to act as negative controls, where no TF binding was expected. We did, however, find strong and frequent binding signals at most NK DNA segments. Analysis of the other DNA strand and the search for partial TF binding sites identified several putative binding motifs for the investigated TFs in the amplified DNA segments or in close proximity, which can easily explain the high amount of obtained binding signals to these DNA segments.

The amount of protein bound to the investigated genes clearly did not depend on the pull-down sample. One example is that on day 7 of differentiation most genes were occupied by the five proteins with lower cumulative binding intensity than on day 5 of differentiation, but the Brachyury gene was occupied with a distinctly higher cumulative binding intensity of almost all TFs on day 7 than on day 5 of differentiation. Especially Desmin bound to the Brachyury gene much more strongly on day 7 than on day 5 of ESC differentiation (Figure 4.13). This suggests that binding signal intensity was not simply a

product of ChIP pull-down efficiency, but highly specific for each investigated TF and gene.

The specificity of the used antibodies is crucial for a successful ChIP experiment. We could show that the **anti-Nanog** and the **anti-Brachyury** antibodies produced highly specific signals in western blot experiments with no significant background signals. Furthermore, the anti-Brachyury antibody had already been used successfully in published ChIP experiments by another group (Evans et al., 2012). The **anti-Mesp1** antibody did not produce background signals in western blot either. The fact that we did not find any signal is most likely due to the fact that we investigated days where no high amounts of Mesp1 protein were present in the cells. Repeating the experiment with samples harvested on day 6 of ESC and day 10 of CVPC differentiation, one day after the obtained Brachyury signals, would very likely yield the expected results. The **anti-Nkx2.5** antibody was expected to produce several signals with a variable pattern (see antibody datasheet). We did find several signals at the correct position, which very likely specifically show the Nkx2.5 protein, because they did not correlate with proteins seen in the Ponceau staining and were not present in undifferentiated ESCs. The **anti-Desmin** antibody generated a signal at the correct position in western blot experiments, but also produced several background signals. This high background is, however, very likely due to the fact that this antibody might not be best suited for western blot experiments. If the antibody produced such a high background in ChIP experiments, more binding signals with a less defined pattern would be expected. An unspecific antibody would be expected to produce signals with very similar binding signal frequency and intensity for every investigated interaction. That was not the case, as Desmin showed very strong binding at the Brachyury gene on most investigated days, but was entirely absent at the Mesp1 gene on day 7 of ESC and on days 9 and 12 of CVPC differentiation.

5.2. The Significance of Differences in the DNA-Protein Interaction in ESCs and CVPCs

5.2.1. Differences Between the Five Genes

5.2.1.1. Comparison of the Occupation of Each Gene with the Five Proteins

When comparing the cumulative binding intensities of all five investigated TFs to the Brachyury, Mesp1, Nkx2.5 and Desmin gene, clear differences can be observed. In order

to illustrate these changes over the course of differentiation, we added all binding signals of all investigated proteins to a certain gene and summarized them in one diagram. The Nanog gene was not included as the results are not comparable to those of the other genes, because only the binding of Nanog and Brachyury has been investigated so far.

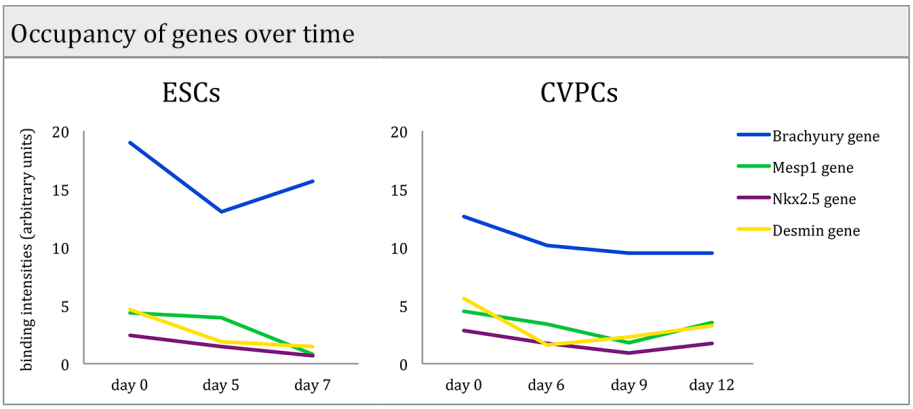


Figure 5.1. Cumulative binding intensities of all five proteins to the Brachyury, Mesp1, Nkx2.5 and Desmin gene over the course of differentiation in both ESCs and CVPCs.

The amount of protein bound to the **Mesp1, Nkx2.5 and Desmin gene** was similar in ESCs and CVPCs, although slightly different binding peaks could be observed. The binding of TFs to the **Brachyury gene,** on the other hand, was markedly stronger than the binding to all other genes (Figure 5.1). In CVPCs, the Brachyury gene was bound by proteins at least twice as strongly as the other genes throughout differentiation and in ESCs at least 4 times as strongly. This suggests that the Brachyury gene might be regulated in a different manner than the other investigated genes, possibly by so-called super enhancers, which usually span several thousand base pairs (Whyte et al., 2013). The Brachyury gene would meet the criteria for the existence of such a super enhancer as it is a key identity gene for the cardiac lineage, and in silico analysis of the DNA sequence identified an unusually high frequency of TF binding sites for not only the investigated proteins, but also for Oct4 and Sox2 (Whyte et al., 2013).

Furthermore, the cumulative binding intensities of the five proteins at the Mesp1, Nkx2.5 and Desmin gene decreased in ESCs as differentiation progressed, but the binding to the Brachyury gene rose sharply on day 7 of differentiation after a drop on day 5. Interestingly, the Mesp1, Nkx2.5 and Desmin gene were occupied with similar amounts of proteins in both ESCs and CVPCs, while the Brachyury gene was bound more strongly in ESCs.

In summary it can be said that there are some differences between the occupancy of the four genes in ESCs and CVPCs but also several similarities. This suggests that most of the investigated genes are regulated in a similar fashion with small differences that might be responsible for the fine-tuning of gene expression.

5.2.1.2. Comparison of the Occupation of Individual DNA Segments of the Five Genes with All Five Proteins

There are several aspects of the ChIP experiments that suggest a different regulation of the respective genes in ESCs and CVPCs. Although it is hard to directly compare the change of TF binding over the course of differentiation due to the fact that different days were investigated and because CVPCs are believed to be several days ahead in development, it is possible to draw some conclusions based on general occupation of DNA segments and the binding behavior in regard to the peaks of protein expression.

The differences between ESCs and CVPCs were best visible in the western blot experiments. In ESCs, Brachyury could be detected on day 5 of differentiation and in CVPCs on day 9 of differentiation, but day 9 of CVPC differentiation and day 5 of CVPC differentiation would also have to be investigated to completely eliminate the possibility that Brachyury is expressed on those days too. An explicit difference can be seen when looking at the Desmin protein: in ESCs, there was no signal at the correct position on days 0 or 5 of differentiation. In CVPCs, Desmin was already expressed in undifferentiated cells and the signal became stronger as differentiation progressed. The same was true for Nkx2.5: no protein could be detected at the expected position in undifferentiated ESCs, but in CVPCs there was a signal in undifferentiated cells.

In order to illustrate the different occupation of certain DNA segments in ESCs and CVPCs, we selected DNA segments that either had an already described regulatory function or were occupied frequently by the investigated proteins. To compare the cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to selected DNA segments of the genes over time in both ESCs and CVPCs, the intensity of the protein interactions of all investigated TFs with a certain DNA segment of the gene was summarized (Figure 5.2 to 5.6). DNA segments which showed a similar trend and are located in close proximity to each other were combined. Only relative changes can be depicted because the affinities of the Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin antibodies used for ChIP are not known.

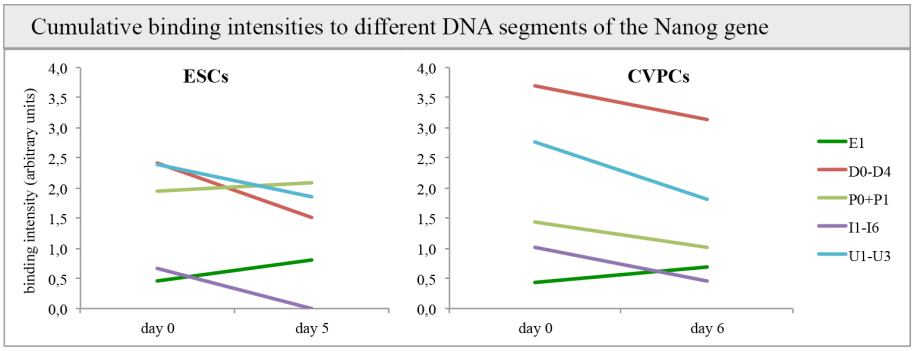


Figure 5.2. Cumulative binding intensities of Nanog and Brachyury to the enhancer (E1), distal promoter (D0, D1, D2, D3 and D4), proximal promoter (P0 and P1), first intron (I1, I2, I3, I4, I5 and I6) and the 3'UTR (U1, U2 and U3) DNA segments of the Nanog gene over time. Only the data generated on days 0 and 5 of ESC and days 0 and 6 of CVPC differentiation were considered here, because the binding of Brachyury to the Nanog gene was not investigated beyond those days.

At the **Nanog gene** most investigated DNA segments had declining cumulative binding intensities of Nanog and Brachyury over the course of differentiation in both ESCs and CVPCs (Figure 5.2). In ESCs, only the binding to the enhancer (E1) and the proximal promoter (P0 and P1) (Kuroda et al., 2005; Rodda et al., 2005; Suzuki et al., 2006) increased on day 5 of differentiation when compared to day 0. In CVPCs, only the binding to the enhancer (E1) increased on day 6 when compared to day 0 of differentiation. The amount of proteins bound to proximal promoter (P0+P1) declined over the course of CVPC differentiation. Furthermore, the distal promoter (D0-D4), the first intron (I1-I6) and the 3'UTR (U1-U3) were occupied with significantly more protein in CVPCs than in ESCs, whereas the proximal promoter (P0+P1) was bound more strongly in ESCs.

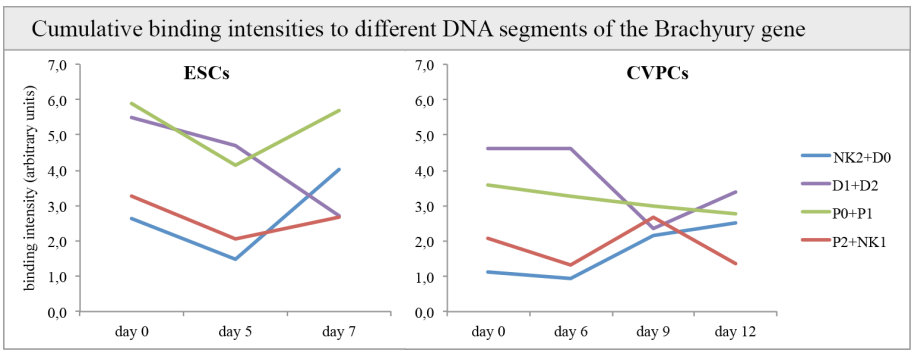


Figure 5.3. Cumulative binding intensities of all five proteins to the distal (NK2+D0 and D1+D2) and more proximal (P0+P1 and P2+NK1) DNA segments of the Brachyury gene over time.

At the **Brachyury gene** (Figure 5.3), both the distal DNA segments (NK2+D0 and D1+D2) and the more proximal DNA segments (P0+P1 and P2+NK1) were occupied with more TFs in undifferentiated ESCs than in undifferentiated CVPCs. The more proximal DNA segments P0 and P1 were occupied with a declining cumulative binding intensity of TFs in CVPCs, whereas in ESCs the binding intensity increased on day 7 when compared to day 5 of differentiation. At the time of maximal Brachyury protein expression on day 5 of ESC differentiation (see 4.1.2), all DNA segments except D1 and D2 were occupied with the lowest cumulative binding intensity of TFs when compared to the other investigated days. At the time of maximal Brachyury protein expression on day 9 of CVPC differentiation, the D1 and D2 DNA segments were occupied with the lowest cumulative binding intensity of TFs, whereas the other distal NK2 and D0 as well as the proximal P2 and NK1 DNA segments were bound quite strongly by TFs on day 9. This suggests a different regulation of the Brachyury gene in ESCs and CVPCs.

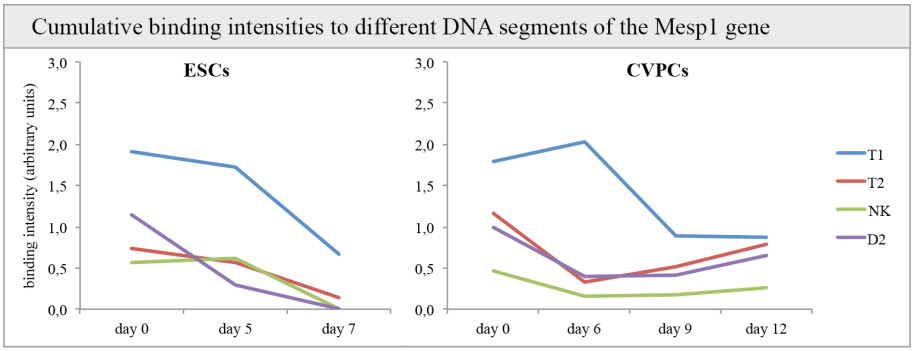


Figure 5.4. Cumulative binding intensities of all five proteins to the distal (NK and D2) and proximal (T1 and T2) DNA segments of the Mesp1 gene over time.

At the **Mesp1 gene** (Figure 5.4), the most obvious difference between ESCs and CVPCs was the fact that the binding of TFs to almost all DNA segments (D2, NK, T2) declined drastically on day 7 of ESC differentiation, whereas in CVPCs the proximal T1 and T2 DNA segments (van den Ameele et al., 2012) as well as the more distal NK and D2 DNA segments were occupied by TFs on every investigated day. In ESCs, only the very proximal T1 DNA segment was still occupied with considerable cumulative binding intensity of TFs on day 7 of differentiation.

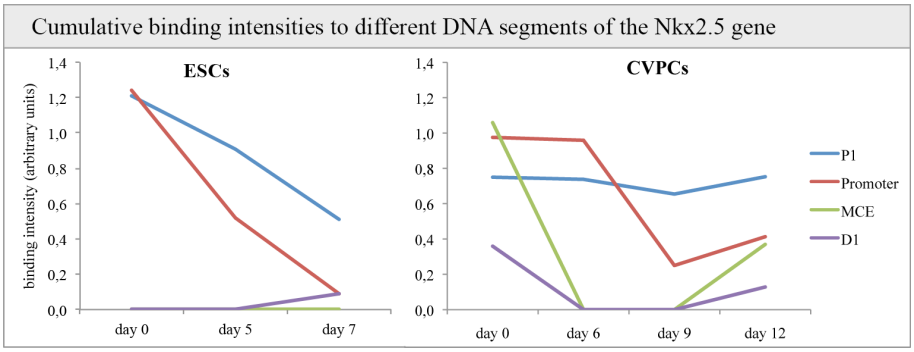


Figure 5.5. Cumulative binding intensities of all five proteins to the P1, promoter (Prom1 and Prom2), D1 and the MCE (E1, E2 and MCE) DNA segments of the Nkx2.5 gene over time. Of the binding signals generated with the overlapping E1, E2 and MCE as well as the overlapping Prom1 and Prom2 primer pairs, respectively, only the strongest signal was considered here.

At the **Nkx2.5 gene** (Figure 5.5), the change of cumulative binding signal intensities to different DNA segments over time most significantly differed between ESCs and CVPCs. The P1 DNA segment had higher cumulative binding intensity of TFs in undifferentiated ESCs than in undifferentiated CVPCs. As differentiation progressed, the cumulative binding intensity of TFs to the P1 DNA segment declined steadily in ESCs, but remained rather constant in CVPCs. The amount of protein bound to distal the promoter (Searcy et al., 1998) also decreased steadily in ESCs, whereas it remained constant until day 6 of CVPC differentiation, before decreasing on day 9.

The MCE (Lien et al., 1999) of the Nkx2.5 gene was bound very strongly by TFs in undifferentiated CVPCs and again on day 12 of CVPC differentiation, but never in ESCs. This was surprising, as previous work from our group demonstrated that Desmin occupied the MCE of the Nkx2.5 gene with high binding intensity between days 6 and 8 of ESC differentiation (unpublished data). Interestingly, the D1 and the MCE DNA segments, which are located only 113 bp apart, showed exactly the same trend in CVPCs: the cumulative binding intensity of TFs was highest on day 0, disappeared on days 6 and 9, and reappeared again on day 12 of differentiation. Furthermore, the D1 DNA segment was always occupied by a TF also bound to the MCE (Nanog on day 0 and Mesp1 on day 12 of CVPC differentiation), but with lower binding signal intensity. This observation suggests that in CVPCs the binding signals to the D1 DNA segment were the product of co-amplification of an actual occupation of the MCE. Therefore, the persistent binding of Desmin to the D1 DNA segment (# of PCRs = 4; # of signals = 4) on day 7 of differentiation probably actually reflects binding to the MCE.

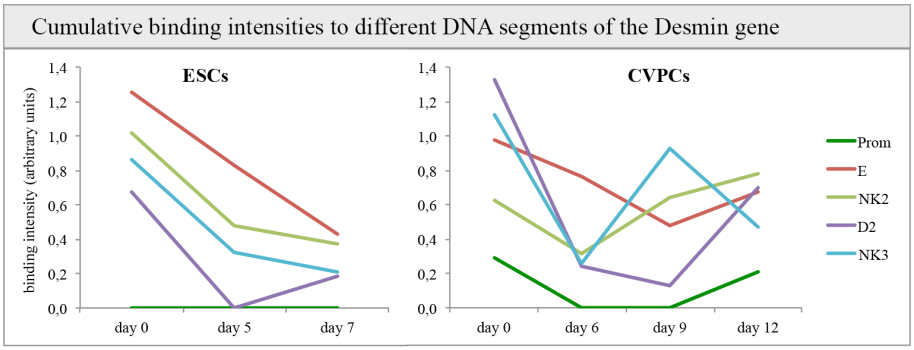


Figure 5.6. Cumulative binding intensities of all five proteins to the promoter (Prom), the enhancer (E1 and E2), the NK2, D2 and NK3 DNA segments of the Desmin gene over time. Of the binding signals generated with the E1 and E2 primer pairs, only the stronger signal was considered here.

At the **Desmin gene** (Figure 5.6), no proteins occupied the promoter (Li and Capetanakil, 1994) in ESCs. All other DNA segments showed declining cumulative binding intensities of TFs in ESCs on day 5 of differentiation when compared to undifferentiated cells, and only the binding intensity of TFs to the D2 DNA segment increased again on day 7. This was not surprising, as the Desmin gene should not be expressed strongly until later in ESC differentiation. In CVPCs, where the Desmin gene should be expressed during every investigated time point, the situation was much more complex. The promoter was occupied with TFs on days 0 and 12 of CVPC differentiation. Interestingly, occupation of the promoter and the enhancer (Li and Capetanakil, 1994) correlated to some extent: cumulative binding intensity of TFs to both DNA segments declined from day 0 to day 6 and increased from day 9 to day 12 of CVPC differentiation.

5.2.1.3. Résumé: Different Roles of Genes

Of the five genes, Brachyury was occupied with by far the highest amounts of proteins. Notably, there were more TFs bound to the Brachyury gene in ESCs than in CVPCs. There were also several differences between ESCs and CVPCs when looking at individual DNA segments of the five genes. Most noteworthy were the following: at the **Nanog gene**, the proximal promoter was bound much more strongly in ESCs than in CVPCs, and while the amount of protein bound to this DNA segment increased over the course of ESC differentiation, it declined in CVPCs. At the **Brachyury gene**, the more proximal DNA segments P0 and P1 had a declining cumulative binding intensity of TFs in CVPCs over time, whereas in ESCs the binding intensity increased on day 7 when compared to day 5 of differentiation. At the **Mesp1 gene**, the binding of TFs to almost all

DNA segments disappeared on day 7 of ESC differentiation, whereas in CVPCs the proximal T1 and T2 DNA segments (van den Ameele et al., 2012) as well as the more distal NK and D2 DNA segments were occupied by TFs on every investigated day. At the **Nkx2.5 gene**, the MCE (Lien et al., 1999) was bound quite strongly by TFs in undifferentiated CVPCs, but not in undifferentiated ESCs. Furthermore, the amount of proteins bound to the P1 DNA segment decreased over the course of ESC differentiation, but remained rather constant in CVPCs. At the **Desmin gene**, no proteins were bound to the promoter (Li and Capetanakil, 1994) in ESCs, but we did find TFs binding to this DNA segment in CVPCs.

5.2.2. Differences between the Five Proteins

5.2.2.1. Quantification of the Presence of the Five Proteins at the Four Genes

Next, we looked at the cumulative binding intensity of each individual protein: we added the amount of a given protein bound to the Brachyury, Mesp1, Nkx2.5 and Desmin gene for each day. Proteins bound to the Nanog gene were not included as only the binding of Nanog and partially that of Brachyury have been investigated so far.

The Brachyury gene was occupied with much more TFs than the other three genes (Figure 5.1). But if we look at the cumulative presence of the five proteins at the four genes it becomes evident that the overall amount of proteins bound to the four genes was similar (Figure 5.7).

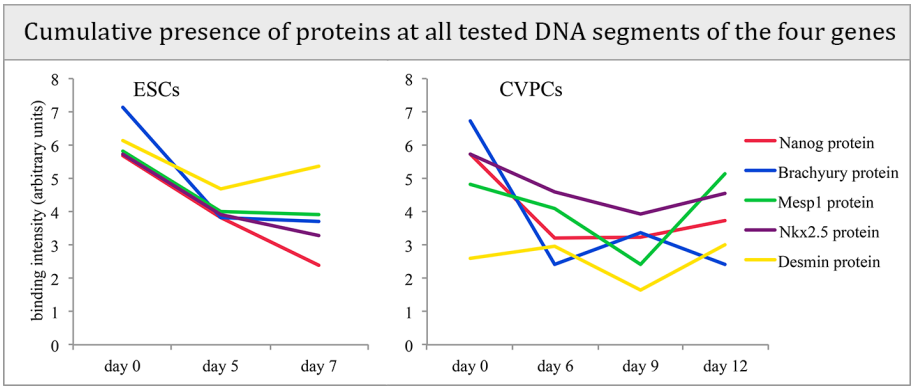


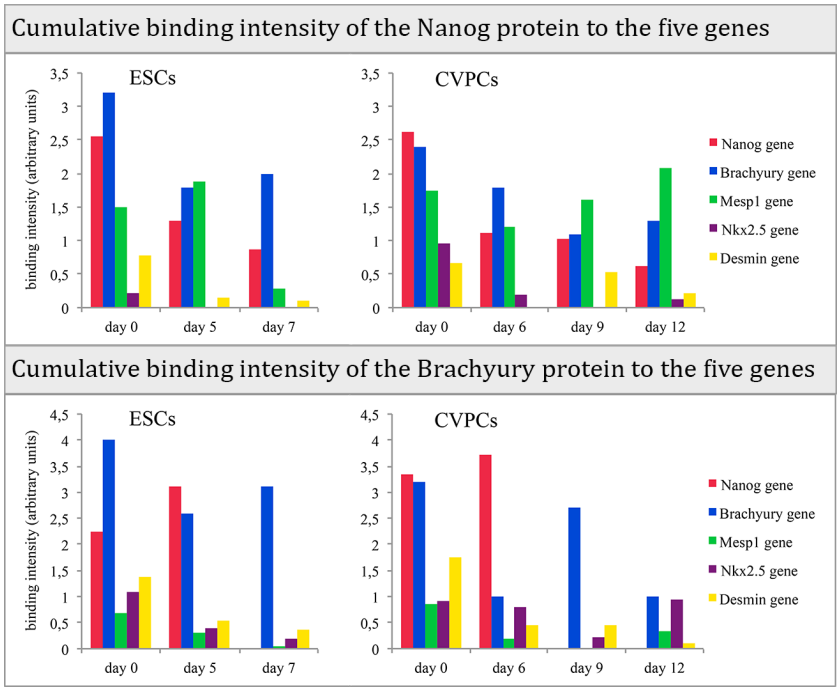
Figure 5.7. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the Brachyury, Mesp1, Nkx2.5 and Desmin gene in both ESCs and CVPCs over time.

In ESCs, the trend of cumulative binding intensities of TFs was very similar for the five investigated proteins. Without exception the binding intensity of each protein was highest

in undifferentiated cells. In CVPCs no clear trend was visible: Nanog, Brachyury and Nkx2.5 occupied the investigated genes with the highest cumulative binding intensity in undifferentiated cells. The binding of Desmin to the four genes was strongest on days 6 and 12 of differentiation. Mesp1 occupied the investigated genes with the highest cumulative binding intensity on day 12. Interestingly, Brachyury binding showed a peak on day 9 of differentiation, which coincides with the peak of protein expression found in western blot experiments. Surprisingly, much higher amounts of Desmin protein were bound to the four genes in ESCs than in CVPCs.

5.2.2.2. Quantification of the Presence of Each Individual Protein at the Five Genes

Looking at the amount of a given protein at each of the five genes revealed more pronounced differences between the investigated proteins. In order to illustrate these differences, we determined the cumulative binding intensities of a given protein to each of the five genes in both ESCs and CVPCs (Figure 5.8).



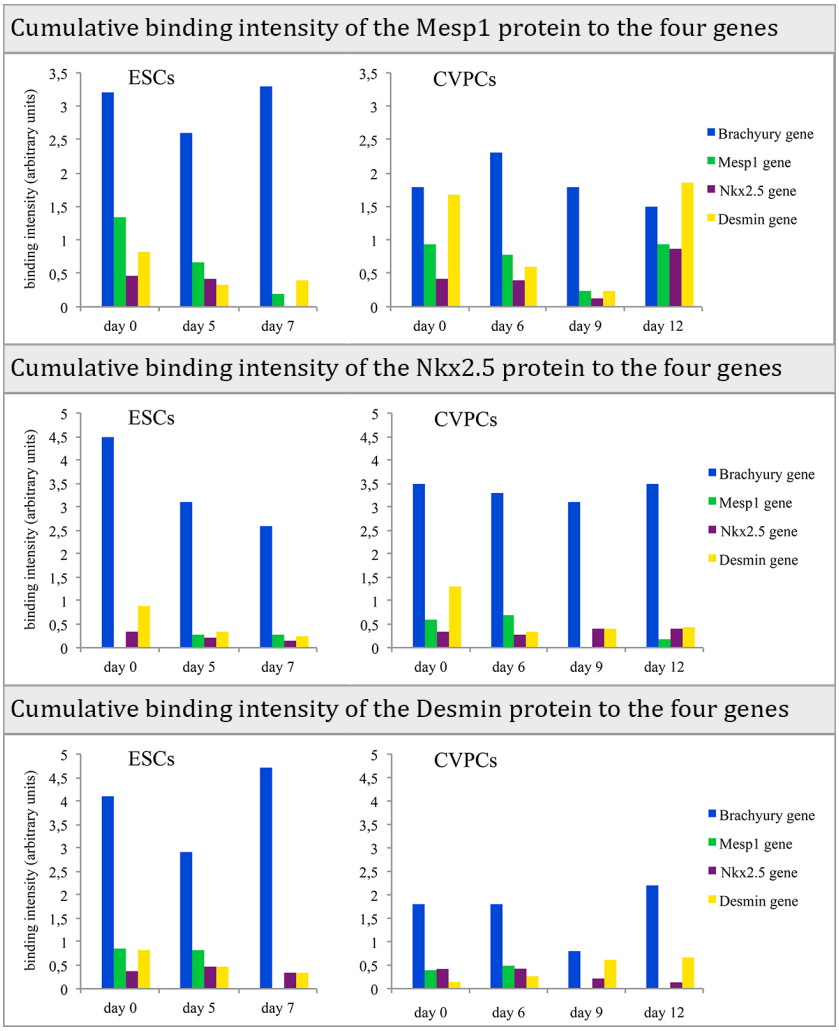


Figure 5.8. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to the investigated genes in ESCs and CVPCs. For Nanog, the binding to all five genes on every investigated day was analyzed. The binding of Brachyury to the Nanog gene was only investigated on days 0 and 5 of ESC and days 0 and 6 of CVPC differentiation. The binding of Mesp1, Nkx2.5 and Desmin to the Nanog gene was not analyzed.

The **Nanog protein** (Figure 5.8) interacted with all five genes in both ESCs and CVPCs, but great differences between the cumulative binding intensities could be observed. In both ESCs and CVPCs, Nanog interacted quite strongly with the Nanog and the Brachyury gene, but only weakly with the Desmin and the Nkx2.5 gene. There were several days when Nanog did not occupy the Nkx2.5 gene. In ESCs, Nanog occupied the Mesp1 gene with high cumulative binding intensity on days 0 and 5 of differentiation, but not on day 7. In CVPCs Nanog interacted strongly with the Mesp1 gene throughout differentiation.

The **Brachyury protein** (Figure 5.8) occupied both the Nanog and the Brachyury gene with high cumulative binding signal intensity, whereas binding to the other three genes was comparably weak. Unlike Nanog, Brachyury was bound only weakly to the Mesp1 gene in both cell lines. This was surprising, as Brachyury has been shown to activate Mesp1 gene expression (David et al., 2011).

The **Mesp1 protein** (Figure 5.8) showed a distinct difference in cumulative binding intensity in ESCs and CVPCs: Mesp1 bound to the Brachyury gene quite strongly in ESCs, whereas the cumulative binding intensity was significantly lower in CVPCs. Mesp1 bound to the Desmin gene only weakly in ESCs but quite strongly in CVPCs when compared to the other genes. On day 12 of CVPC differentiation, the highest amount of Mesp1 protein could be found at the Desmin gene.

The **Nkx2.5 protein** (Figure 5.8) occupied the Brachyury gene with very high cumulative binding signal intensity throughout differentiation in both ESCs and CVPCs, while the other genes were bound comparably weakly.

The **Desmin protein** (Figure 5.8) bound to the Brachyury gene very strongly in ESCs. In CVPCs, Desmin still occupied the Brachyury gene with higher cumulative binding intensity than the other genes, but the amount of Desmin protein at the Brachyury gene was markedly less than in ESCs.

In summary it can be stated that there were apparent differences not only between the investigated proteins and the occupied genes, but also between ESCs and CVPCs. Notably, all five proteins occupied the Brachyury gene much more strongly than any other gene.

5.2.2.3. Presence of the Five Proteins at Individual DNA Segments of the Five Genes

When analyzing data generated by ChIP experiments, several aspects should be considered with respect to the significance of an observed signal.

The presence of one or more binding motifs specific for a certain TF in the investigated DNA segment is always a strong indication of the specificity and significance of an obtained binding signal by this TF. There are, however, also several theories that might explain binding of a TF despite the lack of a binding motif: (1.) TFs could also bind other proteins and protein complexes of the transcriptional machinery without directly interacting with DNA ("Piggyback" recruitment). (2.) A TF could bind to a consensus sequence located at a distant DNA segment of the chromosome and interact with another protein, bound to a different binding site, through looping. In this case, ChIP experiments

could generate a signal for a TF in a DNA segment, which is actually occupied by another TF. 3. A TF might also bind to a sequence with low affinity and be anchored by other proteins or a co-activator. Furthermore, most or even all TFs can recognize multiple different binding motifs, many of which have probably not been discovered yet (Farnham, 2009).

Another aspect that should be considered is that the mere presence of a binding motif is evidently not enough for a TF to bind to a certain DNA segment. For most tested DNA segments we could identify at least one putative binding motif for each investigated TF in silico. Therefore, we would have obtained significantly more signals if the presence of a binding motif were enough for a TF to bind. There are two theories that might explain the lack of TF binding to its binding motif: (1.) Possibly the close packaging of nucleosomes in heterochromatin and DNA methylation prevents binding of TFs to their consensus motifs or (2.) Specific histone modifications might enhance the binding of a TF to a certain DNA segment (Farnham, 2009).

It was also proposed that in most cases DNA binding of a single TF might not have a functional effect, but that binding of several TFs, which act together, is necessary to activate transcription (Farnham, 2009).

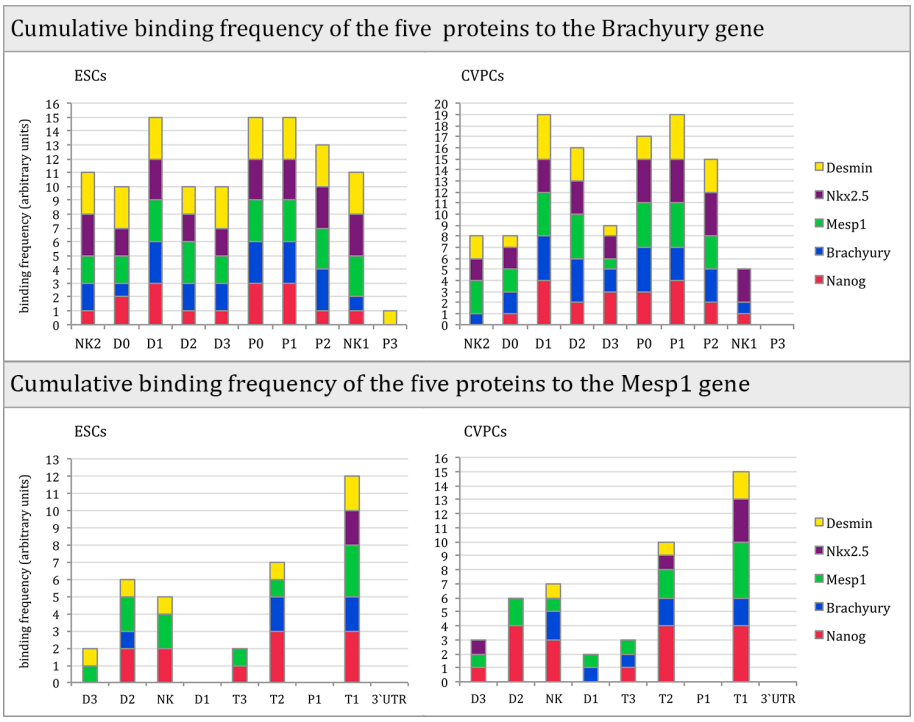
Moreover, it can be excluded that TFs sliding along the DNA strand in search of their respective binding sites (Berg et al., 1981; Marklund et al., 2013) produce strong signals when analyzed with ChIP. This process is very dynamic and it can be assumed that the momentary position of sliding TFs on DNA strands is simply stochastic and, therefore, leads to only very sporadic and weak signals.

Keeping all this in mind, frequent occupation of a DNA segment by several TFs and the presence of binding motifs for these TFs is a strong indication of the significance of an obtained binding signal.

Therefore, we analyzed which of the investigated DNA segments were bound most often throughout the course of differentiation and were therefore likely to be important for the regulation of gene expression. For that we counted the number of times each protein was bound to each DNA segment of the five genes over the course of ESC and CVPC differentiation. Only the statistically relevant binding signals were considered (for numbers of conducted PCRs for each individual protein-DNA interaction see appendix). Additionally, the intensities of the signals were not considered because varying DNA fragment length brings in stochasticity. Each binding signal of a TF was counted as a

score of one. For ESCs, each protein could have a score of zero to three for each DNA segment, with zero meaning no binding on any investigated day and three meaning binding on all three investigated days. For CVPCs, a score between zero and four was possible for a given TF for each DNA segment, with zero meaning no binding on any investigated day and four meaning binding on all four investigated days. In ESCs a total score of 15, and in CVPCs a total score of 20 signified the binding of all 5 TFs on every investigated day. The Nanog gene was excluded from this analysis because only the binding of two TFs was investigated.

As discussed above (5.2.1.1), the **Brachyury gene** seems to be regulated differently from the other investigated genes, possibly by super enhancers (Whyte et al., 2013). In ESCs, three DNA segments (D1, P0 and P1) of the Brachyury gene were bound on every single day by all five TFs. Several other DNA segments were bound by TFs almost as frequently both in ESCs and CVPCs. Notably, the NK2, D0 and D3 DNA segments were occupied more frequently with TFs in ESCs than in CVPCs. This extremely frequent occupation of the Brachyury gene with TFs seems to be specific, because none of the other investigated genes were occupied this frequently at any point in time.



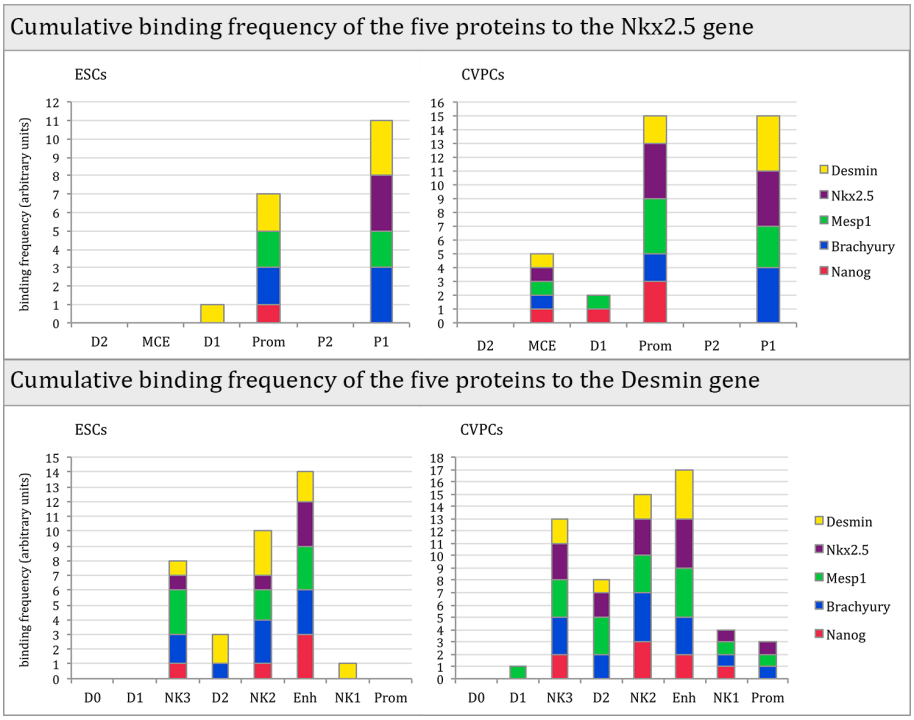


Figure 5.9. Cumulative binding frequency of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to investigated DNA segments of the respective genes in ESCs and CVPCs. One binding signal of a certain TF = one arbitrary unit; 100% = 15 for ESCs and 20 for CVPCs.

At the **Mesp1 gene** the T1 and T2 DNA segments (van den Ameele et al., 2012) were bound most frequently by TFs, followed by the D2 and NK DNA segments. The described T3 DNA segment (van den Ameele et al., 2012) was bound comparably infrequently. Interestingly, in silico analysis of the DNA sequence identified TF binding motifs for Nanog, Brachyury, Mesp1 and Nkx2.5 in or around the D2 DNA segment, but no Nanog and Nkx2.5 binding motifs in close proximity to the NK DNA segment. Nkx2.5 bound neither the D2 nor the NK DNA segment, but Nanog did bind both DNA segments rather frequently. Nanog could have either interacted with another protein bound to the NK DNA segments, or it could have bound to a different DNA segment (possibly T1 or T2) and interacted with proteins bound to the NK DNA segment through looping. There could also be yet unknown binding motifs for Nanog in the NK DNA segment. Interestingly, Nkx2.5 bound only to the T1 DNA segment in ESCs, but occupied the D3, T2 and T1 DNA segments in CVPCs. Surprisingly, the canonical Brachyury binding motif identified in the 3'UTR of the Mesp1 gene by in silico analysis of the DNA sequence was never occupied with TFs.

At the **Nkx2.5 gene**, only two DNA segments were occupied frequently with TFs: the distal promoter (Searcy et al., 1998) and the P1 DNA segment. In silico analysis of the DNA sequence identified TF binding motifs for Brachyury, Mesp1 and Nkx2.5 in the proximity of the P1 DNA segment, but not for Nanog. In the ChIP experiments we did find very frequent binding of Brachyury, Mesp1, Nkx2.5 and Desmin to the P1 DNA segment. Nanog never bound there, which correlates with the lack of binding motifs. Notably, in ESCs the proximal P1 DNA segment was bound by TFs even more frequently than the described distal promoter and as frequently as the promoter in CVPCs.

At the **Desmin gene**, the described enhancer (Li and Capetanakil, 1994) was occupied most frequently with TFs, but other not yet described DNA segments were also occupied very frequently with several TFs: the NK3 and NK2 DNA segments were both bound by all five TFs. In silico analysis of the DNA sequence identified binding motifs for Nanog, Brachyury, Mesp1 and Nkx2.5 in or close to the NK3 DNA segment, but we found only binding motifs for Nanog, Brachyury and Nkx2.5 in or close to the NK2 DNA segment. As discussed above, Mesp1 could either have bound to an unknown binding motif or to other proteins bound to the NK2 DNA segment. The D2 DNA segment at the Desmin gene was also bound by several TFs, but only three times in ESCs (max. 15) and eight times in CVPCs (max. 20). The more frequent binding in CVPCs could hint at a cell-type specific DNA segment, but further experiments would be needed to make a clearer statement.

In summary it can be stated that there are several described regulatory DNA segments as well as not yet described DNA segments within the five genes that were occupied frequently by the investigated TFs. Even more reassuring is the fact that there were also several DNA segments where none of the five TFs bound on any of the investigated days. This indicates that the frequent binding signals found at other DNA segments were indeed highly specific.

5.2.2.4. Résumé: Different Roles of Proteins

The only stark difference between ESCs and CVPCs with respect to the cumulative binding intensities of the five proteins to all investigated genes was the fact that much higher amounts of Desmin protein occupied the five genes in ESCs than in CVPCs.

There were several differences between ESCs and CVPCs when looking at the amount of a given protein at each of the five genes. Most noteworthy were the following: the **Nanog protein** occupied the Mesp1 gene with high cumulative binding intensity on days 0 and 5

of ESC differentiation, but almost completely disappeared together with all other TFs on day 7. In CVPCs, Nanog bound strongly to the Mesp1 gene throughout differentiation. The **Brachyury protein** bound to its own gene strongly throughout ESC differentiation, but only weakly on days 6 and 12 of CVPC differentiation. Higher amounts of **Mesp1 protein** occupied the Brachyury gene in ESCs than in CVPCs, but the Desmin gene was bound more strongly by Mesp1 in CVPCs than in ESCs. The **Nkx2.5 protein** did not show pronounced differences between ESCs and CVPCs with respect to its cumulative binding intensity to the four genes. The **Desmin protein** bound to the Brachyury gene much more strongly in ESCs than in CVPCs.

Furthermore, there were differences between ESCs and CVPCs when looking at the frequency with which the five proteins occupied the individual DNA segments of the investigated genes. Most noteworthy were the following: at the **Brachyury gene**, the NK2, D0 and D3 DNA segments were occupied more frequently with TFs in ESCs than in CVPCs. At the **Mesp1 gene**, only the T1 DNA segment was occupied by Nkx2.5 in ESCs, but in CVPCs Nkx2.5 bound to the D3, T2 and T1 DNA segments. At the **Nkx2.5 gene**, the MCE (Lien et al., 1999) was bound by TFs only in CVPCs and the distal promoter (Searcy et al., 1998) was bound more frequently in CVPCs than in ESCs. At the **Desmin gene**, the promoter (Li and Capetanakil, 1994) was occupied with TFs only in CVPCs and the D2 DNA segment was occupied more frequently in CVPCs than in ESCs.

5.3. Potential New Regulatory Elements

Frequent occupation of a DNA segment by several TFs and the presence of binding motifs for these TFs is still the most reasonable approach when looking for an indication of possible new regulatory DNA segments by analysis of ChIP data.

In the conducted ChIP experiments we did find several not yet described DNA segments, which were occupied with high frequency by several or even all of the investigated TFs (Figure 5.9). With the present data it can only be speculated what their actual function might be. Nevertheless, identification of such frequently occupied DNA segments might be a good foundation for future experiments investigating the potential regulatory function of specific DNA segments.

At the **Brachyury gene** all investigated DNA segments, except the P3 DNA segment, were occupied very frequently by all five proteins. These data strongly suggest that the investigated regions of the Brachyury gene have an actual regulatory function. Further investigations would be necessary to make a clearer assessment. The distribution of

binding signals to the Brachyury gene might also indicate that the gene is regulated through looping. Investigation of the DNA segment between the D3 and the D2 DNA segments (approximately 2750 bp) might yield interesting data with respect to that question.

At the **Mesp1 gene**, the D2 and NK DNA segments were occupied frequently by several TFs. The fact that the D2 and NK DNA segments were occupied much more frequently than other investigated DNA segments in relatively close proximity (distance between D1 and NK DNA segments: 726 bp; distance between D2 and D3 DNA segments: 765 bp) indicates that they might have a regulatory function. It would be interesting to investigate the DNA segment between D2 and NK (approximately 1800 bp) to see if this whole DNA segment is occupied as frequently.

At **Nkx2.5 gene**, apart from the already describe MCE (Lien et al., 1999) and distal promoter (Searcy et al., 1998), only the P1 DNA segment was occupied frequently with TFs. In ESCs, the proximal P1 DNA segment was bound even more frequently than the described distal promoter, and in CVPCs it was bound as frequently as the promoter. This suggests that the chances of finding a regulatory element in this DNA segment, which is active during early cardiomyogenesis, are very high.

At the **Desmin gene**, the NK2 and NK3 DNA segments were occupied almost as frequently as the described enhancer (Li and Capetanakil, 1994). The very frequent binding to the NK2 and NK3 DNA segments makes them both good targets for possible new regulatory elements at the Desmin gene.

In summary it can be said that there are several not yet described DNA segments within the five genes that were bound very frequently by the investigated TFs, and it would not be surprising to find that they have a regulatory function in gene expression.

5.4. The Influence of LIF on the Interaction of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin with Their Respective Genes

In one ChIP experiment conducted with undifferentiated ESCs and CVPCs, we added 2000 units of LIF per ml medium additionally to the cultivation on feeder cells. We expected that the presence of most TFs would be reduced, as less cells should start to differentiate prematurely. Surprisingly, the effects of LIF supplementation manifested in an entirely different way: some TFs occupied the genes with even higher cumulative binding intensity than before and only the binding of Nanog clearly decreased. Notably, this experiment has been performed only once. Nonetheless, the fact that some proteins

showed a persistent pattern, while others did not, does suggest a direct influence of LIF on the binding behavior of some TFs.

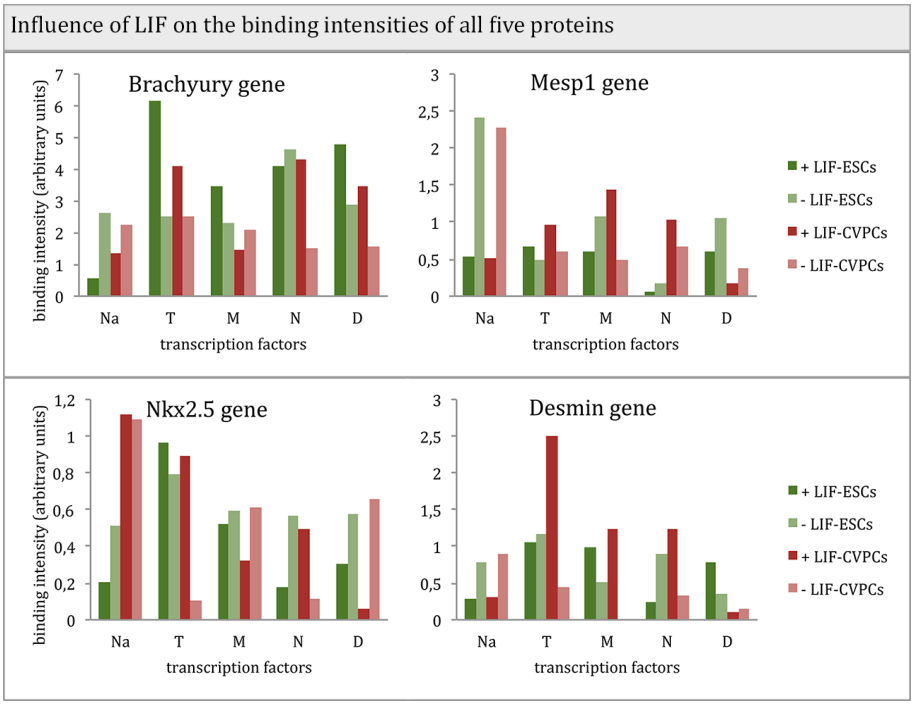


Figure 5.10. Cumulative binding intensities of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin to their respective genes in cells cultivated on feeder cells with supplementary LIF addition. Na, Nanog; T, Brachyury; M, Mesp1; N, Nkx2.5; D, Desmin.

Interestingly, the cumulative binding intensity of **Nanog** was significantly stronger in ESCs and CVPCs without additional LIF supplementation. Only the binding of Nanog to the Nkx2.5 gene in CVPCs showed slightly stronger intensities with LIF supplementation. This general decline in Nanog binding is somewhat surprising, as Nanog is known to be activated by LIF through the PI3K/Akt signaling pathway (Saunders et al., 2013). A possible explanation would be that LIF supports the Oct4/Sox2 dependent self-renewal pathway in the cells to such an extent that Nanog is not as important for the cells to remain in an undifferentiated state as under standard culture conditions on feeder cells. Under standard culture conditions, Nanog might have to bind and, therefore, inhibit genes expressed early during differentiation like that of Brachyury, Mesp1, and Nkx2.5 to maintain self-renewal. This might not be as important in cells supplemented with additional LIF, which are generally more stable. However, less Nanog

binding to the investigated genes does not necessarily implicate reduced Nanog protein levels.

In the presence of additional LIF, the cumulative binding intensity of **Brachyury** rose sharply to all but the Desmin gene in both ESCs and CVPCs. At the Mesp1, Nkx2.5 and Desmin gene, this effect was more pronounced in CVPCs than in ESCs. This strong influence of LIF could depend somehow on the already described binding of Stat3 to the Brachyury gene, as LIF is a known activator of Stat3 expression (Kidder et al., 2008).

Interestingly, **Nkx2.5** showed a different trend in ESCs and CVPCs: in ESCs, Nkx2.5 occupied all genes with a higher cumulative binding intensity without LIF supplementation, but in CVPCs the opposite was true.

For the **Mesp1** and the **Desmin** protein no clear trend could be observed, and the differences were generally not as pronounced as with the other three TFs. It could either be that LIF does not have a direct effect on the binding intensity of these two proteins, or it could be that the effect of LIF is simply more complex with different ramifications in ESCs and CVPCs, and at the different genes.

It can be concluded that the presence of Brachyury, Mesp1, Nkx2.5 and Desmin in undifferentiated ESCs was not due to insufficient LIF concentrations in the culture medium. Furthermore, the proteins were not present in the cells in quantities large enough to detect in western blot experiments. This suggests that the great majority of ESCs was actually in an undifferentiated state. It stands to reason that apart from small subpopulations of somewhat differentiated cells, which are always present under these culture conditions, undifferentiated ESCs and CVPCs express all of the investigated proteins at very low levels and can be detected with ChIP. If the obtained signals came only from the small subpopulations of differentiating cells, it could be assumed that the signals would get markedly stronger as differentiation progressed. This was not the case because the cumulative binding intensities to most genes declined over the course of differentiation (see 5.2.1.1.).

6. Conclusion and Outlook

The data generated by ChIP yielded very interesting insights into certain aspects of the very complex and dynamic processes of the transcriptional network involved in regulating early cardiomyogenesis. We could show that Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin physically interact with their respective genes during differentiation in both ESCs and CVPCs. These interactions clearly changed over the course of differentiation not only with respect to general occupancy of genes, but also when looking at the binding of specific TFs to individual DNA segments. We could show that the five investigated TFs bound to both, already described regulatory DNA segments in the 5'UTRs of the tested genes, and to possible new regulatory DNA segments. Several of these not yet described DNA segments were bound even more strongly than described regulatory regions, which strongly indicates that they have a function in the regulation of gene expression. There are still many open questions in regard to the way Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin gene expression is regulated during development. Further analysis of several DNA segments, which were occupied very frequently and with high binding intensity of TFs in our ChIP experiments, might yield interesting new insights into the way these genes are regulated during cardiomyogenesis.

We could also demonstrate by western blotting that there are clear differences between ESCs and CVPCs with respect to the exactly timed and tightly controlled expression of mesoderm-specific TFs. These differences between ESCs and CVPCs were also visible in the ChIP experiments, where some genes showed distinctly different occupation by the five tested TFs in the two cell lines.

Furthermore, an excess of LIF in the culture medium of undifferentiated stem cells had some surprising but very interesting effects on the binding of the five TFs to their respective genes. Further investigations on the exact role of LIF on the investigated TFs would likely yield interesting data, which might help elucidate some aspects of the complex dynamics involved in maintaining pluripotency. Better understanding of the mechanisms regulating pluripotency might also help to find new ways of inducing directed differentiation of cardiac stem cells into fully functioning cardiomyocytes. The data generated during the course of this thesis could present a valuable foundation for future projects investigating the transcriptional regulation of Nanog, Brachyury, Mesp1, Nkx2.5 and Desmin.

7. References

Van den Ameele, J., Tiberi, L., Bondue, A., Paulissen, C., Herpoel, A., Iacovino, M., Kyba, M., Blanpain, C., and Vanderhaeghen, P. (2012). Eomesodermin induces Mesp1 expression and cardiac differentiation from embryonic stem cells in the absence of Activin. EMBO Rep. *13*, 355–362.

Anversa, P., Leri, A., Rota, M., Hosoda, T., Bearzi, C., Urbanek, K., Kajstura, J., and Bolli, R. (2007). Concise review: stem cells, myocardial regeneration, and methodological artifacts. Stem Cells *25*, 589–601.

Asahina, K., Tsai, S.Y., Li, P., Ishii, M., Maxson, R.E., Sucov, H.M., and Tsukamoto, H. (2009). Mesenchymal origin of hepatic stellate cells, submesothelial cells, and perivascular mesenchymal cells during mouse liver development. Hepatology *49*, 998–1011.

Beddington, R.S.P., and Robertson, E.J. (1998). Anterior patterning in mouse. Trends Genet. *14*, 277–284.

Behrens, A.N., Iacovino, M., Lohr, J.L., Ren, Y., Zierold, C., Harvey, R.P., Kyba, M., Garry, D.J., and Martin, C.M. (2013). Nkx2-5 Mediates Differential Cardiac Differentiation Through Interaction with Hoxa10. Stem Cells Dev. *22*, 2211–2220.

Beltrami, A.P., Urbanek, K., Kajstura, J., Yan, S.M., Finato, N., Bussani, R., Nadal-Ginard, B., Silvestri, F., Leri, A., Beltrami, C.A., et al. (2001). Evidence that human cardiac myocytes divide after myocardial infarction. N. Engl. J. Med. *344*, 1750–1757.

Beltrami, A.P., Barlucchi, L., Torella, D., Baker, M., Limana, F., Chimenti, S., Kasahara, H., Rota, M., Musso, E., Urbanek, K., et al. (2003). Adult Cardiac Stem Cells Are Multipotent and Support Myocardial Regeneration. Cell *114*, 763–776.

Berg, O.G., Winter, R.B., and Von Hippel, P.H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry *20*, 6929–6948.

Van den Berg, D.L.C., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol. Cell. Biol. *28*, 5986–5995.

Ten Berge, D., Koole, W., Fuerer, C., Fish, M., Eroglu, E., and Nusse, R. (2008). Wnt signaling mediates self-organization and axis formation in embryoid bodies. Cell Stem Cell *3*, 508–518.

Bergmann, O., Bhardwaj, R.D., Bernard, S., Zdunek, S., Barnabé-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B.A., Druid, H., et al. (2009). Evidence for cardiomyocyte renewal in humans. Science *324*, 98–102.

Bondue, A., and Blanpain, C. (2010). Mesp1: a key regulator of cardiovascular lineage commitment. Circ. Res. *107*, 1414–1427.

Bondue, A., Lapouge, G., Paulissen, C., Semeraro, C., Iacovino, M., Kyba, M., and Blanpain, C. (2008). Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. Cell Stem Cell *3*, 69–84.

Booth, H.A.F., and Holland, P.W.H. (2004). Eleven daughters of NANOG. Genomics *84*, 229–238.

Boyer, L. a, Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Bratt-leal, M., Carpenedo, R.L., and Mcdevitt, T.C. (2009). Engineering the Embryoid Body Microenvironment to Direct Embryonic Stem Cell Differentiation. Biotechnol. Prog. *25*, 43–51.

Breckler, J., and Lazarides, E. (1982). Isolation of a new high molecular weight protein associated with desmin and vimentin filaments from avian embryonic skeletal muscle. J. Cell Biol. *92*, 795–806.

Bruneau, B.G., Nemer, G., Schmitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D.A., Gessler, M., Nemer, M., Seidman, C.E., et al. (2001). A murine model of Holt-Oram syndrome defines roles of the T-Box transcription factor Tbx5 in cardiogenesis and disease. Cell *106*, 709–721.

Buckley, S.M., Aranda-Orgilles, B., Strikoudis, A., Apostolou, E., Loizou, E., Moran-Crusio, K., Farnsworth, C.L., Koller, A.A., Dasgupta, R., Silva, J.C., et al. (2012). Regulation of Pluripotency and Cellular Reprogramming by the Ubiquitin-Proteasome System. Cell Stem Cell *11*, 783–798.

Can, A. (2008). A concise review on the classification and nomenclature of stem cells. Turkish J. Hematol. *25*, 57–59.

Capetanaki, Y., Milner, D.J., and Weitzer, G. (1997). Desmin in Muscle Formation and Maintenance : Knockouts and Consequences. Cell Struct. Funct. *22*, 103–116.

Carey, M.F., Peterson, C.L., and Smale, S.T. (2009). Chromatin immunoprecipitation (ChIP). Cold Spring Harb. Protoc. *4*, pdb.prot5279.

Cartaud, A., Jasmin, B.J., Changeux, J.P., and Cartaud, J. (1995). Direct involvement of a lamin-B-related (54 kDa) protein in the association of intermediate filaments with the postsynaptic membrane of the Torpedo marmorata electrocyte. J. Cell Sci. *108*, 153–160.

Chan, S.S.-K., Shi, X., Toyama, A., Arpke, R.W., Dandapat, A., Iacovino, M., Kang, J., Le, G., Hagen, H.R., Garry, D.J., et al. (2013). Mesp1 patterns mesoderm into cardiac, hematopoietic, or skeletal myogenic progenitors in a context-dependent manner. Cell Stem Cell *12*, 587–601.

Chen, Y., and Cao, X. (2009). NFAT directly regulates Nkx2-5 transcription during cardiac cell differentiation. Biol. Cell *101*, 335–349.

Chen, Y.C., and Schwartz, R.J. (1995). Identification of novel DNA binding targets and regulatory domains of a murine tinman homeodomain factor, nkx2.5. J. Biol. Chemsitry *270*, 15628–15633.

Chen, C.Y., Croissant, J., Majesky, M., Topouzis, S., McQuinn, T., Frankovsky, M.J., and Schwartz, R.J. (1996). Activation of the cardiac alpha-actin promoter depends upon serum response factor, Tinman homologue, Nkx-2.5, and intact serum response elements. Dev. Genet. *19*, 119–130.

Chugh, A.R., Beache, G.M., Loughran, J.H., Mewton, N., Elmore, J.B., Kajstura, J., Pappas, P., Tatooles, A., Stoddard, M.F., Lima, J. a C., et al. (2012). Administration of cardiac stem cells in patients with ischemic cardiomyopathy: the SCIPIO trial: surgical aspects and interim analysis of myocardial function and viability by magnetic resonance. Circulation *126*, 54–64.

Chung, D., Kuan, P.F., Li, B., Sanalkumar, R., Liang, K., Bresnick, E.H., Dewey, C., and Keleş, S. (2011). Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-Seq data. PLoS Comput. Biol. *7*, e1002111.

Costa, M.L., Escaleira, R., Cataldo, A., Oliveira, F., and Mermelstein, C.S. (2004). Desmin : molecular interactions and putative functions of the muscle intermediate filament protein. Brazilian J. Med. Biol. Res. *37*, 1819–1830.

Costello, I., Pimeisl, I.-M., Dräger, S., Bikoff, E.K., Robertson, E.J., and Arnold, S.J. (2011). The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. Nat. Cell Biol. *13*, 1084–1091.

Das, S., Jena, S., and Levasseur, D.N. (2011). Alternative Splicing Produces Nanog Protein Variants with Different Capacities for Self-renewal and Pluripotency in Embryonic Stem Cells. J. Biol. Chem. *286*, 42690–42703.

David, R., Jarsch, V.B., Schwarz, F., Nathan, P., Gegg, M., Lickert, H., and Franz, W.-M. (2011). Induction of MesP1 by Brachyury(T) generates the common multipotent cardiovascular stem cell. Cardiovasc. Res. *92*, 115–122.

Desbaillets, I., Ziegler, U., Groscurth, P., and Gassmann, M. (2000). Embryoid bodies : an in vitro model of mouse embryogenesis. Exp. Physiol. *85*, 645–651.

Dickson, V.V., Nocella, J., Yoon, H., Hammer, M., Melkus, G.D.E., and Chyun, D. (2013). Cardiovascular Disease Self-Care Interventions. Nurs. Res. Pract. *2013*, 1–16.

Dietrich, J.-E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development *134*, 4219–4231.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Daniel, J., Dash, C., Bazett-jones, D.P., Grice, S. Le, Ronald, D.G., Buetow, K.H., et al. (2009). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437–447.

Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A., Cheung, M.-S., Day, D.S., Gadel, S., Gorchakov, A.A., et al. (2011). An assessment of histone-modification antibody quality. Nat. Struct. Mol. Biol. *18*, 91–93.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Evans, A.L., Faial, T., Gilchrist, M.J., Down, T., Vallier, L., Pedersen, R. a, Wardle, F.C., and Smith, J.C. (2012). Genomic targets of Brachyury (T) in differentiating mouse embryonic stem cells. PLoS One *7*, e33346.

Farnham, P.J. (2009). Insights from genomic profiling of transcription factors. Nat. Rev. Genet. *10*, 605–616.

Fraenkel-Conrat, H., and Olcott, H.S. (1948). The reaction of formaldehyde with proteins; cross-linking between amino and primary amide or guanidyl groups. J. Am. Chem. Soc. *70*, 2673–2684.

Franco, D., and Kelly, R.G. (2011). Contemporary cardiogenesis: new insights into heart development. Cardiovasc. Res. *91*, 183–184.

Furey, T.S. (2013). ChIP-seq and Beyond: new and improved methodologies to detect and characterize protein-DNA interactions. Nat. Rev. Genet. *13*, 840–852.

Fürst, D.O., Osborn, M., and Weber, K. (1989). Myogenesis in the mouse embryo: differential onset of expression of myogenic proteins and the involvement of titin in myofibril assembly. J. Cell Biol. *109*, 517–527.

Galvez, B.G., Sampaolesi, M., Barbuti, A., Crespi, A., Covarello, D., Brunelli, S., Dellavalle, A., Crippa, S., Balconi, G., Cuccovillo, I., et al. (2008). Cardiac mesoangioblasts are committed, self-renewable progenitors, associated with small vessels of juvenile mouse ventricle. Cell Death Differ. *15*, 1417–1428.

Gambini, E., Pompilio, G., Biondi, A., Alamanni, F., Capogrossi, M.C., Agrifoglio, M., and Pesce, M. (2011). C-kit+ cardiac progenitors exhibit mesenchymal markers and preferential cardiovascular commitment. Cardiovasc. Res. *89*, 362–373.

Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E., et al. (2009). Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature *460*, 863–868.

Geisler, N., Schünemann, J., and Weber, K. (1992). Chemical cross-linking indicates a staggered and antiparallel protofilament of desmin intermediate filaments and characterizes one higher-level complex between protofilaments. Eur. J. Biochem. *206*, 841–852.

Georgatos, S.D., Weber, K., Geisler, N., and Blobel, G. (1987). Binding of two desmin derivatives to the plasma membrane and the nuclear envelope of avian erythrocytes: evidence for a conserved site-specificity in intermediate filament-membrane interactions. Proc. Natl. Acad. Sci. U. S. A. *84*, 6780–6784.

Getsios, S., Huen, A.C., and Green, K.J. (2004). Working out the strength and flexibility of desmosomes. Nat. Rev. Mol. Cell Biol. *5*, 271–281.

Go, A.S., Mozaffarian, D., Roger, V.L., Benjamin, E.J., Berry, J.D., Borden, W.B., Bravata, D.M., Dai, S., Ford, E.S., Fox, C.S., et al. (2013). Heart disease and stroke statistics--2013 update: a report from the American Heart Association. Circulation *127*, e6–e245.

Golob, J.L., Paige, S.L., Muskheli, V., Pabon, L., and Murry, C.E. (2011). Chromatin Remodeling During Mouse and Human Embryonic Stem Cell Differentiation. Dev. Dyn. *237*, 1389–1398.

Hansson, E.M., Lindsay, M.E., and Chien, K.R. (2009). Regeneration Next: Toward Heart Stem Cell Therapeutics. Cell Stem Cell *5*, 364–377.

Harel, I., Nathan, E., Tirosh-Finkel, L., Zigdon, H., Guimarães-Camboa, N., Evans, S.M., and Tzahor, E. (2009). Distinct Origins and Genetic Programs of Head Muscle Satellite Cells. Dev. Cell *16*, 822–832.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C., and Stam, M. (2007). Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. Plant Methods *3*, 11.

Heallen, T., Zhang, M., Wang, J., Bonilla-Claudio, M., Klysik, E., Johnson, R.L., and Martin, J.F. (2011). Hippo pathway inhibits Wnt signaling to restrain cardiomyocyte proliferation and heart size. Science *332*, 458–461.

Herrmann, B.G. (1991). Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. Development *113*, 913–917.

Herrmann, H., and Aebi, U. (2000). Intermediate filaments and their associates: Multi-talented structural elements specifying cytoarchitecture and cytodynamics. Curr. Opin. Cell Biol. *12*, 79–90.

Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. Nature *343*, 617–622.

Herrmann, H., Häner, M., Brettel, M., Ku, N.O., and Aebi, U. (1999). Characterization of distinct early assembly units of different intermediate filament proteins. J. Mol. Biol. *286*, 1403–1420.

Hoebaus, J., Heher, P., Gottschamel, T., Scheinast, M., Auner, H., Walder, D., Wiedner, M., Taubenschmid, J., Miksch, M., Sauer, T., et al. (2013). Embryonic Stem Cells Facilitate the Isolation of Persistent Clonal Cardiovascular Progenitor Cell Lines and Leukemia Inhibitor Factor Maintains Their Self-Renewal and Myocardial Differentiation Potential in vitro. Cells Tissues Organs Tissues *197*, 249–268.

Hosoda, T., Rota, M., Kajstura, J., Leri, A., and Anversa, P. (2011). Role of stem cells in cardiovascular biology. J. Thromb. Haemost. *9*, 151–161.

Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375–386.

Jauch, R., Ng, C.K.L., Saikatendu, K.S., Stevens, R.C., and Kolatkar, P.R. (2008). Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. J. Mol. Biol. *376*, 758–770.

Johnson, M.H., and Ziomek, C.A. (1981). The foundation of two distinct cell lineages within the mouse morula. Cell *24*, 71–80.

Kajstura, J., Urbanek, K., Rota, M., Bearzi, C., Hosoda, T., Bolli, R., Anversa, P., and Leri, A. (2008). Cardiac stem cells and myocardial disease. J. Mol. Cell. Cardiol. *45*, 505–513.

Kajstura, J., Urbanek, K., Perl, S., Hosoda, T., Zheng, H., Ogórek, B., Ferreira-Martins, J., Goichberg, P., Rondon-Clavo, C., Sanada, F., et al. (2010). Cardiomyogenesis in the adult human heart. Circ. Res. *107*, 305–315.

Kasahara, H., Usheva, a, Ueyama, T., Aoki, H., Horikoshi, N., and Izumo, S. (2001). Characterization of homo- and heterodimerization of cardiac Csx/Nkx2.5 homeoprotein. J. Biol. Chem. *276*, 4570–4580.

Kidder, B.L., Yang, J., and Palmer, S. (2008). Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells. PLoS One *3*, e3932.

Kim, H.-S., Park, S.Y., Choi, Y.R., Kang, J.G., Joo, H.J., Moon, W.K., and Cho, J.W. (2009). Excessive O-GlcNAcylation of proteins suppresses spontaneous cardiogenesis in ES cells. FEBS Lett. *583*, 2474–2478.

Kim, H.S., Woo, J.S., Joo, H.J., and Moon, W.K. (2012). Cardiac transcription factor Nkx2.5 is downregulated under excessive O-GlcNAcylation condition. PLoS One *7*, e38053.

Kiss, B., Röhlich, P., and Kellermayer, M.S.Z. (2011). Structure and elasticity of desmin protofibrils explored with scanning force microscopy. J. Mol. Recognit. *24*, 1095–1104.

Kitajima, S., Miyagawa-Tomita, S., Inoue, T., Kanno, J., and Saga, Y. (2006). Mesp1-nonexpressing cells contribute to the ventricular cardiac conduction system. Dev. Dyn. *235*, 395–402.

Konieczny, P., Fuchs, P., Reipert, S., Kunz, W.S., Zeöld, A., Fischer, I., Paulin, D., Schröder, R., and Wiche, G. (2008). Myofiber integrity depends on desmin network targeting to Z-disks and costameres via distinct plectin isoforms. J. Cell Biol. *181*, 667–681.

Kouloumenta, A., Mavroidis, M., and Capetanaki, Y. (2007). Proper perinuclear localization of the TRIM-like protein myospryn requires its binding partner desmin. J. Biol. Chem. *282*, 35211–35221.

Kouskoff, V., Lacaud, G., Schwantz, S., Fehling, H.J., and Keller, G. (2005). Sequential development of hematopoietic and cardiac mesoderm during embryonic stem cell differentiation. Proc. Natl. Acad. Sci. U. S. A. *102*, 13170–13175.

Kuisk, I.R., Li, H., Tran, D., and Capetanaki, Y. (1996). A Single MEF2 Site Governs Desmin Transcription in Both Heart and Skeletal Muscle during Mouse Embryogenesis. Dev. Biol. *174*, 1–13.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S., Suemori, H., Nakatsuji, N., and Tada, T. (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol. Cell. Biol. *25*, 2475–2485.

Langley, R.C., and Cohen, C.M. (1986). Association of spectrin with desmin intermediate filaments. J. Cell. Biochem. *30*, 101–109.

Lapouge, K., Fontao, L., Champliaud, M.-F., Jaunin, F., Frias, M. a, Favre, B., Paulin, D., Green, K.J., and Borradori, L. (2006). New insights into the molecular basis of desmoplakin- and desmin-related cardiomyopathies. J. Cell Sci. *119*, 4974–4985.

Laugwitz, K.-L., Moretti, A., Caron, L., Nakano, A., and Chien, K.R. (2008). Islet1 cardiovascular progenitors: a single source for heart lineages? Development *135*, 193–205.

Lauss, M., Stary, M., Tischler, J., Egger, G., Puz, S., Bader-allmer, A., Seiser, C., and Weitzer, G. (2005). Single inner cell masses yield embryonic stem cell lines differing in lifr expression and their developmental potential. Biochem. Biophys. Res. Commun. *331*, 1577–1586.

Li, H., and Capetanaki, Y. (1993). Regulation of the mouse desmin gene: transactivation by MyoD, myogenin, MRF4 and Myf5. Nucleic Acids Res. *21*, 335–343.

Li, H., and Capetanakil, Y. (1994). An E box in the desmin promoter cooperates with the E box and MEF-2 sites of a distal enhancer to direct muscle-specific transcription. EMBO J. *13*, 3580–3589.

Li, Q., and Johnston, S.A. (2001). Are All DNA Binding and Transcription Regulation by an Activator Physiologically Relevant ? Mol. Cell. Biol. *21*, 2467–2474.

Li, H., Choudhary, S.K., Milner, D.J., Munir, M.I., Kuisk, I.R., and Capetanaki, Y. (1994). Inhibition of desmin expression blocks myoblast fusion and interferes with the myogenic regulators MyoD and myogenin. J. Cell Biol. *124*, 827–841.

Li, X., MacArthur, S., Bourgon, R., Nix, D., Pollard, D. a, Iyer, V.N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C.L., et al. (2008). Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. PLoS Biol. *6*, e27.

Li, Y., Yu, W., Cooney, A.J., Schwartz, R.J., and Liu, Y. (2013). Brief report: Oct4 and canonical Wnt signaling regulate the cardiac lineage factor Mesp1 through a Tcf/Lef-Oct4 composite element. Stem Cells *31*, 1213–1217.

Lien, C.L., Wu, C., Mercer, B., Webb, R., Richardson, J.A., and Olson, E.N. (1999). Control of early cardiac-specific transcription of Nkx2-5 by a GATA-dependent enhancer. Development *126*, 75–84.

Limana, F., Zacheo, A., Mocini, D., Mangoni, A., Borsellino, G., Diamantini, A., De Mori, R., Battistini, L., Vigna, E., Santini, M., et al. (2007). Identification of myocardial and vascular precursor cells in human and mouse epicardium. Circ. Res. *101*, 1255–1265.

Lindsley, R.C., Gill, J.G., Murphy, T.L., Langer, E.M., Cai, M., Mashayekhi, M., Wang, W., Niwa, N., Nerbonne, J.M., Kyba, M., et al. (2008). Mesp1 coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. Cell Stem Cell *3*, 55–68.

Mabuchi, K., Li, B., Ip, W., and Tao, T. (1997). Association of calponin with desmin intermediate filaments. J. Biol. Chem. *272*, 22662–22666.

MacArthur, B.D., Sevilla, A., Lenz, M., Müller, F.-J., Schuldt, B.M., Schuppert, A. a, Ridden, S.J., Stumpf, P.S., Fidalgo, M., Ma'ayan, A., et al. (2012). Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. Nat. Cell Biol. *14*, 1139–1147.

Markham, K., Bai, Y., and Schmitt-Ulms, G. (2007). Co-immunoprecipitations revisited: an update on experimental concepts and their implementation for sensitive interactome investigations of endogenous proteins. Anal. Bioanal. Chem. *389*, 461–473.

Marklund, E.G., Mahmutovic, A., Berg, O.G., Hammar, P., Spoel, D. Van Der, Fange, D., and Elf, J. (2013). Transcription-factor binding and sliding on DNA studied using micro- and macroscopic models. Proc. Natl. Acad. Sci. U. S. A. *110*, 19796–19801.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Dev. Biol. *78*, 7634–7638.

McBratney-Owen, B., Iseki, S., Bamforth, S.D., Olsen, B.R., and Morriss-Kay, G.M. (2008). Development and tissue origins of the mammalian cranial base. Dev. Biol. *322*, 121–132.

McCullagh, K.J.A., Edwards, B., Poon, E., Lovering, R.M., Paulin, D., and Davies, K.E. (2007). Intermediate filament-like protein syncoilin in normal and myopathic striated muscle. Neuromuscul. Disord. NMD *17*, 970–979.

McMahon, A.P., and Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. Cell *62*, 1073–1085.

Meng, X.-L., Shen, J.-S., Kawagoe, S., Ohashi, T., Brady, R.O., and Eto, Y. (2010). Induced pluripotent stem cells derived from mouse models of lysosomal storage disorders. Proc. Natl. Acad. Sci. U. S. A. *107*, 7886–7891.

Messina, E., De Angelis, L., Frati, G., Morrone, S., Chimenti, S., Fiordaliso, F., Salio, M., Battaglia, M., Latronico, M.V.G., Coletta, M., et al. (2004). Isolation and expansion of adult cardiac stem cells from human and murine heart. Circ. Res. *95*, 911–921.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Miyanari, Y., and Torres-Padilla, M.-E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470–473.

Moretto-Zita, M., Jin, H., Shen, Z., Zhao, T., Briggs, S.P., and Xu, Y. (2010). Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. Proc. Natl. Acad. Sci. U. S. A. *107*, 13312–13317.

Müller, C.W., and Herrmann, B.G. (1997). Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. Nature *389*, 884–888.

Mullin, N.P., Yates, A., Rowe, A.J., Nijmeijer, B., Colby, D., Barlow, P.N., Walkinshaw, M.D., and Chambers, I. (2008). The pluripotency rheostat Nanog functions as a dimer. Biochem. J. *411*, 227–231.

Naiche, L. a, Harrelson, Z., Kelly, R.G., and Papaioannou, V.E. (2005). T-box genes in vertebrate development. Annu. Rev. Genet. *39*, 219–239.

Nguyen, Q.-G. V, Buskin, J.N., Himeda, C.L., Shield, M. a, and Hauschka, S.D. (2003). Differences in the function of three conserved E-boxes of the muscle creatine kinase gene in cultured myocytes and in transgenic mouse skeletal and cardiac muscle. J. Biol. Chem. *278*, 46494–46505.

Nibu, Y., José-Edwards, D.S., and Di Gregorio, A. (2013). From notochord formation to hereditary chordoma: the many roles of Brachyury. Biomed Res. Int. *2013*, 826435.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

Nishioka, N., Inoue, K. ichi, Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., et al. (2009). The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. Dev. Cell *16*, 398–410.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell *123*, 917–929.

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118–122.

Oginuma, M., Hirata, T., and Saga, Y. (2008). Identification of presomitic mesoderm (PSM)-specific Mesp1 enhancer and generation of a PSM-specific Mesp1/Mesp2-null mouse using BAC-based rescue technology. Mech. Dev. *125*, 432–440.

Orlando, V. (2000). Mapping chromosomal proteins in vivo by formaldehyde-immunoprecipitation. Trends Biochem. Sci. *25*, 99–104.

Pan, D. (2010). The hippo signaling pathway in development and cancer. Dev. Cell *19*, 491–505.

Pauken, C.M., and Capco, D.G. (2000). The expression and stage-specific localization of protein kinase C isotypes during mouse preimplantation development. Dev. Biol. *223*, 411–421.

Pereira, L., Yi, F., and Merrill, B.J. (2006). Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. Mol. Cell. Biol. *26*, 7479–7491.

Plusa, B., Frankenberg, S., Chalmers, A., Hadjantonakis, A.-K., Moore, C. a, Papalopulu, N., Papaioannou, V.E., Glover, D.M., and Zernicka-Goetz, M. (2005). Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. J. Cell Sci. *118*, 505–515.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408–1419.

Rodda, D.J., Chew, J.-L., Lim, L.-H., Loh, Y.-H., Wang, B., Ng, H.-H., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J. Biol. Chem. *280*, 24731–24737.

Saga, Y., Hata, N., Kobayashi, S., Magnuson, T., Seldin, M.F., and Taketo, M.M. (1996). MesP1: a novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. Development *122*, 2769–2778.

Saga, Y., Miyagawa-Tomita, S., Takagi, a, Kitajima, S., Miyazaki, J.I., and Inoue, T. (1999). MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. Development *126*, 3437–3447.

Saga, Y., Kitajima, S., and Miyagawa-Tomita, S. (2000). Mesp1 expression is the earliest sign of cardiovascular development. Trends Cardiovasc. Med. *10*, 345–352.

Saunders, A., Faiola, F., and Wang, J. (2013). Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. Stem Cells *31*, 1227–1236.

Schmidt, C., Wilson, V., Stott, D., and Beddington, R.S. (1997). T promoter activity in the absence of functional T protein during axis formation and elongation in the mouse. Dev. Biol. *189*, 161–173.

Schwartz, R.J., and Olson, E.N. (1999). Building the heart piece by piece: modularity of cis-elements regulating Nkx2-5 transcription. Development *126*, 4187–4192.

Schweitzer, S.C., Klymkowsky, M.W., Bellin, R.M., Robson, R.M., Capetanaki, Y., and Evans, R.M. (2001). Paranemin and the organization of desmin filament networks. J. Cell Sci. *114*, 1079–1089.

Searcy, R.D., Vincent, E.B., Liberatore, C.M., and Yutzey, K.E. (1998). A GATA-dependent nkx-2.5 regulatory element activates early cardiac gene expression in transgenic mice. Development *125*, 4461–4470.

Sedmera, D. (2011). Function and form in the developing cardiovascular system. Cardiovasc. Res. *91*, 252–259.

Sjöberg, G., Jiang, W.Q., Ringertz, N.R., Lendahl, U., and Sejersen, T. (1994). Colocalization of nestin and vimentin/desmin in skeletal muscle cells demonstrated by three-dimensional fluorescence digital imaging microscopy. Exp. Cell Res. *214*, 447–458.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688–690.

Spradling, A., Drummond-Barbosa, D., and Kai, T. (2001). Stem cells find their niche. Nature *414*, 98–104.

Steinhoff, G. (2013). Regenerative Medicine: From Protocol to Patient (2nd edition). Dordrecht. Springer.

Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F.H., Rodríguez-Esteban, C., and Izpisúa Belmonte, J.C. (2006). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *103*, 10294–10299.

Tada, M., and Smith, J.C. (2001). T-targets: clues to understanding the functions of T-box proteins. Dev. Growth Differ. *43*, 1–11.

Takaoka, K., and Hamada, H. (2012). Cell fate decisions and axis determination in the early mouse embryo. Development *139*, 3–14.

Tam, P.P., and Behringer, R.R. (1997). Mouse gastrulation: the formation of a mammalian body plan. Mech. Dev. *68*, 3–25.

Tanaka, T.S. (2009). Transcriptional heterogeneity in mouse embryonic stem cells. Reprod. Fertil. Dev. *21*, 67–75.

Tanaka, M., Wechsler, S.B., Lee, I.W., Yamasaki, N., Lawitts, J. a, and Izumo, S. (1999). Complex modular cis-acting elements regulate expression of the cardiac specifying homeobox gene Csx/Nkx2.5. Development *126*, 1439–1450.

Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature *455*, 1124–1128.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Titeux, M., Brocheriou, V., Xue, Z., Gao, J., Pellissier, J.F., Guicheney, P., Paulin, D., and Li, Z. (2001). Human synemin gene generates splice variants encoding two distinct intermediate filament proteins. Eur. J. Biochem. *268*, 6435–6449.

Tolstonog, G. V, Sabasch, M., and Traub, P. (2002). Cytoplasmic intermediate filaments are stably associated with nuclear matrices and potentially modulate their DNA-binding function. DNA Cell Biol. *21*, 213–239.

Torres, J., Prieto, J., Durupt, F.C., Broad, S., and Watt, F.M. (2012). Efficient differentiation of embryonic stem cells into mesodermal precursors by BMP, retinoic acid and Notch signalling. PLoS One *7*, e36405.

Vinckevicius, A., and Chakravarti, D. (2012). Chromatin immunoprecipitation: advancing analysis of nuclear hormone signaling. J. Mol. Endocrinol. *49*, R113–23.

Vinot, S., Le, T., Ohno, S., Pawson, T., Maro, B., and Louvet-Vallée, S. (2005). Asymmetric distribution of PAR proteins in the mouse embryo begins at the 8-cell stage during compaction. Dev. Biol. *282*, 307–319.

Wang, Q., Tolstonog, G. V, Shoeman, R., and Traub, P. (2001). Sites of nucleic acid binding in type I-IV intermediate filament subunit proteins. Biochemistry *40*, 10342–10349.

Warren, S. a, Terada, R., Briggs, L.E., Cole-Jeffrey, C.T., Chien, W.-M., Seki, T., Weinberg, E.O., Yang, T.P., Chin, M.T., Bungert, J., et al. (2011). Differential role of Nkx2-5 in activation of the atrial natriuretic factor gene in the developing versus failing heart. Mol. Cell. Biol. *31*, 4633–4645.

WHO (2013). WHO - Factsheet of Cardiovascular Diseases (CVDs). url as of 3.4. 2014: http://www.who.int/mediacentre/factsheets/fs317/en/.

Whyte, W. a, Orlando, D. a, Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R. a (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Wilkinson, D.G., Bhatt, S., and Herrmann, B.G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. Nature *343*, 657–659.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684–687.

Winter, D.L., Paulin, D., Mericskay, M., and Li, Z. (2013). Posttranslational modifications of desmin and their implication in biological processes and pathologies. Histochem. Cell Biol. *141*, 1–16.

Wu, S.M., Chien, K.R., and Mummery, C. (2008). Origins and Fates of Cardiovascular Progenitor Cells. Cell *132*, 537–543.

Xin, M., Olson, E.N., and Bassel-Duby, R. (2013). Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. Nat. Rev. Mol. Cell Biol. *14*, 529–541.

Xue, Z.G., Cheraud, Y., Brocheriou, V., Izmiryan, A., Titeux, M., Paulin, D., and Li, Z. (2004). The mouse synemin gene encodes three intermediate filament proteins generated by alternative exon usage and different open reading frames. Exp. Cell Res. *298*, 431–444.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281–292.

Yoshida, T., Vivatbutsiri, P., Morriss-Kay, G., Saga, Y., and Iseki, S. (2008). Cell lineage in mammalian craniofacial mesenchyme. Mech. Dev. *125*, 797–808.

Yuasa, S., Onizuka, T., Shimoji, K., Ohno, Y., Kageyama, T., Yoon, S.H., Egashira, T., Seki, T., Hashimoto, H., Nishiyama, T., et al. (2010). Zac1 is an essential transcription factor for cardiac morphogenesis. Circ. Res. *106*, 1083–1091.

Zhang, J., Wang, X., Li, M., Han, J., Chen, B., Wang, B., and Dai, J. (2006). NANOGP8 is a retrogene expressed in cancers. FEBS J. *273*, 1723–1730.

8. Appendix

Table 8.1. ChIP data

Protein	Gene	DNA segment	Cell line	Day of differentiation	# of PCRs	# of visible Signals	% of input (mean value)
Nanog	Nanog	E1	E[1]	0	9	1	/[3]
				5	2	1	/
				7	4	0	/
			C[2]	0	7	1	/
				6	5	2	/
				9	1	0	/
				12	1	0	/
		D0	E	0	11	5	0,399
				5	6	2	/
				7	4	0	/
			C	0	10	4	0,366
				6	5	1	/
				9	4	0	/
				12	4	1	/
		D1	E	0	8	3	/
				5	4	2	/
				7	4	1	/
			C	0	9	4	0,405
				6	5	2	/
				9	4	1	/
				12	4	0	/
		D2	E	0	8	7	0,822
				5	4	1	/
				7	4	0	/
			C	0	9	8	0,613
				6	4	3	0,455
				9	4	2	0,098
				12	4	1	/
		D3	E	0	10	7	0,637
				5	6	3	0,728
				7	4	1	/
			C	0	10	9	0,493
				6	6	3	0,303
				9	4	2	0,251
				12	4	1	/
		D4	E	0	6	2	/
				5	4	2	/
				7	4	4	0,861
			C	0	8	3	/
				6	4	1	/
				9	4	3	0,672
				12	3	2	0,612
		P0	E	0	11	9	0,276
				5	6	2	/
				7	4	1	/
			C	0	10	5	0,269
				6	6	1	/

		9	4	1	/
		12	4	0	/
P1	E	0	11	8	0,430
		5	6	3	0,571
		7	4	1	/
	C	0	10	10	0,475
		6	6	3	0,357
		9	4	1	/
		12	4	2	/
I1	E	0	4	0	/
		5	2	0	/
		7	-[4]	-	-
	C	0	5	1	/
		6	3	0	/
		9	-	-	-
		12	1	0	/
I2	E	0	6	2	/
		5	3	0	/
		7	-	-	-
	C	0	5	0	/
		6	5	1	/
		9	1	0	/
		12	1	0	/
I3	E	0	1	0	/
		5	1	0	/
		7	-	-	-
	C	0	3	0	/
		6	3	0	/
		9	-	-	-
		12	-	-	-
I4	E	0	7	0	/
		5	5	2	/
		7	4	1	/
	C	0	9	5	0,321
		6	4	1	/
		9	4	0	/
		12	4	0	/
I5	E	0	7	2	/
		5	4	1	/
		7	4	0	/
	C	0	8	3	/
		6	4	1	/
		9	4	0	/
		12	4	0	/
I6	E	0	9	4	0,181
		5	6	1	/
		7	4	0	/
	C	0	8	3	/
		6	6	0	/
		9	3	0	/
		12	4	0	/
U1	E	0	9	8	0,867
		5	4	4	0,873
		7	4	4	0,677
	C	0	9	6	0,889

			6	4	4	0,777
			9	4	4	0,509
			12	4	4	0,391
	U2	E	0	3	0	/
			5	5	2	/
			7	4	0	/
		C	0	6	2	/
			6	4	1	/
			9	1	0	/
			12	2	0	/
	U3	E	0	4	1	/
			5	5	2	/
			7	4	0	/
		C	0	6	4	0,358
			6	5	3	0,154
			9	4	0	/
			12	3	0	/
	U0	E	0	8	1	/
			5	4	1	/
			7	4	0	/
		C	0	8	2	/
			6	4	1	/
			9	4	0	/
			12	4	0	/
Brachyury	NK2	E	0	5	2	/
			5	4	1	/
			7	4	2	0,274
		C	0	5	0	/
			6	5	0	/
			9	4	0	/
			12	4	0	/
	D0	E	0	5	3	0,673
			5	4	1	/
			7	4	2	0,526
		C	0	3	1	/
			6	5	3	0,445
			9	4	0	/
			12	4	0	/
	D1	E	0	5	4	0,613
			5	4	2	0,809
			7	4	4	0,426
		C	0	5	4	0,539
			6	5	4	0,648
			9	4	3	0,152
			12	4	2	0,346
	D2	E	0	5	3	0,593
			5	4	0	/
			7	4	1	/
		C	0	4	3	0,210
			6	5	4	0,369
			9	4	1	/
			12	4	0	/
	D3	E	0	5	1	/
			5	4	1	/
			7	4	3	0,197

		C	0	5	3	0,589
			6	5	2	/
			9	4	2	0,179
			12	4	2	0,129
	P0	E	0	5	3	0,476
			5	4	2	0,450
			7	4	4	0,397
		C	0	5	4	0,375
			6	5	2	/
			9	4	4	0,211
			12	4	4	0,163
	P1	E	0	5	3	0,496
			5	4	2	0,323
			7	4	3	0,219
		C	0	4	2	0,221
			6	5	4	0,310
			9	4	4	0,217
			12	3	3	0,297
	P2	E	0	5	4	0,298
			5	4	1	/
			7	4	1	/
		C	0	5	3	0,239
			6	5	2	/
			9	4	4	0,297
			12	4	2	0,352
	NK1	E	0	5	2	/
			5	4	2	0,201
			7	4	1	/
		C	0	4	2	0,210
			6	5	1	/
			9	4	0	/
			12	4	0	/
	P3	E	0	5	0	/
			5	4	1	/
			7	4	0	/
		C	0	4	0	/
			6	5	1	/
			9	3	0	/
			12	4	0	/
Mesp1	D3	E	0	4	1	/
			5	4	1	/
			7	4	0	/
		C	0	4	0	/
			6	4	1	/
			9	4	2	0,163
			12	4	3	0,629
	D2	E	0	4	4	0,381
			5	4	2	0,204
			7	4	0	/
		C	0	4	2	0,819
			6	4	4	0,397
			9	4	4	0,257
			12	4		0,432
	NK	E	0	4	3	0,259
			5	4	2	0,556

			7	4	1	/
		C	0	4	3	0,227
			6	4	3	0,152
			9	4	3	0,173
			12	4	32	/
	D1	E	0	4	1	/
			5	4	0	/
			7	4	1	/
		C	0	4	0	/
			6	4	0	/
			9	4	1	/
			12	4	0	/
	T3	E	0	4	2	/
			5	4	2	0,183
			7	4	0	/
		C	0	4	1	/
			6	4	4	0,136
			9	4	0	/
			12	4	0	/
	T2	E	0	4	4	0,445
			5	4	2	0,464
			7	4	3	0,092
		C	0	4	4	0,411
			6	4	4	0,197
			9	4	4	0,512
			12	4	3	0,524
	P1	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
	T1	E	0	4	4	0,405
			5	4	4	0,465
			7	4	4	0,186
		C	0	4	3	0,279
			6	4	4	0,328
			9	4	4	0,662
			12	4	4	0,489
	3’UTR	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
Nkx2.5	D2	E	0	2	0	/
			5	2	0	/
			7	-	-	-
		C	0	5	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
	E1	E	0	4	0	/

		5	6	0	/
		7	4	0	/
	C	0	5	0	/
		6	4	0	/
		9	4	0	/
		12	4	0	/
E2	E	0	8	1	/
		5	6	0	/
		7	3	0	/
	C	0	5	3	0,219
		6	4	0	/
		9	4	0	/
		12	4	0	/
MCE	E	0	-	-	-
		5	-	-	-
		7	-	-	-
	C	0	-	-	-
		6	-	-	-
		9	-	-	-
		12	-	-	-
D1	E	0	8	3	/
		5	5	0	/
		7	4	0	/
	C	0	5	3	0,359
		6	4	0	/
		9	4	11	/
		12	4	0	/
Prom1	E	0	8	5	0,215
		5	6	2	/
		7	4	1	/
	C	0	5	3	0,240
		6	4	2	0,200
		9	4	1	/
		12	4	3	0,128
Prom2	E	0	8	2	/
		5	6	1	/
		7	4	1	/
	C	0	5	3	0,371
		6	4	0	/
		9	4	0	/
		12	4	0	/
P2	E	0	8	0	/
		5	6	0	/
		7	4	0	/
	C	0	5	0	/
		6	4	0	/
		9	4	0	/
		12	4	0	/
P1	E	0	8	2	/
		5	6	2	/
		7	4	0	/
	C	0	5	1	/
		6	4	0	/
		9	4	1	/
		12	4	0	/

Desmin	D1	E	0	6	1	/
			5	5	0	/
			7	3	0	/
		C	0	5	1	/
			6	5	1	/
			9	3	0	/
			12	4		/
	NK3	E	0	6	3	0,468
			5	5	0	/
			7	3	0	/
		C	0	5	3	0,218
			6	5	0	/
			9	3	2	0,207
			12	4	0	/
	D2	E	0	6	2	/
			5	5	0	/
			7	3	0	/
		C	0	4	1	/
			6	5	0	/
			9	3	0	/
			12	4	0	/
	NK2	E	0	6	4	0,173
			5	5	1	/
			7	3	1	/
		C	0	5	2	/
			6	5	0	/
			9	3	3	0,322
			12	4	2	0,117
	E2	E	0	6	2	/
			5	5	1	/
			7	3	0	/
		C	0	4	1	/
			6	5	0	/
			9	3	0	/
			12		0	/
	E1	E	0	6	3	0,135
			5	5	4	0,144
			7	3	2	0,114
		C	0	5	3	0,215
			6	5	0	/
			9	3	1	/
			12	4	2	0,106
	NK1	E	0	6	0	/
			5	5	0	/
			7	3	0	/
		C	0	5	3	0,229
			6	5	0	/
			9	3	0	/
			12	4	0	/
	Prom	E	0	6	1	/
			5	5	0	/
			7	3	0	/
		C	0	5	0	/
			6	5	0	/
			9	3	0	/

				12	4	0	/
Brachyury	Nanog	E1	E	0	1	1	0,459
				5	1	1	0,818
			C	0	1	1	0,423
				6	1	1	0,689
		D0	E	0	1	0	/
				5	1	0	/
			C	0	1	1	0,683
				6	1	1	0,687
		D1	E	0	1	0	/
				5	1	0	/
			C	0	1	0	/
				6	1	1	0,792
		D2	E	0	1	0	/
				5	1	0	/
			C	0	1	1	0,761
				6	1	0	/
		D3	E	0	1	0	/
				5	1	1	0,783
			C	0	1	1	0,792
				6	1	0	/
		D4	E	0	1	1	0,558
				5	1	0	/
			C	0	1	0	/
				6	1	1	0,891
		P0	E	0	1	1	0,359
				5	1	1	0,856
			C	0	1	1	0,687
				6	1	1	0,655
		P1	E	0	1	1	0,879
				5	1	1	0,656
			C	0	1	0	/
				6	1	0	/
		I1	E	0	1	0	/
				5	1	0	/
			C	0	1	0	/
				6	1	1	0,581
		I2	E	0	1	0	/
				5	1	0	/
			C	0	1	0	/
				6	1	0	/
		I3	E	0	1	0	/
				5	1	0	/
			C	0	1	0	/
				6	1	0	/
		I4	E	0	1	1	0,673
				5	1	0	/
			C	0	1	1	0,568
				6	1	0	/
		I5	E	0	1	0	/
				5	1	0	/
			C	0	1	1	0,705
				6	1	1	0,455
		I6	E	0	1	0	/

			5	1	0	/
		C	0	1	0	/
			6	1	0	/
	U1	E	0	1	1	0,793
			5	1	1	0,98
		C	0	1	1	0,869
			6	1	1	0,883
	U2	E	0	1	1	0,738
			5	1	0	/
		C	0	1	1	0,647
			6	1	0	/
	U3	E	0	1	0	/
			5	1	0	/
		C	0	1	0	/
			6	1	0	/
	U0	E	0	1	1	0,4
			5	1	0	/
		C	0	1	0	/
			6	1	0	/
Brachyury	NK2	E	0	9	3	/
			5	6	5	0,313
			7	4	4	0,455
		C	0	6	2	/
			6	6	0	/
			9	4	2	0,601
			12	4	1	/
	D0	E	0	9	3	/
			5	6	2	/
			7	4	4	0,231
		C	0	6	3	0,333
			6	6	1	/
			9	4	2	0,394
			12	4	0	/
	D1	E	0	9	6	0,887
			5	6	5	0,411
			7	4	3	0,543
		C	0	6	5	0,59
			6	6	3	0,503
			9	4	4	0,435
			12	4	4	0,298
	D2	E	0	9	7	0,556
			5	6	3	0,451
			7	4	0	0,051
		C	0	6	6	0,517
			6	6	5	0,388
			9	4	4	0,229
			12	4	3	0,182
	D3	E	0	9	6	0,513
			5	6	5	0,209
			7	4	4	0,115
		C	0	6	6	0,318
			6	6	0	/
			9	4	2	0,169
			12	4	2	/
	P0	E	0	9	6	0,536

			5	6	3	0,428
			7	4	4	0,431
		C	0	6	5	0,497
			6	6	3	0,142
			9	4	4	0,296
			12	4	4	0,153
	P1	E	0	9	7	0,536
			5	6	5	0,579
			7	4	4	0,827
		C	0	6	6	0,196
			6	6	2	/
			9	4	4	0,276
			12	4	4	0,145
	P2	E	0	9	6	0,496
			5	6	4	0,214
			7	4	4	0,459
		C	0	6	5	0,346
			6	6	2	/
			9	4	3	0,276
			12	4	4	0,233
	NK1	E	0	9	5	0,35
			5	6	1	/
			7	4	1	/
		C	0	6	5	0,355
			6	6	1	/
			9	4	2	/
			12	4	0	/
	P3	E	0	9	1	/
			5	6	0	/
			7	4	0	/
		C	0	6	2	/
			6	6	0	/
			9	4	0	/
			12	4	0	/
Mep1	D3	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	3	/
			6	4	1	/
			9	4	0	/
			12	4	0	/
	D2	E	0	4	2	0,256
			5	4	1	/
			7	4	0	/
		C	0	4	2	/
			6	4	1	/
			9	4	0	/
			12	4	0	/
	NK	E	0	4	2	/
			5	4	0	/
			7	4	0	/
		C	0	4	4	0,079
			6	4	0	/
			9	4	0	/
			12	4	3	0,1

	D1	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	3	0,114
			6	4	0	/
			9	4	0	/
			12	4	0	/
	T3	E	0	4	1	/
			5	4	1	/
			7	4	0	/
		C	0	4	2	/
			6	4	3	0,199
			9	4	0	/
			12	4	0	/
	T2	E	0	4	2	0,179
			5	4	0	/
			7	4	2	0,045
		C	0	4	2	0,187
			6	4	0	/
			9	4	0	/
			12	4	3	0,151
	P1	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	1	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
	T1	E	0	4	4	0,24
			5	4	2	0,295
			7	4	1	/
		C	0	4	4	0,485
			6	4	1	/
			9	4	0	/
			12	4	2	0,095
	3’UTR	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
Nkx2.5	D2	E	0	1	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	0	/
			6	4	0	/
			9	3	0	/
			12	-	-	-
	E1	E	0	5	1	/
			5	6		-
			7	-	-	-
		C	0	3	0	/
			6	6	0	/
			9	-	-	-

			12	4	0	/
	E2	E	0	5	0	/
			5	8	0	/
			7	3	0	/
		C	0	3	0	/
			6	6	0	/
			9	4	0	/
			12	4	0	/
	MCE	E	0	2	0	/
			5	2	0	/
			7	4	0	/
		C	0	2	1	0,676
			6	2	0	/
			9	4	-	-
			12	4	1	/
	D1	E	0	7	2	/
			5	7	3	/
			7	4	1	/
		C	0	4	2	0,07
			6	6	1	/
			9	4	0	/
			12	4	0	/
	Prom1	E	0	7	2	/
			5	7	1	/
			7	4	0	/
		C	0	4	2	0,132
			6	6	4	0,203
			9	4	1	/
			12	4	1	/
	Prom2	E	0	7	4	0,722
			5	8	1	/
			7	4	2	0,087
		C	0	5	3	0,094
			6	6	2	/
			9	4	0	/
			12	4	0	/
	P2	E	0	7	0	/
			5	8	0	/
			7	4	0	/
		C	0	5	0	/
			6	6	0	/
			9	4	0	/
			12	4	0	/
	P1	E	0	7	5	0,356
			5	6	4	0,302
			7	4	4	0,113
		C	0	5	3	0,223
			6	6	5	0,168
			9	4	3	0,143
			12	4	3	/
Desmin	D1	E	0	3	1	/
			5	4	0	/
			7	4	0	/
		C	0	2	0	/
			6	6	0	/

				9	5	0	/
				12	4	0	/
		NK3	E	0	8	6	0,163
				5	9	2	/
				7	4	3	0,121
			C	0	6	3	0,276
				6	10	6	0,093
				9	5	5	0,144
				12	4	0	/
		D2	E	0	3	3	0,334
				5	4	0	/
				7	4	0	/
			C	0	2	2	0,445
				6	6	2	/
				9	5	3	0,129
				12	4	0	/
		NK2	E	0	8	5	0,269
				5	9	6	0,372
				7	4	3	0,121
			C	0	6	4	0,247
				6	10	7	0,197
				9	5	4	0,114
				12	4	3	0,155
		E2	E	0	8	6	0,188
				5	8	3	/
				7	4	1	/
			C	0	6	4	0,179
				6	10	3	/
				9	5	0	/
				12	4	0	/
		E1	E	0	8	6	0,423
				5	8	4	0,177
				7	4	4	0,122
			C	0	6	4	0,204
				6	10	6	0,158
				9	5	3	0,06
				12	4	1	/
		NK1	E	0	8	1	/
				5	7	0	/
				7	4	0	/
			C	0	6	3	0,119
				6	10	0	/
				9	5	0	/
				12	4	2	/
		Prom	E	0	8	0	/
				5	7	2	/
				7	4	1	/
			C	0	6	3	0,077
				6	10	0	/
				9	5	1	/
				12	4	1	/
Mesp1	Brachyury	NK2	E	0	4	4	0,274
				5	4	0	/
				7	4	4	0,387

	C	0	6	2	/
		6	5	3	0,221
		9	4	4	0,241
		12	4	4	0,226
D0	E	0	4	2	0,374
		5	4	3	0,082
		7	4	4	0,255
	C	0	6	0	/
		6	5	3	0,152
		9	4	1	/
		12	4	3	0,156
D1	E	0	4	4	0,510
		5	4	4	0,739
		7	4	4	0,467
	C	0	6	6	0,336
		6	5	5	0,419
		9	4	4	0,266
		12	4	4	0,377
D2	E	0	4	4	0,233
		5	4	4	0,483
		7	4	2	0,245
	C	0	6	4	0,484
		6	5	5	0,431
		9	4	4	0,274
		12	4	3	0,236
D3	E	0	4	4	0,220
		5	4	4	0,245
		7	4	3	0,102
	C	0	6	1	/
		6	5	2	/
		9	4	4	0,251
		12	4	1	/
P0	E	0	4	4	0,612
		5	4	4	0,310
		7	4	4	0,505
	C	0	6	5	0,306
		6	5	5	0,255
		9	4	3	0,250
		12	4	4	0,169
P1	E	0	4	4	0,603
		5	4	4	0,465
		7	4	4	0,508
	C	0	6	6	0,286
		6	5	5	0,528
		9	4	4	0,338
		12	4	4	0,305
P2	E	0	4	2	0,247
		5	4	4	0,397
		7	4	4	0,351
	C	0	6	5	0,248
		6	5	5	0,309
		9	4	4	0,221
		12	4	4	0,093
NK1	E	0	4	4	0,168
		5	4	4	0,102

			7	4	3	0,434
		C	0	6	3	/
			6	5	1	/
			9	4	1	/
			12	4	1	/
	P3	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	6	0	/
			6	5	0	/
			9	4	0	/
			12	4	0	/
Mep1	D3	E	0	4	1	/
			5	4	2	0,116
			7	4	0	/
		C	0	4	1	/
			6	4	0	/
			9	4	0	/
			12	5	3	0,082
	D2	E	0	4	2	0,214
			5	4	2	0,082
			7	4	0	/
		C	0	4	4	0,176
			6	4	1	/
			9	4	1	/
			12	5	4	0,212
	NK	E	0	4	4	0,203
			5	4	4	0,054
			7	4	0	/
		C	0	4	2	/
			6	4	1	/
			9	4	0	/
			12	5	4	0,159
	D1	E	0	4	1	/
			5	4	0	/
			7	4	0	/
		C	0	4	2	/
			6	4	0	/
			9	4	0	/
			12	5	3	0,127
	T3	E	0	4	1	/
			5	4	3	0,118
			7	4	0	/
		C	0	4	1	/
			6	4	1	/
			9	4	1	/
			12	5	5	0,149
	T2	E	0	4	3	0,120
			5	4	2	/
			7	4	1	/
		C	0	4	2	0,436
			6	4	1	/
			9	4	2	/
			12	5	5	0,121
	P1	E	0	4	0	/

			5	4	0	/
			7	4	0	/
		C	0	4	2	/
			6	4	0	/
			9	4	0	/
			12	5	0	/
	T1	E	0	4	2	0,808
			5	4	4	0,286
			7	4	2	0,188
		C	0	4	3	0,315
			6	4	2	0,783
			9	4	3	0,230
			12	5	4	0,094
	3'UTR	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	4	0	/
			6	4	0	/
			9	4	0	/
			12	5	0	/
Nkx2.5	D2	E	0	6	0	/
			5	6	0	/
			7	4	0	/
		C	0	4	0	/
			6	4	0	/
			9	2	0	/
			12	-	-	-
	E1	E	0	4	0	/
			5	4	0	/
			7	-	-	-
		C	0	6	0	/
			6	6	0	/
			9	2	0	/
			12	2	1	/
	E2	E	0	6	0	/
			5	6	0	/
			7	4	0	/
		C	0	6	0	/
			6	6	0	/
			9	4	0	/
			12	4	2	0,225
	MCE	E	0	2	0	/
			5	2	1	/
			7	4	0	/
		C	0	2	0	/
			6	2	0	/
			9	4	0	/
			12	4	1	/
	D1	E	0	6	2	/
			5	6	0	/
			7	4	1	/
		C	0	6	2	/
			6	6	0	/
			9	4	0	/
			12	4	3	0,128

	Prom1	E	0	6	3	0,081
			5	6	4	0,262
			7	4	0	/
		C	0	6	5	0,194
			6	6	6	0,162
			9	4	1	/
			12	4	4	0,172
	Prom2	E	0	6	4	0,171
			5	6	1	/
			7	4	0	/
		C	0	6	2	/
			6	6	3	0,182
			9	4	2	0,134
			12	4	3	0,131
	P2	E	0	6	0	/
			5	6	0	/
			7	4	0	/
		C	0	6	0	/
			6	6	0	/
			9	4	0	/
			12	4	1	/
	P1	E	0	6	3	0,295
			5	6	4	0,156
			7	4	1	/
		C	0	6	4	0,214
			6	6	6	0,223
			9	4	1	/
			12	4	4	0,336
Desmin	D1	E	0	2	0	/
			5	2	0	/
			7	5	0	/
		C	0	2	1	/
			6	4	0	/
			9	3	0	/
			12	4	2	0,423
	NK3	E	0	7	2	/
			5	6	3	0,121
			7	5	3	0,085
		C	0	4	2	0,629
			6	7	2	/
			9	3	3	0,176
			12	4	4	0,252
	D2	E	0	2	2	0,147
			5	2	0	/
			7	5	2	/
		C	0	2	2	0,371
			6	4	3	0,244
			9	3	1	/
			12	4	4	0,416
	NK2	E	0	7	5	0,220
			5	6	2	/
			7	5	3	0,096
		C	0	4	2	0,170
			6	7	4	0,115
			9	3	1	/

				12	4	3	0,296
		E2	E	0	7	4	0,260
				5	6	2	/
				7	5	3	0,107
			C	0	4	3	0,180
				6	7	3	/
				9	3	1	/
				12	4	3	0,133
		E1	E	0	7	4	0,202
				5	6	3	0,218
				7	5	4	0,117
			C	0	4	2	0,208
				6	7	4	0,146
				9	3	3	0,068
				12	4	3	0,260
		NK1	E	0	7	0	/
				5	6	0	/
				7	5	0	/
			C	0	4	2	0,125
				6	7	0	/
				9	3	0	/
				12	4	2	/
		Prom	E	0	7	1	/
				5	6	2	/
				7	5	0	/
			C	0	4	0	/
				6	7	0	/
				9	3	0	/
				12	4	3	0,210
Nkx2.5	Brachyury	NK2	E	0	6	4	0,541
				5	6	3	0,504
				7	4	3	0,295
			C	0	4	2	0,533
				6	6	2	/
				9	4	3	0,491
				12	4	4	0,590
		D0	E	0	6	3	0,172
				5	6	1	/
				7	4	4	0,127
			C	0	5	1	/
				6	6	2	/
				9	4	4	0,427
				12	4	4	0,699
		D1	E	0	6	5	0,538
				5	6	5	0,650
				7	4	4	0,359
			C	0	4	2	0,645
				6	6	5	0,566
				9	4	4	0,594
				12	4	4	0,570
		D2	E	0	6	6	0,541
				5	6	4	0,391
				7	4	1	/
			C	0	5	5	0,518

			6	6	6	0,524
			9	4	0	/
			12	4	2	/
	D3	E	0	6	6	0,392
			5	6	3	0,169
			7	4	1	/
		C	0	5	4	0,261
			6	6	4	0,196
			9	4	4	0,084
			12	4	4	0,093
	P0	E	0	6	5	0,798
			5	6	6	0,474
			7	4	4	0,855
		C	0	5	4	0,677
			6	6	5	0,791
			9	4	4	0,681
			12	4	4	0,508
	P1	E	0	6	6	0,684
			5	6	4	0,251
			7	4	4	0,440
		C	0	5	5	0,405
			6	6	6	0,567
			9	4	4	0,513
			12	4	4	0,567
	P2	E	0	6	6	0,423
			5	6	4	0,234
			7	4	4	0,384
		C	0	5	4	0,310
			6	6	4	0,366
			9	4	4	0,281
			12	4	4	0,288
	NK1	E	0	6	4	0,429
			5	6	4	0,411
			7	4	4	0,192
		C	0	5	3	0,193
			6	6	5	0,285
			9	4	4	0,142
			12	4	4	0,163
	P3	E	0	6	0	/
			5	6	1	/
			7	4	0	/
		C	0	5	0	/
			6	6	0	/
			9	4	0	/
			12	4	0	/
Mep1	D3	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	3	1	/
			6	4	0	0,121
			9	4	0	/
			12	4	0	/
	D2	E	0	4	0	/
			5	4	1	/
			7	4	1	/

		C	0	3	1	/
			6	4	0	/
			9	4	0	/
			12	4	1	/
	NK	E	0	4	1	/
			5	4	1	/
			7	4	1	/
		C	0	3	1	/
			6	4	0	/
			9	4	0	/
			12	4	1	/
	D1	E	0	4	0	/
			5	4	1	/
			7	4	0	/
		C	0	3	1	/
			6	4	0	/
			9	4	0	/
			12	4	2	/
	T3	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	3	1	/
			6	4	0	/
			9	4	0	/
			12	4	2	/
	T2	E	0	4	1	/
			5	4	1	/
			7	4	0	/
		C	0	3	2	0,123
			6	4	0	/
			9	4	0	/
			12	4	1	/
	P1	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	3	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
	T1	E	0	4	1	/
			5	4	3	0,268
			7	4	3	0,292
		C	0	3	3	0,479
			6	4	4	0,570
			9	4	1	/
			12	4	3	0,188
	3’UTR	E	0	4	0	/
			5	4	0	/
			7	4	0	/
		C	0	3	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
Nkx2.5	D2	E	0	5	0	/
			5	5	0	/

		7	3	0	/
	C	0	5	0	/
		6	4	0	/
		9	4	0	/
		12	4	0	
E1	E	0	1	0	/
		5	-	-	-
		7	3	0	/
	C	0	5	0	/
		6	2	0	/
		9	4	0	/
		12	4	0	
E2	E	0	3	0	/
		5	5	0	/
		7	3	0	/
	C	0	5	0	/
		6	4	0	/
		9	4	0	/
		12	4	0	
MCE	E	0	4	0	/
		5	2	0	/
		7	-	-	-
	C	0	2	0	/
		6	2	0	/
		9	-	-	-
		12	2	2	0,144
D1	E	0	5	0	/
		5	5	0	/
		7	3	0	/
	C	0	5	2	/
		6	4	0	/
		9	4	0	/
		12	4	0	
Prom1	E	0	5	0	/
		5	5	1	/
		7	3	1	/
	C	0	5	3	0,147
		6	4	3	0,134
		9	4	2	0,116
		12	4	2	0,113
Prom2	E	0	5	0	/
		5	5	0	/
		7	3	0	/
	C	0	5	0	/
		6	4	0	/
		9	4	0	/
		12	4	0	
P2	E	0	5	0	/
		5	5	0	/
		7	3	0	/
	C	0	5	0	/
		6	4	0	/
		9	4	0	/
		12	4	0	
P1	E	0	5	5	0,333

			5	5	5	0,225
			7	3	2	0,147
		C	0	5	5	0,188
			6	4	4	0,136
			9	4	3	0,296
			12	4	3	0,150
Desmin	D1	E	0	2	0	/
			5	2	0	/
			7	4	0	/
		C	0	3	0	/
			6	2	0	/
			9	4	0	/
			12	3	0	/
	NK3	E	0	6	0	/
			5	6	4	0,198
			7	4	0	/
		C	0	6	1	/
			6	2	2	0,162
			9	4	3	0,138
			12	3	3	0,101
	D2	E	0	2	0	/
			5	2	0	/
			7	4	0	/
		C	0	3	2	0,513
			6	2	0	/
			9	4	1	/
			12	3	2	0,069
	NK2	E	0	6	3	0,217
			5	6	1	/
			7	4	2	/
		C	0	6	3	0,205
			6	2	0	/
			9	4	4	0,084
			12	3	3	0,087
	E2	E	0	6	4	0,279
			5	6	1	/
			7	4	4	0,153
		C	0	6	3	0,219
			6	2	0	/
			9	4	3	0,085
			12	3	2	0,080
	E1	E	0	6	4	0,221
			5	6	3	0,144
			7	4	4	0,092
		C	0	6	2	/
			6	2	2	0,170
			9	4	3	0,113
			12	3	2	0,099
	NK1	E	0	6	1	/
			5	6	0	/
			7	4	0	/
		C	0	6	4	0,171
			6	2	0	/
			9	4	0	/
			12	3	0	/

		Prom	E	0	6	1	/
				5	6	0	/
				7	4	0	/
			C	0	6	5	0,210
				6	2	0	/
				9	4	0	/
				12	3	0	/
Desmin	Brachyury	NK2	E	0	6	3	0,401
				5	5	3	0,138
				7	4	4	0,849
			C	0	5	4	0,255
				6	4	2	0,120
				9	4	1	/
				12	4	4	0,508
		D0	E	0	6	3	0,210
				5	5	2	0,436
				7	4	3	0,633
			C	0	5	0	/
				6	4	1	/
				9	4	0	/
				12	4	2	0,324
		D1	E	0	6	5	0,757
				5	5	4	0,544
				7	4	4	0,731
			C	0	5	4	0,543
				6	4	4	0,323
				9	4	3	0,398
				12	4	4	0,296
		D2	E	0	6	4	0,254
				5	5	5	0,220
				7	4	0	/
			C	0	5	5	0,243
				6	4	2	0,447
				9	4	0	/
				12	4	4	0,397
		D3	E	0	6	6	0,305
				5	5	5	0,211
				7	4	4	0,134
			C	0	5	3	/
				6	4	0	/
				9	4	2	0,216
				12	4	3	0,054
		P0	E	0	6	6	0,312
				5	5	3	0,268
				7	4	4	0,780
			C	0	5	5	0,362
				6	4	4	0,334
				9	4	4	0,093
				12	4	4	0,081
		P1	E	0	6	6	0,683
				5	5	5	0,580
				7	4	4	0,717
			C	0	5	4	0,266
				6	4	3	0,342

			9	4	4	0,210
			12	4	4	0,382
	P2	E	0	6	6	0,378
			5	5	5	0,303
			7	4	4	0,544
		C	0	5	4	0,175
			6	4	4	0,352
			9	4	1	/
			12	4	4	0,248
	NK1	E	0	6	3	0,495
			5	5	3	0,194
			7	4	3	0,324
		C	0	5	2	/
			6	4	1	/
			9	4	0	/
			12	4	1	/
	P3	E	0	6	3	0,350
			5	5	1	/
			7	4	0	/
		C	0	5	0	/
			6	4	0	/
			9	4	0	/
			12	4	0	/
Mep1	D3	E	0	4	0	/
			5	4	2	0,312
			7	3	0	/
		C	0	4	0	/
			6	6	0	/
			9	3	0	/
			12	4	0	/
	D2	E	0	4	2	0,290
			5	4	1	/
			7	3	0	/
		C	0	4	0	/
			6	6	3	/
			9	3	0	/
			12	4	0	/
	NK	E	0	4	3	0,102
			5	4	0	/
			7	3	0	/
		C	0	4	2	0,153
			6	6	0	/
			9	3	0	/
			12	4	0	/
	D1	E	0	4	0	/
			5	4	0	/
			7	3	0	/
		C	0	4	1	/
			6	6	0	/
			9	3	0	/
			12	4	0	/
	T3	E	0	4	1	/
			5	4	1	/
			7	3	0	/
		C	0	4	0	/

			6	6	2	/
			9	3	0	/
			12	4	0	/
	T2	E	0	4	1	/
			5	4	2	0,101
			7	3	1	/
		C	0	4	0	/
			6	6	5	0,132
			9	3	0	/
			12	4	0	/
	P1	E	0	4	0	/
			5	4	0	/
			7	3	0	/
		C	0	4	0	/
			6	6	2	/
			9	3	0	/
			12	4	0	/
	T1	E	0	4	4	0,466
			5	4	2	0,406
			7	3	1	/
		C	0	4	2	0,239
			6	6	4	0,356
			9	3	1	/
			12	4	1	/
	3’UTR	E	0	4	0	/
			5	4	0	/
			7	3	0	/
		C	0	4	0	/
			6	6	0	/
			9	3	0	/
			12	4	0	/
Nkx2.5	D2	E	0	6	0	/
			5	7	0	/
			7	4	0	/
		C	0	4	0	/
			6	6	0	/
			9	-	-	-
			12	-	-	-
	E1	E	0	4	1	/
			5	5	1	/
			7	-	-	-
		C	0	6	3	0,164
			6	4	1	/
			9	2	0	/
			12	2	0	
	E2	E	0	6	1	/
			5	7	1	/
			7	4	0	/
		C	0	6	2	/
			6	6	1	/
			9	2	0	/
			12	4	0	
	MCE	E	0	2	1	/
			5	2	0	/
			7	4	0	/

		C	0	2	0	/
			6	2	0	/
			9	4	1	/
			12	4	0	
	D1	E	0	6	3	/
			5	7	1	/
			7	4	4	0,093
		C	0	6	0	/
			6	6	1	/
			9	4	0	/
			12	4	0	
	Prom1	E	0	6	2	/
			5	7	7	0,253
			7	4	1	/
		C	0	6	5	0,131
			6	6	3	0,240
			9	4	1	
			12	4	0	
	Prom2	E	0	6	3	0,137
			5	7	2	/
			7	4	1	/
		C	0	6	1	/
			6	6	1	/
			9	4	0	/
			12	4	0	
	P2	E	0	6	1	/
			5	6	0	/
			7	4	0	/
		C	0	6	0	/
			6	6	1	/
			9	4	0	/
			12	4	0	/
	P1	E	0	6	3	0,223
			5	6	3	0,226
			7	4	3	0,247
		C	0	6	4	0,125
			6	6	4	0,210
			9	4	3	0,215
			12	4	0	0,134
Desmin	D1	E	0	6	0	/
			5	5	0	/
			7	4	0	/
		C	0	7	1	/
			6	5	0	/
			9	3	0	/
			12	4	0	
	NK3	E	0	6	4	0,230
			5	5	1	/
			7	4	0	/
		C	0	7	2	/
			6	5	2	/
			9	3	2	0,260
			12	4	4	0,121
	D2	E	0	6	3	0,190
			5	5	0	/

			7	4	3	0,183
		C	0	7	2	/
			6	5	0	/
			9	3	0	/
			12	4	2	0,211
	NK2	E	0	6	4	0,137
			5	5	3	0,107
			7	4	3	0,152
		C	0	7	2	/
			6	5	1	/
			9	3	2	0,120
			12	4	4	0,123
	E2	E	0	6	3	0,157
			5	5	3	0,129
			7	4	3	0,043
		C	0	7	3	/
			6	5	3	0,125
			9	3	1	/
			12	4	1	/
	E1	E	0	6	3	0,121
			5	5	3	0,151
			7	4	1	/
		C	0	7	6	0,126
			6	5	4	0,138
			9	3	3	0,233
			12	4	3	0,210
	NK1	E	0	6	0	/
			5	5	3	0,080
			7	4	0	/
		C	0	7	1	/
			6	5	0	/
			9	3	0	/
			12	4	1	/
	Prom	E	0	6	1	/
			5	5	0	/
			7	4	0	/
		C	0	7	1	/
			6	5	0	/
			9	3	1	/
			12	4	0	/

[(1)] E = ESCs; [(2)] C = CVPCs; [(3)] / = no significant signal; [(4)] - = no PCR done

Acknowledgements

I want to thank my supervisor Ao. Univ. Prof. Dr. Georg Weitzer for giving me the great opportunity to work in his lab and for always welcoming my own ideas regarding experimental approaches and interpretation of data. His enthusiasm and knowledge about science are both inspiring and motivating. Furthermore, I am very grateful for his proofreading this thesis and for his very helpful suggestions for improvement.

I would also like to thank my co-workers Brigitte, Lucia and Marie-Thérèse for sharing their expertise and ideas as well as for creating a not only productive but also friendly working atmosphere. I am also very grateful to my co-worker Florian, who did a great job performing a vast amount of PCRs analyzing ChIP samples. I also want to thank Mario for some very helpful advice.

I am also very thankful to Thomas and Sharif, who were always ready to help with a wide variety of questions and requests.

Last but not least I would like to thank my family and Leo for their continuous support and encouragement throughout my studies. I am especially grateful to my mother for proofreading this thesis.

Curriculum Vitae

PERSONAL DATA

Name	Hannah Paar
Date of Birth	March 11th, 1988
Nationality	Austria
Address	Hetzendorferstraße 97/house 2/door 5
	1120 Wien
e-mail	h.paar@gmx.net

EDUCATION

1994 - 1998	Primary school
	Dunkelstein, Austria
1998 - 2006	Bundesgymnasium und Bundesrealgymnasium
	Neunkirchen, Austria
2007 - 2011	University of Vienna
	Bachelor Studies Molecular Biology
2011 - present	University of Vienna
	Master Studies Molecular Biology
10/2012 – 11/2013	Master Thesis (Georg Weitzer's Lab)
	Department of Medical Biochemistry
	Max F. Perutz Laboratories,
	Medical University of Vienna

SCIENTIFIC POSTER PRESENTATION

September 2013	Poster Presentation at the 5th Annual ÖGMBT Meeting
	Transcriptional Regulation in Cardiac Stem Cells
	Paar, H., Tabery, F. and Weitzer, G.