Migration of Mesodermal Cells and Axial Elongation in Mouse Embryos Depend on the Serum Response Factor

Inaugural-Dissertation to Obtain the Academic Degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry and Pharmacy

of Freie Universität Berlin

by

Benedikt Schwartz

from Kelkheim

November, 2011

This work was performed at the Max-Planck Institute for Molecular Genetics in Berlin

Between January 2007 and November 2011

Under the supervision of Prof. Dr. Bernhard G. Herrmann and Dr. Phillip Grote

1st reviewer: **Prof. Dr. Bernhard G. Herrmann**; Max-Planck Institute for Molecular Genetics, Ihnestr. 73, 14195 Berlin

2nd reviewer: **Prof. Dr. Stephan Sigrist**; Institute of Biology-Freie Universität Berlin, Takustr.6, 14195 Berlin

Date of Defense: 17.01.2012

Acknowledgements

Table of Contents

1. Introduction

1.1. Formation of the vertebrate body

The mouse embryo is a valuable model for the analysis of all processes during early vertebrate development, including the processes of gastrulation and axial elongation. This is mainly due to the fact that the mouse genome has been completely sequenced and that there exist elaborate genetic tools which allow for the manipulation of its genome. During mouse embryogenesis, development of the body axis proceeds in a head-to-tail direction. Extension of the anterior-posterior axis (A-P axis) begins during the process of gastrulation, when the three germ layers known as the ectoderm, endoderm, and mesoderm are generated (Snow & Bennett 1978). Gastrulation leads to the generation of the embryo proper, where the area of the rostral portion of the embryo is initially much longer with respect to the rest of the body. Over the next few days, the posterior region is expanded in a progressive manner, giving rise to posterior trunk structures (Fig.1). The continuous generation of new tissue by a posterior growth zone in the embryonic caudal end is a crucial step for this extension. The emerging tissue expresses a specific set of genes, required to induce its differentiation into mesoderm and its derivatives, such as the presomitic mesoderm (psm). Additionally, the expression of specific genes such as the Hox genes confers positional identity along the A-P axis (reviewed in Kmita & Duboule 2003). Coinciding with the extension of the A-P axis at the embryonic caudal end is the periodic segmentation of the psm at its anterior end. Segmentation gives rise to pairs of somites, which are blocks of epithelial tissue located along either side of the neural tube which will form the ribs, vertebrae, and trunk musculature (reviewed in Christ & Ordahl 1995) (Fig.1,B).

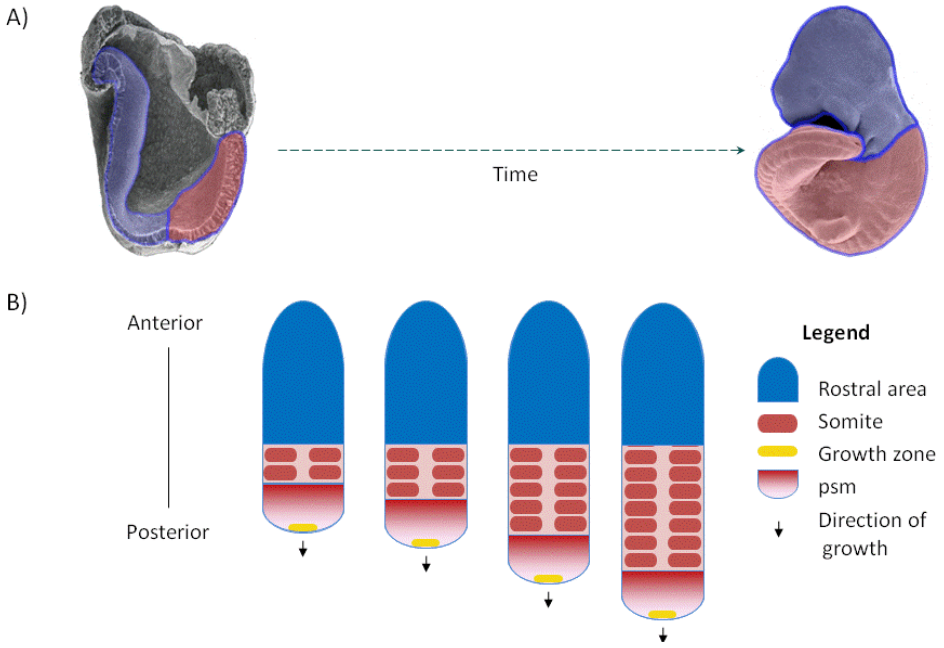

Figure 1: *A) During mouse embryogenesis, the body grows in a head-to-tail orientation. At embryonic day 8, the future head area (blue) is still overrepresented with respect to the trunk area (red). By embryonic day 10 (right), the rostral and caudal portions of the embryo have become more equally represented as a result of progressive axis extension. B) Schematic representation of axis extension. While the size of the head region (blue) remains the same, the trunk area (red) is progressively extended in a posterior direction. The extension of the axis is maintained by the activity of a growth zone (yellow) at the caudal end of the embryo, which continuously generates new tissue. At its anterior end, the psm is segmented into somites.*

1.1.1. Early gastrulation

During mouse embryogenesis, specification of the A-P axis is established prior to gastrulation, at embryonic day (E) 6.0. At this stage, the embryo consists only of the epiblast, a cup-shaped layer of cells, which represents the embryo proper (Fig.2,B), and the surrounding extraembryonic tissues, such as the visceral endoderm (VE). The anterior part of the VE (anterior visceral endoderm, AVE) is a structure, which at E6.0 secrets antagonists of Wnt (wingless-type MMTV integration site) and TGFβ (transforming growth factor beta) signaling (reviewed in Beddington & Robertson 1999; Belo et al. 1997; Bertocchini & Stern 2002; Perea-Gomez et al. 2002). These signals were initially active in all epiblast cells, but now their activity becomes restricted to the caudal part, where they establish a posterior identity. TGFβ, Wnt, and FGF (Fibroblast Growth Factor) signals induce the intercalation of epiblast cells which leads to the formation of a temporary structure, called the primitive streak (ps), by E6.5 (Conlon et al. 1994; Bertocchini et al. 2004; Lawson et al. 1991; Liu et al. 1999). The ps is a

hallmark of gastrulation, since all cells that differentiate into mesoderm and endoderm will ingress through it, while cells that remain in the epiblast will form the ectodermal layer (Fig.2). Cells within the ps activate genetic programs that affect both, cell fate specification as well as migratory behavior, and both processes are under the control of the TGFβ, Wnt, and FGF signaling pathways.

Cells in the epiblast have an epithelial morphology, which is marked by tight cell-cell contacts, presence of a basal lamina and apicobasal polarization. In contrast, cells of the emerging mesoderm and endoderm have mesenchymal characteristics (reviewed in Nakaya & Sheng 2008). They display much looser cell-cell contacts and gained the ability to migrate as individual cells. The transition between these two states (epithelial to mesenchymal transition, EMT) occurs in the ps (in more detail in 1.1.3.). Cells which have completed the EMT process migrate laterally from the ps and localize between the epiblast and the visceral endoderm. A subset of cells differentiate into definitive endoderm, gradually replacing the visceral endoderm, while the vast majority of ingressing cells differentiate into mesoderm (Poelmann 1980) (Fig.2,A).

As gastrulation proceeds, more and more epiblast cells migrate through the ps, undergo EMT, and differentiate into endoderm and mesoderm; resulting in the continuous generation of new tissue. At the same time, the ps increases in length, until it reaches the middle of the A-P axis of the embryo by E7.5. At the anterior end of the ps, the node is formed, an organizing structure with axis-inducing properties and crucial for the subsequent elongation of the A-P axis (Beddington 1994) (Fig.2,B). Cells that ingress through the node form the foregut, head mesoderm and a mesodermal midline structure termed the notochord (Schoenwolf et al. 1992). The notochord produces secreted factors that signal to all surrounding tissues, providing position and fate information, for example for the synchronization of neural tube development. Additionally, it serves as the axial skeleton of the embryo until other elements, such as the vertebrae, form (reviewed in Stemple 2005).

The elongation of the A-P axis begins when a complete layer of mesoderm is formed between ectoderm and endoderm (Fig.2,B). The tissues at this stage are mainly derived from relocalization of epiblast cells that were generated before the onset of gastrulation. They form extraembryonic structures such as the allantois and the yolk sac, as well as the anterior structures of the embryo, namely the head, the heart, and the upper trunk (Kinder et al. 1999; Tam et al 1999).

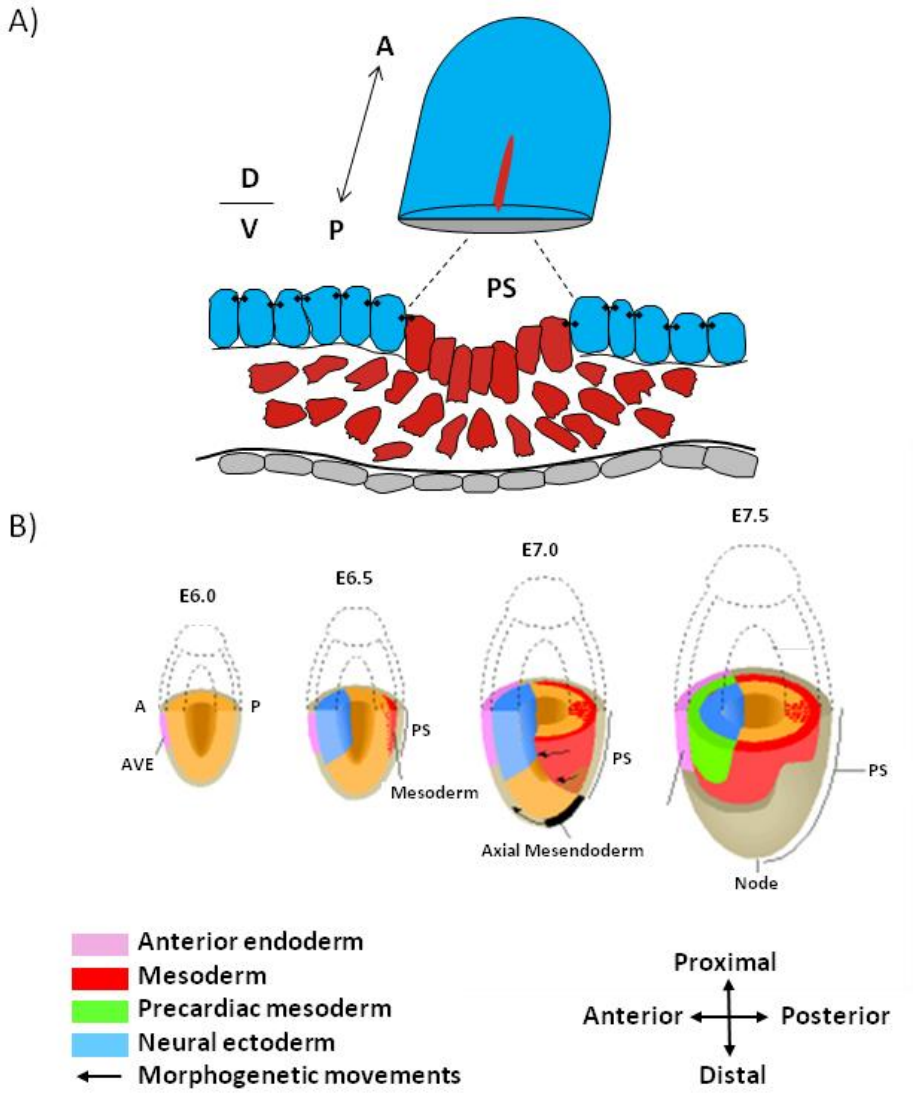

Figure 2: *Gastrulation in mouse embryos. A) At E6.5 the embryo consists of a single layer of cells called the epiblast (blue), which lies above the extraembryonic tissue (grey) (both depicted as a flattened disc for simplification). Cells in the posterior part of the epiblast form the ps (red) where they undergo EMT. During EMT they break down the basal lamina (depicted as a grey line at the apical side of the epithelium) and the tight cell-cell contacts (black) and migrate between the epiblast and the visceral endoderm to form mesoderm (red) (adapted from Chuai & Weijer, 2009). B) Schematic representation of a gastrulating embryo. Extraembryonic regions overlying the epiblast are indicated with dashed lines, and the epiblast (orange) is shown in its classical cup shaped form. At the posterior side of the embryo, epiblast cells migrate through the ps from E6.5 onwards. They form mesoderm (red), definitive endoderm (grey), and axial mesendoderm (black). Mesodermal cells move along the epiblast (indicated with arrows). At E7.5 a complete layer of mesoderm is formed between the epiblast and the definite endoderm, and the node is visible at the anterior site of the ps. Precardiac mesoderm (green) ingresses through the ps at E7.5 and moves towards the anterior most region to form the heart. A: anterior; AVE: anterior visceral endoderm; P: posterior; ps: primitive streak* (adapted from Yamaguchi 2001)

1.1.2. Elongation of the A-P axis

During the course of axial elongation, mesodermal cells are continuously generated at the caudal end of the embryo, successively contributing to more posterior structures of the trunk. Concurrent with mesoderm generation, the ps regresses posteriorly and is replaced by the tailbud at around E9.25-E9.5 (Fig.3,B), which forms after the closure of the neuropore at the posterior tip of the embryo (Catala et al. 1995). In contrast to the generation of early mesoderm, the generation of posterior mesoderm not only depends on the ingression of epiblast cells, but is also based on the activity of mesodermal progenitor cells that are located in and adjacent to the regressing ps, and later on in the tailbud (Tam & Beddington 1987; Wilson & Beddington 1996). Ingression of epiblast cells stops around E10.5 when EMT is inhibited, indicating that further axial elongation relies solely on the release of differentiating cells from the progenitors (Ohta et al. 2007).

Cells, which have ingressed through the ps and differentiated into psm proceed to migrate. Recently, it has been shown in chick embryos, that movement of posterior psm cells is crucial for the extension of the postcranial AP axis (Bénazéraf et al. 2010). A key signaling pathway for the regulation of cellular motility is the FGF pathway. FGF signaling induces the process of EMT in the ps, enabling the cells to detach from neighboring cells. It is also thought to control cell migration within the psm (Delfini et al. 2005). Consequently, disruption of cell motility, either by interference with the FGF pathway or by treatment with cell motility inhibitors, slows down axial elongation in chick embryos (Bénazéraf et al. 2010).

In concert with extension of the body axis at the posterior end of the embryo, the psm is subdivided into repeating segments, called somites, at its anterior end (Fig.3). The generation of a somite pair is a periodic process consisiting of maturation and reepithelialization, and the number of somite paires directly reflects the length of the axis. Patterning of the psm occurs at the determination front, dividing the psm into immature, loose mesenchyme posteriorly, and a more structured anterior part with cells that are progressively activating the segmentation program (Fig.3,A). Somitogenesis is regulated by a complex mechanism that includes opposing gradients of signaling molecules from the FGF and the Wnt families posteriorly and retinoic acid (RA) signaling anteriorly, as well as cell-intrinsic cyclic gene expression. This process is referred to as the segmentation clock (reviewed in Aulehla & Herrmann, 2004).

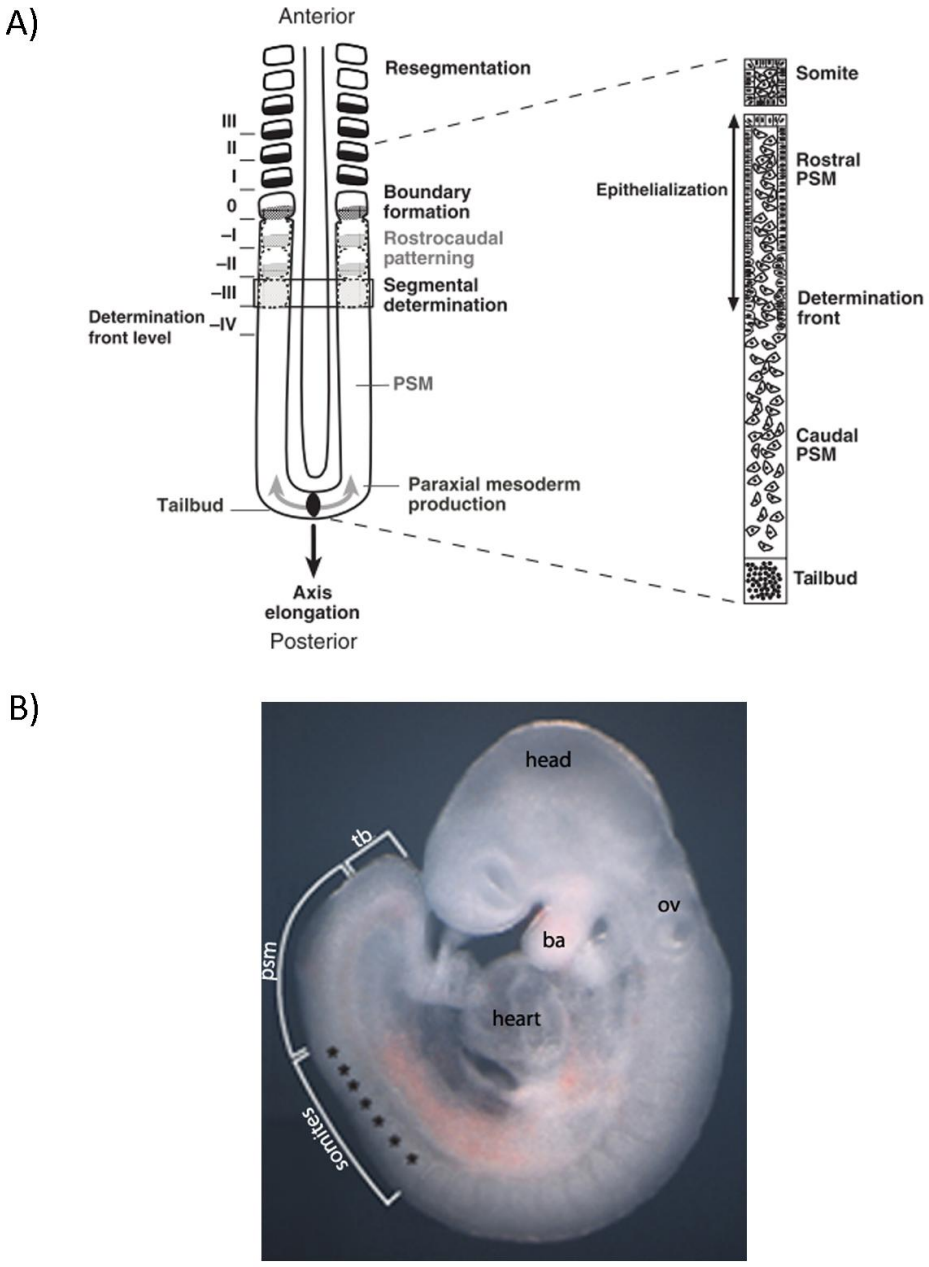

Figure 3: *Elongation of the vertebrate A-P axis. A) Schematic representation of axis extension in vertebrates. Left: New paraxial mesoderm is generated in the tailbud at the posterior end of the embryo, contributing to the unsegmented presomitic mesoderm (psm). More anteriorly, the psm cells mature into somites. Right: A parasaggital section through the psm illustrates the cellular organization of cells along the A-P axis. While the posterior part of the psm is composed of loose mesenchyme, cells anteriorly to the determination front start to reepithelialize* (Dubrulle & Pourquié, 2004). *B) At E9.5, the mouse embryo is in the process of axial elongation. Mesoderm that formed the head, the heart, and the upper trunk were generated during gastrulation. More caudally, somites have emerged from the psm. Asterisks indicate the most recently formed somites. The psm is*

continuously renewed by cells generated in the tb (Aulehla & Herrmann 2004). Ba: branchial arch, ov: otic vesicle, tb: tailbud

1.1.3. Cellular events during mesoderm development and axial elongation

The cellular processes that are fundamental to the extension of the vertebrate body axis can be divided into EMT, specification, proliferation, and migration. These processes are coordinated by a network of signaling pathways, including TGFβ, FGF, and Wnt signaling. A protein that was shown to be fundamental to mesoderm development and axis extension and which is involved in the regulation of all of the processes mentioned above is the T-box transcription factor Brachyury (T) (Kispert et al. 1995; Martin & Kimelman 2008; Wilson et al. 1993; Wilson & Beddington 1997). *T* expression is initially found in the ps and later on in the tailbud, as well as in cells of the unsegmented psm (Herrmann 1991). Embryos homozygous for a *T* mutation generate insufficient amounts of mesodermal cells. Consequently, they fail to extend the A-P axis and they lack the notochord and the extraembryonic allantois (Yanagisawa et al. 1981).

Prior to EMT, cells in the epithelial layer are connected to each other by adherens junctions. A key factor in adherens junction maintenance is the cell adhesion molecule E-cadherin, encoded by the gene *Cdh1*. In the ps, FGF signals induce the expression of the transcriptional repressor *Snai1*, which, in turn, represses *Cdh1* expression, a process that is sufficient to induce EMT (Ciruna & Rossant 2001). The cells lose the tight cell-cell contacts and instead establish transient adhesive interactions with the extracellular matrix (ECM) as well as with neighboring cells, mediated by the expression of mesenchyme-specific genes encoding vimentin (*Vim*), Matrix metalloproteinases 2, 3 and 9 (*Mmp2, Mmp3, Mmp9*) and the ECM protein Fibronectin (*Fn*) (reviewed in Lee et al. 2006). Additionally, the cytoskeleton is reorganized, confering to these cells the ability to migrate individually away from the ps. At the same time, cells that ingress through the ps and contribute to axial elongation differentiate into the mesodermal lineages. Specification of trunk and tail paraxial mesoderm is regulated, at least in part, by T and its downstream target T-box 6 transcription factor (Tbx6). *T* expression is activated by the Wnt ligand Wnt3a, which is thougt to be the main player in mesoderm specification. Mutant embryos that lack T, Tbx6 or Wnt3a do not form caudal somites (Takada et al. 1994; Yamaguchi et al. 1999; Wilkinson et al. 1990; Chapman & Papaioannou 1998). While early mesoderm is generated from epiblast cells that ingress through the ps, further axis extension depends on the generation of new mesodermal cells by mesodermal progenitors. Maintainance of the pool of mesodermal progenitor cells relies on the control of cell proliferation

and also on cell survival in the ps and later on in the tailbud. Fundamental to the maintainance of this cell population is an autoregulatory loop between T and Wnt signaling, and the expression of FGF8. Interference with the T-Wnt loop or a lack of FGF8 in the progenitor population leads to a premature termination of axis extension and hence to body truncation (Martin & Kimelman 2008).

The extension of the A-P axis requires migration of both, cells that ingress through the ps as well as cells that originate from mesodermal progenitors. Although there are different modes of mesenchymal cell migration, some common features are displayed by all migrating mesenchymal cells. Importantly, they are polarized, having a front, or leading edge, and a rear, or trailing edge. In order to move, the cells extend their plasma membrane in the form of protrusions at the leading edge (Fig.4). The generation of protrusions requires extensive actin polymerization, a process in which monomeric globular actin (G-actin) is integrated into filamentous actin (F-actin). Therefore, proper cell migration requires the availability of sufficient amounts of actin monomers. They are generated by breaking down older filaments, for example those at the rear of the cell, by the proteins Gelsolin and Cofilin, as well as by *de novo* synthesis (reviewed in Pollard & Cooper 2009; Kislauskis et al. 1997). Actin filaments are not only found in cell protrusions but also in the body of migrating cells where they act in cooperation with Myosins to form stress fibers (Fig.4). Stress fibers traverse the whole cell and usually terminate in focal adhesions (FA), a multiprotein complex consisting of integrins and adapter proteins, such as Talin and Paxillin, which link the actin filaments to the ECM (reviewed in Hammerschmidt & Wedlich 2008). FA are formed at the leading edge of the cell. At first, they are small protein complex called focal complexes. Focal complexes have a high turnover rate, but some mature into stable FA, wherein they recruit more proteins such as Zyxin. Once in place, a FA remains stationary with respect to the ECM, and the cell uses this as an anchor through which it can pull itself over the ECM. Actomyosin contractile forces generated from stress fibers pull against FA and induce retraction of the rear of the cell and detachment of trailing adhesions (McHardy et al. 2005).

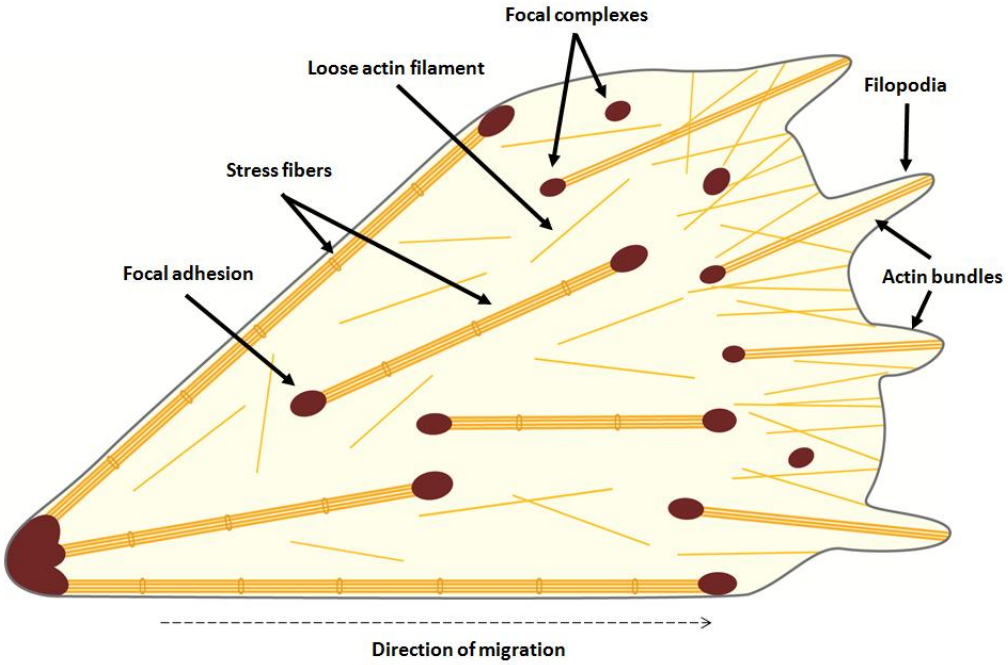

***Figure 4**: Schematic representation of a migrating cell. For a detailed description of the different components see text.*

There are various modes of cell migration, which mainly differ in the form of the protrusion formed by the cells at their leading edge. In zebrafish gastrulation, mesodermal cells use a combination of different protrusion types, such as lamellopodia, filopodia, and blebs (Diz-Muñoz et al. 2010). However, independently of which protrusion is used, all forms of cell migration require a dynamic actin cytoskeleton (Ridley 2011). Also, the regulation of cell migration and cytoskeletal rearrangement is common to all eukaryotic cells, with members of the Rho GTPase family proteins being key factors. The way that Rho signaling is controlled in a particular cell can vary. Numerous signaling pathways have been shown to regulate Rho activity, including FGF, TGFβ, Wnt and E-cadherin signaling (reviewed in Olson & Nordheim 2010). The Rho GTPases encompass members of the Rho, Rac, and Cdc42 subfamilies, which have both common and specific functions in the regulation of cell motility. Briefly, they promote actin polymerization at the cell periphery to generate protrusive forces (Cdc42), the formation of focal complexes (Rac) and their maturation to focal adhesions (Rho) as well as stress fiber assembly (Rho) (Jaffe & Hall 2005). Thus, coordination of signaling cascades is essential to ensure that the right cells are able to travel particular distances in order for proper axis extenison to occur.

1.2. The transcription factor Srf

The progressive extension of the body axis is coordinated by a network of signaling pathways, including Wnt, TGFβ and FGF signaling. Secreted signaling molecules, called morphogens, bind to specific receptors in the cell membrane and activate intracellular cascades. At the end of each cascade is the transcriptional regulation of target genes, which can be mediated by one or several transcription factors. Induction of transcription factor activity can occur by its nuclear translocation, as in the case of β-catenin, which is downstream of the canonical Wnt signaling cascade (Yost et al. 1996). It can also include phosphorylation and subsequent nuclear translocation, as in the case of the TGFβ mediators Smad1 and Smad2 (Candia et al. 1997).

The transcription factor Serum Response Factor (Srf) is implicated in the mediation of several signaling pathways, comprising TGFβ, FGF, Integrin and E-cadherin signaling (reviewed in Olson & Nordheim 2010). It is encoded by a single gene, which is conserved from yeast to humans (Müller & Nordheim, 1991). In the mouse, Srf is composed of 508 amino acids, and includes a DNA binding domain and a transactivation domain (Fig.5,A). Srf binds to the DNA as a homodimer, and has a high specificity for a palindromic 10-bp DNA sequence called the CArG-box (Fig.5,B). Srf also binds to "imperfect" CArG boxes, which usually differ by one, in rare cases by two, bases from the conventional CArG element (CC A/T_6 GG) (Pellegrini et al. 1995; Reinke et al. 2008).

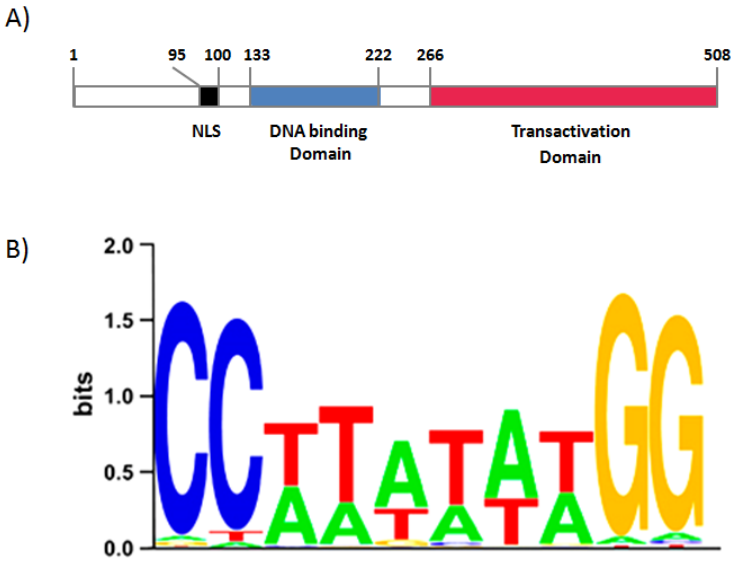

***Figure 5**: A) Schematic representation of the Srf protein. Srf is composed of 508 amino acids, which are represented by a nuclear localization signal (NLS), a DNA binding domain and a*

transactivation domain. B) Consensus sequence of the Srf binding site, called the CArG box (Miano, 2010).

To date more than 200 Srf target genes have been identified and experimentally validated (Miano, 2010). Srf targets can be divided into several classes. Among them are immediate-early genes (IEG), whose expression is activated by mitogenic stimuli independently of new protein synthesis, cytoskeletal genes, and cell type-specific genes. The latter are mainly involved in the regulation of differentiation, e. g. into smooth muscle, skeletal muscle, and cardiac muscle (Miano 2003).

The ability of Srf to regulate different sets of downstream genes depends on its association with different types of cofactors. While Srf itself has a relatively low intrinsic transcriptional activity, the interaction with cofactors confers strong activation potential in a cell- and context-specific manner (Miano 2010). Srf cofactors can be tissue-specifically expressed; otherwise they are rather ubiquitous, but specifically activated in response to mitogens. Hence, Srf serves as a platform to mediate cell intrinsic signals as well as extracellular stimuli at the transcriptional level by engaging various partners. Two families of mitogen-regulated Srf cofactors have been characterized: the ternary complex factor (TCF) family, and the myocardin related transcription factor (MRTF) family. Recently it was shown that both MRTF and TCF interact with the same hydrophobic groove and pocket on the Srf-DNA binding domain, which makes their binding to Srf mutually exclusive (Zaromytidou et al. 2006).

The TCF family of Srf cofactors comprises three different proteins, Elk1, Elk3 (also called Net or Sap-2), and Elk4 (Sap1), and is a subgroup of the superfamiliy of E twenty-six (ets) factors (reviewed in Buchwalter et al. 2004). TCF proteins elicit sequence-specific DNA contacts with ets binding sites, located in close proximity to the Srf-specific CArG-boxes, and additionally interact directly with Srf (Latinkic et al. 1996). TCF activity is regulated by FGF/Mitogen activated protein kinase (MAPK) signaling, which induces the phosphorylation of TCF by extracellular signal-regulated kinase (ERK). In concert, activated TCF and Srf induce the transcription of target genes, including genes that are implicated in the control of cell growth such as the IEGs *Fos* and *Egr1* (Shaw et al. 1989)(Fig.6).

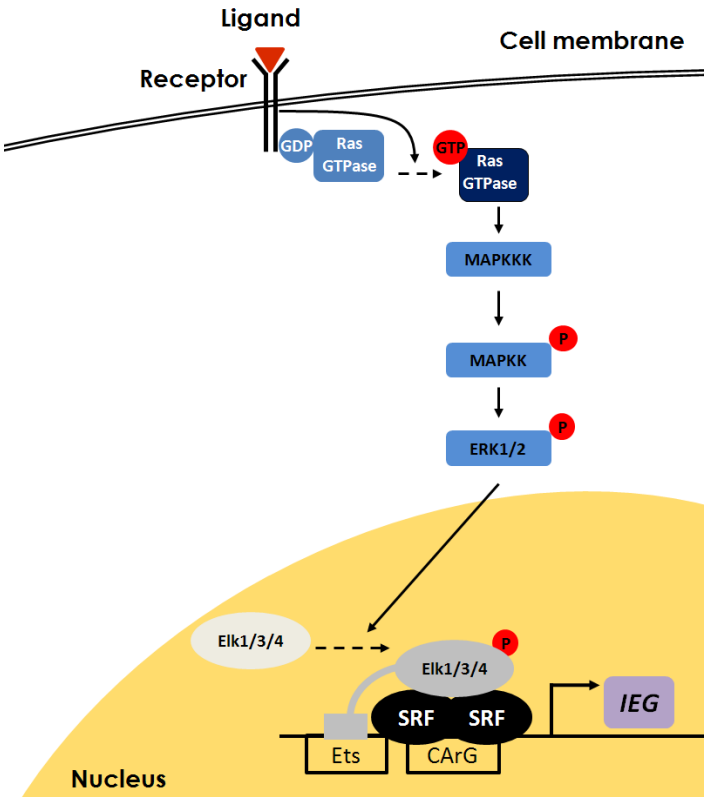

Figure 6: *The TCF family of Srf cofactors is activated by the Ras-MAPK-ERK pathway. Ligand binding activates the G protein Ras, which in turn activates the MAPK pathway. In the nucleus, the extracellular-signal-regulated kinase (ERK) phosphorylates the TCF factors (Elk1, Elk3, and Elk4). Phosphorylation increases their binding affinity to Srf and to specific DNA sequences called Ets sites which are adjacent to the CArG site* (Sharrocks 1995)*. The Srf-TCF complex induces transcription of IEGs such as* Egr1 *and* Fos*.*

The second family of signal-regulated Srf cofactors comprises Myocardin and the Myocardin related transcription factors MRTF-A (also called MAL, MKL1 or BSAC) and MRTF-B (also called MKL2 or MAL16). MRTFs were shown to bind to DNA adjacent to CArG sites, however, no sequence-specific interaction has been identified so far (Zaromytidou et al. 2006). Myocardin, the founding member of this group of cofactors, is expressed in cardiac and smooth muscle tissue and induces, in combination with Srf, the transcription of muscle-specific genes. MRTF-A and MRTF-B are expressed in numerous embryonic and adult tissues and, similar to myocardin, they have essential roles in muscle cell differentiation (Selvaraj & Prywes 2003; Schratt et al. 2002). Additionally, MRTF-A and MRTF-B, in concert with Srf, are regulators of the assembly of the actin cytoskeleton. Many genes encoding monomeric actins (*Actb, Actc, Actg, Acta*) or actin binding proteins such as vinculin and tropomyosin,

are under Srf-MRTF directed transcription (Posern & Treisman 2006; Miano et al. 2007; Knöll & Nordheim 2009) (Fig.7).

The actin cytoskeleton however is not only a downstream effector of Srf-MRTF activity, but also regulates Srf-MRTF activity. Unpolymerized G-actin binds to a region within the MRTF protein, which contains there RPEL (encoding Arg, Pro, Glu, and Leu) motifs, and sequesters it in the cytoplasm (Guettler et al. 2008). It thereby inhibits the formation of the Srf-MRTF complex in the nucleus, and consequently the transcription of target genes (Miralles et al. 2003) (Fig.7). As mentioned above, the assembly of G-actin into actin filaments is regulated by Rho-signaling, and the incorporation of monomeric actin into F-actin releases MRTF, which translocates into the nucleus and binds to Srf. Therefore, Srf-MRTF activity is regulated by Rho-actin signaling. This makes MRTF an "actin sensor" which, together with Srf, mediates the requirement of components of the cytoskeleton at the transcriptional level (reviewed in Knöll 2010).

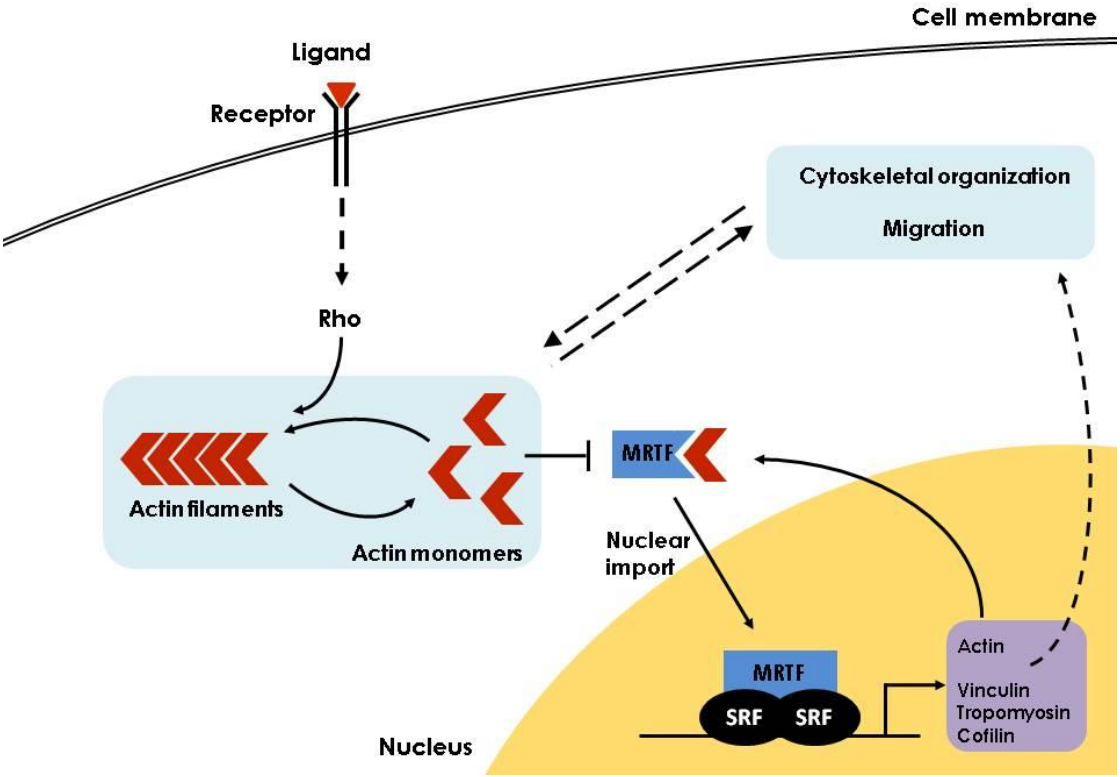

Figure 7: *Srf and MRTF are both regulated by and regulators of the actin cytoskeleton (Adapted from Pipes et al. 2006): The MRTF family of Srf cofactors is regulated by Rho-actin signaling. Members of the Rho family induce the assembly of monomeric G-actin into F-actin, releasing the Srf cofactor MRTF, which in turn binds to Srf in the nucleus and induces transcription of downstream target genes.*

Many of the Srf-MRTF targets encode actin monomers or proteins that are regulators of actin cytoskeletal assembly.

In addition to the signal-regulated TCF and MRTF cofactor families, multiple other factors can bind to Srf and change its transcription factor activity. Examples are Gata4 and Nkx2-5, which bind to Srf and regulate early cardiac gene activity (Sepulveda et al. 2002). In addition, the protein four and a half LIM domains 2 (Fhl2), itself encoded by an Srf target gene, is an example of a factor that negatively regulates Srf activity and antagonizes MRTF-Srf-dependent transcription (Philippar et al. 2004).

1.3. Diverse roles for Srf during embryonic development

Srf is a versatile transcription factor, implicated in many processes during embryogenesis and adulthood, and it can regulate diverse genetic programs in the control of proliferation and apoptosis, differentiation, cell spreading, adhesion, and migration.

During embryogenesis, Srf was expected to be required for cell proliferation, because of its ability to activate the cell cycle regulators *Egr1* and *Fos* (Richard Treisman 1986; Poser et al. 2000). However, ES cells that lack both alleles of the *Srf* gene (*Srf*$^{-/-}$ ES cells*)* proliferate normally and *Srf*$^{-/-}$ mouse embryos are able to grow and develop normally until E6.5, indicating that *Srf* is not essential for proliferation during early embryonic development (Arsenian et al. 1998). The earliest event during embryogenesis which was found to be dependent on Srf activity was the onset of gastrulation. *Srf*$^{-/-}$ mouse embryos fail to form a ps, do not develop any mesodermal cells, and die shortly thereafter. *T* expression cannot be detected, and expression of the Srf target genes *Egr1*, *Fos,* and *α-actin* is severely impaired (Arsenian et al. 1998). Therefore, Srf is thought to be an essential regulator of mesoderm formation at the onset of gastrulation. However, *Srf*$^{-/-}$ ES cells that are cultivated as cell aggregates known as embryoid bodies are capable of activating mesodermal genes, and teratomas that formed upon injection of *Srf*$^{-/-}$ ES cells under the skin of nude mice contained various mesodermal lineages. These results suggested that although Srf is required at the onset of gastrulation, it is not essential for mesoderm induction (Weinhold et al. 2000).

Instead of having a defect in cell proliferation or in mesoderm induction, ES cells deficient for Srf display impaired cell spreading, adhesion, and migration. The defects correlate with a lack of

formation of cytoskeletal structures, such as actin stress fibers and focal adhesion (FA) plaques, indicating that in ES cells Srf is required for the organization of the actin cytoskeleton (Schratt, Philippar, et al. 2002). Schratt and colleagues also showed that Srf regulates anti-apoptotic genes such as *Bcl-2* and thus promotes cell survival in murine ES cells that differentiate *in vitro* (Schratt et al. 2004).

Further information about Srf and its roles during embryo development comes from conditional knock out (KO) studies, using the *Cre/loxP* system. Wiebel and colleagues created a genomic floxed *Srf* allele (named Srf^{flex1} for floxed exon1), and demonstrated that Cre-mediated deletion of the floxed sequence converts the Srf^{flex1} allele into a null allele (Wiebel et al. 2002) (Fig.8). Instead of knocking out the gene in the entire embryo, and thus causing an arrest in development at E6.5, employment of a Cre recombinase that is expressed under a tissue-specific and/or an inducible promoter allows one to target the gene in a certain tissue and at a certain time.

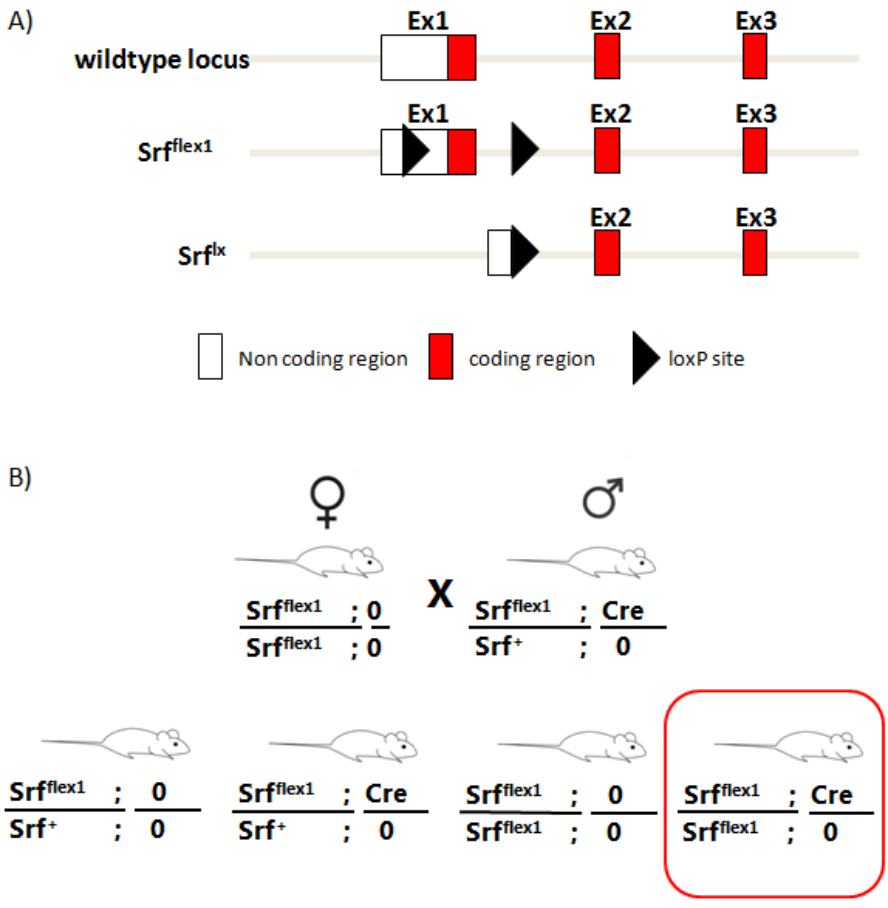

Figure 8: *Generation of conditional* Srf *KO embryos. A) In the* $Srf^{flex1/flex1}$ *line, the first exon, which also includes the translation start site of the* Srf *gene, is flanked by* loxP *sites (black arrowheads). Cre mediated recombination results in embryos that have lost Srf activity* (Srf^{lx}). *B) Breading scheme for*

conditional targeting of Srf *using the* Srf^{flex1} *line and a Cre recombinase under control of a tissue-specific promoter. Crossing of a mouse homozygous for the* Srf^{flex1} *allele with a mouse that expresses the Cre recombinase and has one* Srf^{flex1} *allele generates 25% conditional KO (red box) embryos in the next generation.*

Several conditional *Srf* KO mouse lines have been analyzed in recent years. Cardiac muscle-specific inactivation of *Srf* resulted in embryos with cardiac defects that were lethal between E10.5 and E13.5. The embryos displayed a thin myocardium, dilated cardiac chambers, poor trabeculation, and a disorganized interventricular septum (Parlakian et al. 2004). Molecular analysis of the conditional KO embryos revealed that Srf is essential for cardiomyocyte differentiation, while it does not affect cardiomyocyte proliferation. Conditional KO of *Srf* using the forebrain-specific *Camk2a* Cre resulted in mice with severe balance impairments and lack of interest in feeding, leading to a reduced body weight and finally death around postnatal day (P) 21. This phenotype was shown to be caused by a lack of neuronal migration. $Srf^{-/-}$ neurons accumulated in the subventricular zone instead of migrating into the olfactory bulb, and *in vitro* matrigel assays confirmed an impaired migration of *Srf*-deficient neurons. Molecular analysis revealed a reduced actin fiber density in *Srf*-deficient neurons, caused by the downregulation of *Actb* and impaired activity of the actin severing factors Cofilin and Gelsolin (Alberti et al. 2005). Analysis of the *Alfp*-Cre meditated KO of *Srf* in the liver revealed that, in hepatocytes, Srf regulates the expression of genes that are implicated in the control of cell proliferation. Additionally, it regulates the expression of several genes with roles in hepatic function (Sun et al. 2009). Further conditional KO studies demonstrated roles for Srf in endothelial cells, keratinocytes, skeletal muscle cells, T- and B lymphocytes and megakaryocytes (reviewed in Miano 2010), and it has also been linked to human diseases such as cancer (Bai et al. 2009) and heart disease (Chang et al. 2003; Parlakian et al. 2005). In human heart failure, an increased expression of a natural occurring dominant negative version of SRF has been shown to be elevated. The dominant negative version results from alternative splicing of the *SRF* transcript and inhibits expression of SRF dependent genes (F. J. Davis et al. 2002).

1.4. What role does Srf play during axial elongation?

Although *Srf* is often described as being ubiquitously expressed, its expression during vertebrate embryogenesis is restricted to certain domains. At E7.5, strong *Srf* expression is found in the ps and

the lateral plate mesoderm (LPM). The LPM contains cardiac progenitor cells that migrate to the midline of the embryo and form a linear heart tube. At E8.5, *Srf* expression can be found in the heart tube, the allantois, in developing somites, the ps, and in nascent mesoderm (Barron et al. 2005). Similarly, in chick and frog embryos, strong *Srf* expression was detected in the ps and mesodermal derivatives (Mohun et al. 1991; Croissant et al. 1996), which indicates a conserved role for *Srf* in the processes of mesoderm development and/or axial elongation. In the course of this work, I aimed to clarify this role. The analysis of Srf function was based on two main approaches:

(i) The genome-wide analysis of Srf binding sites in the ps and nascent mesoderm in order to confirm known and identify novel putative Srf target genes.

(ii) The identification of genes that are dysregulated in embryos lacking *Srf* specifically in the ps and mesodermal cells. This involved the generation of conditional *Srf* KO embryos using the Cre/loxP system. To overcome the early lethality of the full KO of *Srf* (Arsenian et al. 1998), the Cre recombinase was under control of a tissue-specific promoter.

The combination of the data sets obtained from both methods allowed for the identification of genes that are directly dependent on Srf activity. Furthermore, analysis of those genes and their functions provided insight into the role of Srf during mesoderm development and axial elongation.

Moreover, since Srf was shown to mediate FGF signals in various processes (Treisman 1996; Olson & Nordheim 2010; Sutherland et al. 1996; Guillemin et al. 1996; Affolter et al. 1994), and FGF signaling is crucial during mesoderm development and axis extension, I aimed to study the possible connection between Srf and FGF signaling during embryonic caudal development.

2. Results

2.1. Identification of direct Srf target genes

2.1.1. *Chromatin Immunoprecipitation*

Transcription factors bind to regulatory regions of their target genes and thereby either induce or repress their transcription. Since Srf is a transcription factor, the identification of genes that are directly bound by Srf in the ps and in nascent mesoderm is essential to understand its role during axis extension. Chromatin immunoprecipitation (ChIP) is a powerful approach to assess the target genes of a transcription factor. In a typical ChIP assay, DNA and proteins are cross-linked to maintain the association of transcription factors with their target DNA sequence. Next, chromatin is fragmented and used for immunoprecipitation of transcription factor-DNA complexes with antibodies specific for the transcription factor of interest. The precipitated (ChIP) DNA is purified and analyzed by quantitative polymerase chain reaction (qPCR) for single gene analysis or by massive parallel sequencing (ChIP-Seq) for genome-wide analysis. In many cases a limiting factor for the ChIP experiment is the availability of a suitable antibody, which is specific to the transcription factor of interest. Not all antibodies are applicable for ChIP, since cross-linking the transcription factor during the ChIP experiment may change the accessibility of the epitope. Moreover, the antibody-transcription factor complex has to be strong enough to endure the stringent washing procedure.

In the course of this work, I aimed to identify Srf target genes in the primitive streak and nascent mesoderm. To this end, I performed Srf-ChIP on embryonic caudal ends at E9.5 and pluripotent embryo cells in culture. The workflow of the ChIP experiment is shown in Figure 9.

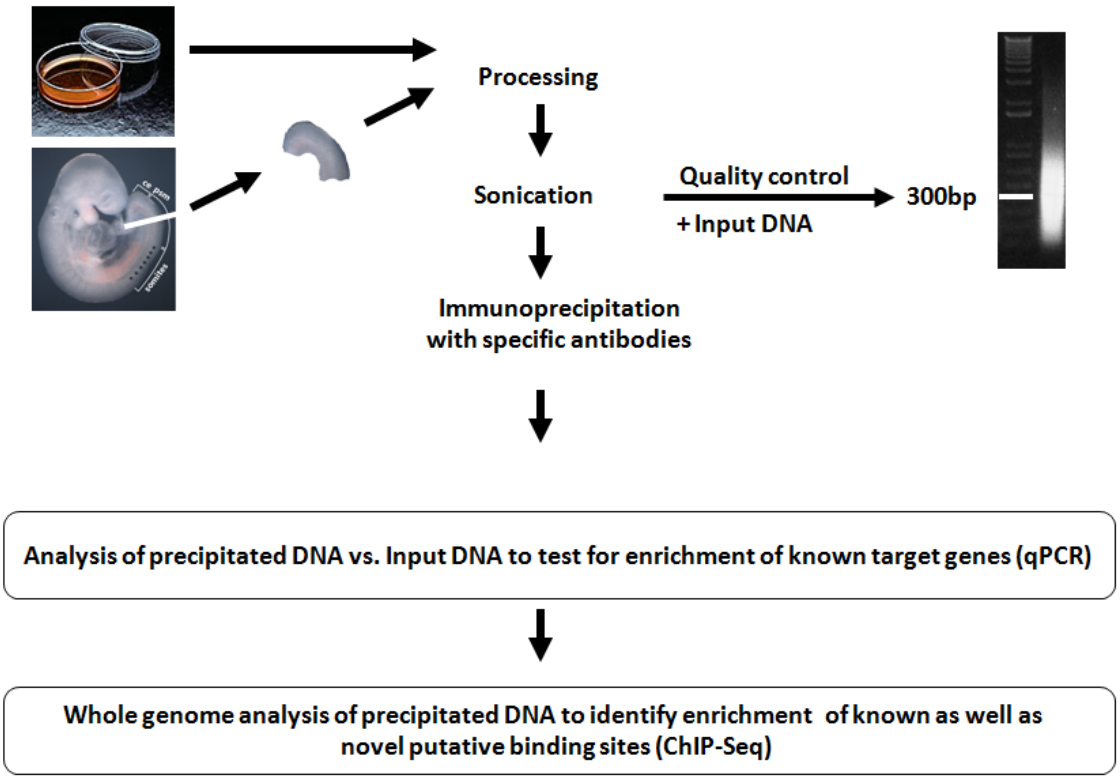

Figure 9: *ChIP experiments are performed using chromatin from P19 cells or from embryonic caudal ends. Caudal ends include the tailbud and the last third of the unsegmented paraxial mesoderm of E9.5 embryos (see white line in the embryo image). The samples are cross-linked, lysed, and the chromatin is fragmented by sonication into fragments with an average size of 300bp. A portion of the DNA is purified for quality control (QC) and as Input DNA, which is used for normalization in the qPCR analysis, while the rest is used for ChIP. ChIP-DNA is analyzed by qPCR or by ChIP-Seq to determine Srf binding sites.*

2.1.2. Srf antibody ChIP

Srf expression in the ps and in newly formed mesoderm has previously been reported for embryos at E7.5 and E8.5 (Barron et al. 2005). However, for ChIP experiments we aimed to use caudal ends from embryos at E9.5. At this stage, there is substantially more tissue, and therefore more chromatin for the ChIP experiments. Therefore, I initially analyzed whether *Srf* is still expressed in the ps and nascent mesoderm at this stage. Whole mount *in situ* hybridization with an *Srf*-specific probe revealed expression in the tailbud and the psm, and in early somites. Further domains of strong expression were seen in the branchial arches, the forebrain, limb bud mesenchyme and intermediate mesoderm (Fig.10). Hence, caudal ends from embryos at this stage should be suitable for identifying Srf targets during the process of axis extension.

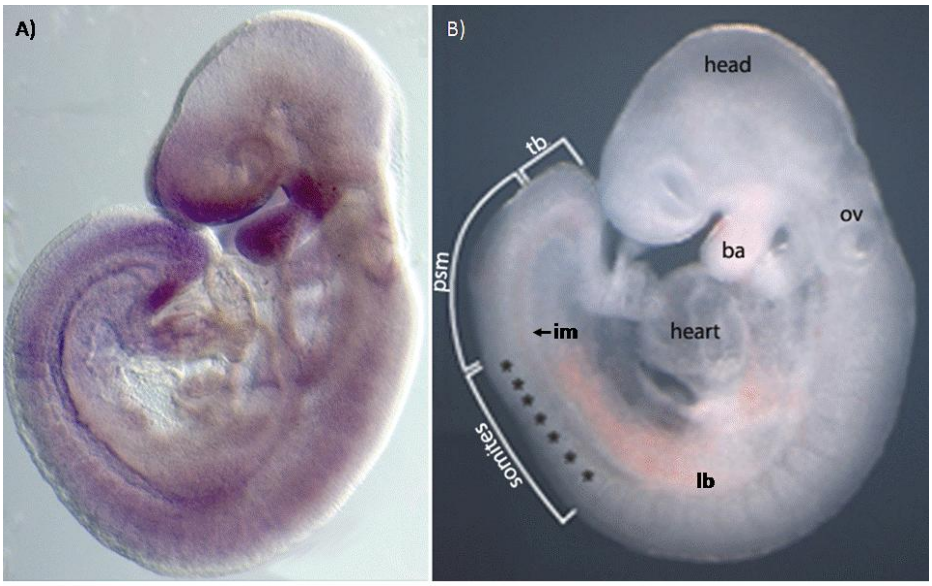

Figure 10: *A) Whole mount* in situ *hybridization of a mouse embryo at E9.5 using an* Srf*-specific probe, revealing* Srf *expression in the tailbud, presomitic mesoderm, early somites, branchial arches, forebrain, limb bud mesenchyme and intermediate mesoderm. B) Bright field image of a mouse embryo at E9.5 (ba: branchial arches, im: intermediate mesoderm, lb: limb bud mesenchyme, ov: otic vesicle, psm: presomitic mesoderm, tb: tailbud)*

2.1.2.1. *Establishment of an* in vitro *Srf ChIP assay*

Before performing ChIP with chromatin from embryonic *in vivo* tissue, I aimed to establish and optimize the Srf ChIP with *in vitro* material. We chose to use chromatin from embryonic carcinoma cells (P19 cells), because undifferentiated P19 cells had been shown to express *Srf* (Zhang et al. 2005), and because the developmental potential of P19 cells resembles that of early embryonic cells, allowing for the differentiation into all three germ layers, including mesoderm (Mcburney 1993). Therefore, they are similar to ES cells, which were already shown to require Srf activity (Schratt, Philippar, et al. 2002), however, they are much easier to cultivate in large quantities than ES cells. The Srf antibody that was used in the following experiments (commercially available from Santa Cruz) was previously reported to be suitable for ChIP experiments (Wycuff et al. 2004).

Following Srf ChIP with chromatin from P19 cells, the precipitated DNA was analyzed by qPCR for enrichment of the Srf target genes *Egr1* and *Fos* (Alexandre et al., 1991; Treisman, 1987). These sites are bound by Srf in most, if not all, cellular contexts and are therefore ideal positive controls for the qPCR analysis. Srf ChIP with chromatin from P19 cells resulted in more than 100 fold enrichment of

Egr1 and around tenfold enrichment of *Fos* (Fig.11). The differences in enrichment reflect distinct binding affinities of Srf to these targets, i.e. as a result of the number of Srf binding sites. *Egr1* has six Srf binding sites, whereas *Fos* has one Srf binding site, explaining that enrichment is higher for *Egr1* than the enrichment of most of the other Srf targets (Datta et al. 1992; Tullai et al. 2004).

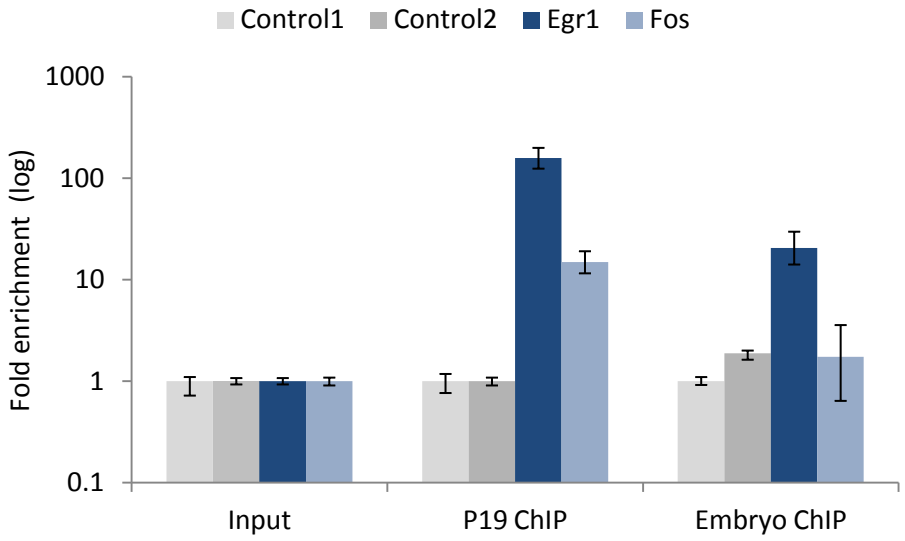

Figure 11: *qPCR analysis of Srf ChIP using chromatin from P19 cells and embryonic tailbuds. ChIP performed on P19 cells yielded 100 and 10 fold enrichment of the known Srf target genes* Egr1 *(red) and* Fos *(yellow), respectively, over background (grey), while ChIP on embryonic tissue resulted only in slight enrichment of* Egr1 *and no enrichment of* Fos*.*

2.1.2.2. *Srf ChIP using embryonic tissue resulted in insufficient enrichment*

After establishing the Srf ChIP with *in vitro* material, I performed the same ChIP protocol with chromatin from embryonic caudal ends. For a single ChIP experiment four to eight micrograms of chromatin, isolated from 20-40 caudal ends from embryos at E9.5, were used. This amount of chromatin was sufficient for Srf ChIP from P19 cells, however, the ChIP performance for the caudal ends was not as efficient as for the P19 cells. Maximal enrichment of the strong target *Egr1* was around 20 fold, while enrichment of the weak target *Fos* could not be detected (Fig.11). In general, this could be due to biological reasons, such as weaker binding of Srf to *Egr1* and *Fos* in the examined tissue, however, other Srf targets such as *Tpm1* were also not enriched (data not shown), indicating that Srf ChIP is not suitable for the identification of most Srf target genes in embryonic caudal ends. Poor performance of antibody ChIP on *in vivo* tissue was a problem also observed by others (P. Tam, T. Jenuwein, personal communication) and could be caused by a multitude of different factors.

Unlike the P19 cell sample, the embryonic tissue is heterogeneous. Cells in the caudal end are in a transient state, in which they are exposed to different extracellular signals. Srf binding to its target genes depends on the cellular context, i.e. the presence of specific cofactors, and cofactor activity and localization is triggered by extracellular signals. Hence, loss of the signal would result in detachment of Srf from its genomic target. Therefore, it is likely that not all cells of the caudal end contain Srf binding to a particular target, which would lead to a weaker enrichment of this target.

Taken together, the data from the Srf ChIP experiments from P19 cells show the enrichment of strong and weak targets, while the data from embryonic caudal ends show only low enrichment of strong Srf targets and no enrichment of weaker targets.

2.1.3. An alternative ChIP method: Avi[Bio] ChIP in vitro

Since ChIP with an Srf-specific antibody performed poorly using embryonic tissue, we aimed to establish a more robust method to identify Srf target genes. Generally, ChIP with an antibody against the transcription factor of interest can be replaced by precipitation of epitope tagged transcription factors. The transcription factor is fused with a tag for which specific high affinity antibodies are available. This approach has often been used to increase ChIP efficiency (Rigaut et al. 1999; reviewed in Terpe 2003), but it still involves the use of an antibody, which always means a limitation in binding affinity. In contrast, it was previously shown that biotin can be used as a high affinity tag for ChIP (Kolodziej et al. 2009), which avoids the usage of antibodies.

Unlike other epitope tags, biotin is not bound by an antibody but by avidin and streptavidin. The dissociation constant of the interaction of biotin with these molecules is 10^{15} $L*mol^{-1}$, which is 10^{3} to 10^{6} times stronger than that of epitopes with their specific antibodies. Consequently, the biotin-streptavidin complex is not disrupted by changes in pH, introduction of detergents, or high salt concentration and therefore remains stable even under very stringent washing conditions (Kolodziej et al. 2009). We reasoned that, due to the high biotin-streptavidin affinity, precipitation of biotinylated Srf with streptavidin coated beads should increase the enrichment of Srf targets from embryonic tissue.

In order to tag a protein with biotin, it is fused to a short 13 amino acid-long peptide sequence (SLNDIFEAQKIEW), called Avidin-tag (Avi-tag). This short peptide is the minimal recognition sequence

that is able to be biotinylated on lysine 10 by the *E. coli* Biotin ligase A (BirA) (Beckett et al. 1999). For biotin tagging, the ligase has to be coexpressed with the Avi-tagged protein of interest (Fig.12,A).

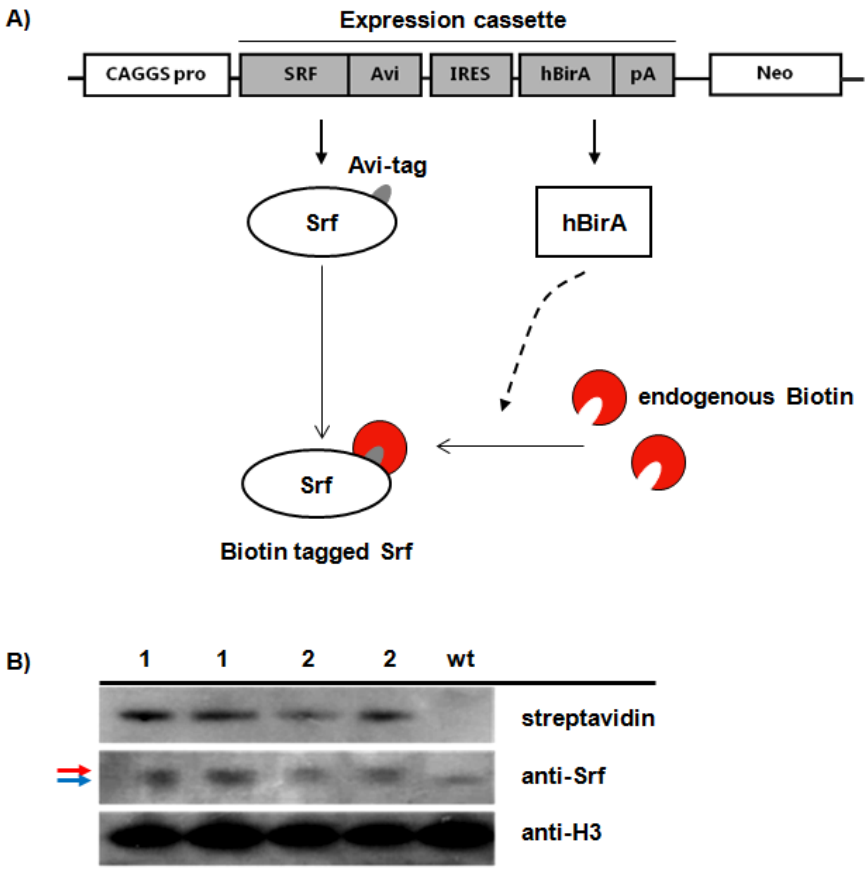

Figure 12: *Two stably transfected P19 cell lines express Srf-Avi[Bio]. A) Schematic representation of the expression cassette including a strong and constitutively active CAGGS promoter, the Srf-Avidin fusion sequence and the biotin ligase (hBirA) encoded by the same transcript. Translation of this transcript generates an Avi-tagged Srf protein as well as a humanized version of a bacterial Biotin Ligase (hBirA). The ligase catalyses the binding of endogenous biotin molecules, which are present in the cell, and thus generates a biotin-tagged version of Srf. B) Detection of endogenous Srf using an Srf-specific antibody and of Srf-Avi[Bio] using streptavidin in protein lysates from two P19 cell lines (both lines were tested in duplicates) and from wt P19 cells. Srf-Avi[Bio] can be detected in both transgenic P19 cell lines, while it is absent from wt P19 cells. Note that there are two bands detected with the anti-Srf antibody (indicated with red and blue arrow) in the lanes of the transgenic P19 cell lines 1 and 2, while only the lower band is present in the wt lane. The higher band represents the tagged Srf protein, which has a higher molecular weight. Histone 3 (anti-H3) was used for normalization. hBirA: humanized version of bacterial biotin ligase A, IRES: internal ribosomal entry site, Neo: Neomycin resistance gene, pA: poly-adenylation site, pro: promoter, wt: wild type*

2.1.3.1. *Srf-Avi[Bio] ChIP in vitro*

The general suitability of the biotin-streptavidin system for precipitation of the Srf-Avi[Bio]-DNA complex was tested *in vitro*. Again, we used P19 cells, since this allows the usage of the previously obtained Srf ChIP data for reference. We assembled an expression cassette which generates a transcript encoding Srf fused to the Avi peptide at its C-terminal end. In the same transcript, we integrated the cDNA of a codon optimized (human) version of the Biotin ligase A (hBirA), with efficient transcription enabled by an internal ribosomal entry site (IRES) (Fig.12,A).

To minimize enrichment variation depending on transfection efficiency, I generated P19 cell lines stably expressing Srf-Avi[Bio] under control of the constitutively active CAGGS promoter (Fig.12,A). The expression cassette was linked to a neomycin-resistance gene, to allow for selection of cells that had stably integrated the transgene. We used two of the stable transfected cell lines (termed P19 Srf#1 and P19 Srf#2) and verified the expression of Srf-Avi[Bio] by immunoblot analysis. While streptavidin-meditated detection of the biotin molecule results in a single band, the Srf antibody detects both the endogenous as well as the biotinylated Srf. Since the Srf-Avi[Bio] molecule has a higher molecular weight than the endogenous Srf, two bands could be distinguished (Fig.12,B). Next, the P19 Srf#1 cell line was applied for Srf-Avi[Bio] ChIP.

In a first step, I used high amounts of chromatin (10µg) from P19 Srf#1 cells and from wt P19 cells and performed Srf-Avi[Bio] ChIP as well as Srf ChIP, the latter as described in 2.1.2.2. The Srf-Avi[Bio] ChIP resulted in approximately 100 fold enrichment of *Egr1* and tenfold enrichment of *Fos*, but as expected, no enrichment was detected when chromatin from wt P19 cells was used (Fig.13). As seen before, Srf ChIP, which uses the Srf antibody, also results in high enrichment of *Egr1* and lower enrichment of *Fos.* Notably, the enrichment obtained by Srf ChIP was comparably high when performed with chromatin from wt P19 cells or Srf#1 P19 cells, showing that a higher abundance of Srf due to its CAGGS promoter driven exogenous expression does not result in higher enrichment.

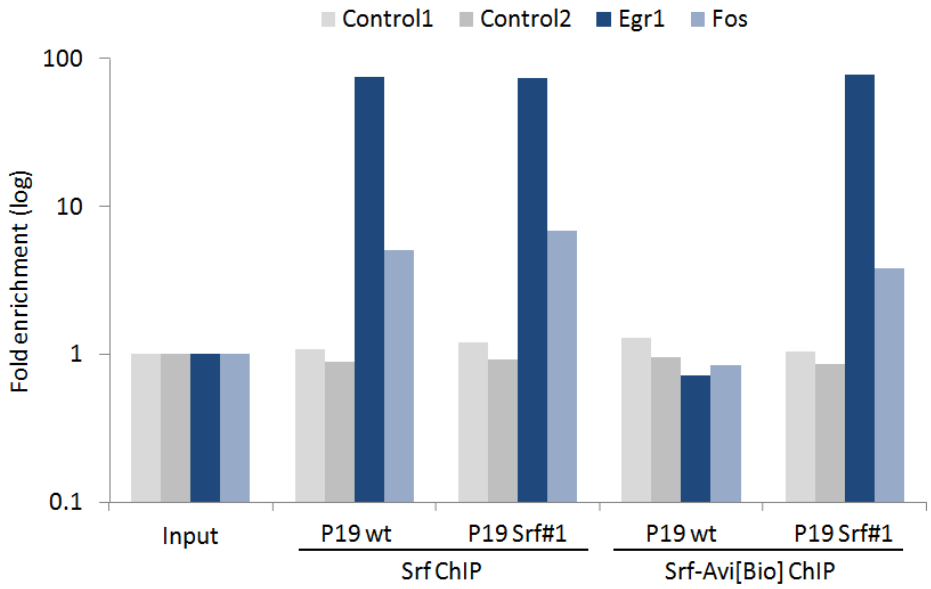

Figure 13: *The Biotin-Streptavidin system is suitable for enriching Srf target genes. Srf ChIP with high amounts of chromatin from P19 cells expressing the Srf-Avi[Bio] expression cassette (P19 Srf#1) or from wild type (wt) P19 cells result in comparable enrichment of the Srf target genes Egr1 and Fos. Similar enrichment is achieved with Srf-Avi[Bio] ChIP when performed on lysate from P19 Srf#1 cells but not with lysate from wild type P19 cells.*

Together, the results show that the biotin-streptavidin system is generally suitable for the identification of Srf target genes. Also, they validate the use of the expression cassette. However, since the enrichment with Srf-Avi[Bio] ChIP and Srf ChIP is similarly high, it does not show that the system can increase the ChIP efficiency.

The experiments that are shown above were performed with high amounts of chromatin. The reason why we used the Biotin-Streptavidin system was to increase ChIP enrichment from heterogeneous embryonic tissue. To challenge the Srf-Avi[Bio] ChIP performance in an *in vitro* situation, I gradually decreased the amount of chromatin used in the ChIP experiments.

Indeed, Srf ChIP and Srf-Avi[Bio] ChIP performed comparably when either 10µg or 2µg of chromatin were used, but with amounts as low as 0.4µg of chromatin, only Srf-Avi[Bio] ChIP resulted in enrichment of both *Egr1* and *Fos*, while Srf ChIP lost *Fos* enrichment (Fig.14,A and B). This indicates that Srf-Avi[Bio] ChIP is more efficient for the identification of weak Srf target genes than conventional Srf ChIP when working with low amounts of material.

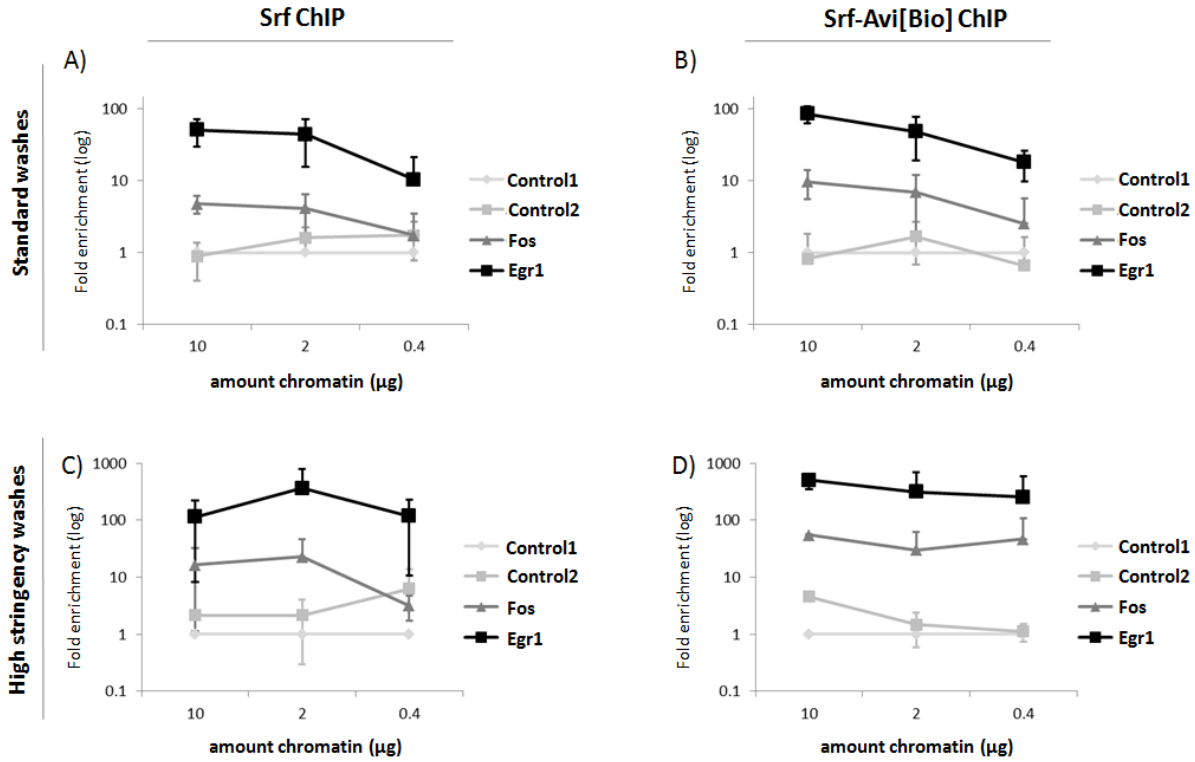

Figure 14: *Srf-Avi[Bio] ChIP performs better than Srf ChIP when low amounts of chromatin are used. Srf ChIP experiments (A and C) and Srf-Avi[Bio] ChIP experiments (B and D) were performed with decreasing amounts of chromatin from P19 cells (x-axis). High stringency washing increased Srf target enrichment in the Srf-Avi[Bio] ChIP (D). All ChIPs were performed twice, and error bars represent the deviation between both experiments.*

Generally, ChIP enrichment represents the ratio of DNA specifically bound to the transcription factor of interest versus nonspecifically bound DNA. Given the high biotin-streptavidin affinity, we reasoned that stringency and washing time could be increased to reduce the background without losing specifically bound DNA. In fact, serial tests of washings with different stringencies and washing times led to the establishment of a new protocol for Srf-Avi[Bio] ChIP (see Methods). It increases the enrichment of Srf target genes by a factor of five, regardless of the amount of chromatin used (Fig.14,C and D). Higher stringency washes also increased the enrichment of Srf ChIP when high amounts of chromatin were used, however, weaker Srf targets were still lost with lower amounts of chromatin (Fig.14,C).

To summarize, Srf ChIP and Srf-Avi[Bio] ChIP result in similar enrichment of the Srf target genes *Egr1* and *Fos* when performed with high amounts of chromatin, but Srf-Avi[Bio] ChIP performance is

superior when the amount of chromatin is decreased. This suggests that the high binding affinity of streptavidin to Srf-Bio should also increase the enrichment of Srf targets from embryonic tissue.

2.1.3.2. *Comparison of Srf and Srf-Avi[Bio] targets*

Before using the Biotin-Streptavidin system for *in vivo* experiments, we wanted to be sure that the method does not create any artifacts. In mammalian cells, there is a small number of endogenously naturally biotinylated proteins, including nuclear proteins such as splice factors (de Boer et al. 2003). The employment of Streptavidin for ChIP could therefore result in enrichment of DNA sequences that are bound by the endogenously biotinylated molecules but not by Srf. On the other hand, the biotinylation of Srf could inhibit binding events to specific DNA sequences, for example by blocking the interaction of Srf with certain cofactors. Both scenarios can be ruled out if Srf ChIP and Srf-Avi[Bio] ChIP lead to the identification of the same target genes.

To identify Srf and Srf-Avi[Bio] targets on a genome-wide level, we performed ChIP-Seq experiments according to the Illumina protocol. Briefly, the precipitated DNA fragments were attached to specific adapter sequences which enabled binding to the surface of a flow cell. Here, the fragments were amplified, resulting in clusters of identical DNA sequences. In the next step, all fragments are sequenced simultaneously using a genome sequencer (Invitrogen), yielding approximately 10-20 million short reads per run. The sequences are then mapped against the mouse genome, identifying those sequences that were enriched by the ChIP. Bioinformatic tools can be used to identify those regions and to visualize them as peaks.

For massive parallel sequencing, several ChIP experiments were performed and the precipitated DNA was pooled, resulting in 8ng of Srf ChIP DNA and 9ng of Srf-Avi[Bio] ChIP DNA. The minimal *Egr1*-enrichment in each experiment was 100 fold as quantified by qPCR. In total, 4288 and 2498 enriched sequences were identified by sequencing the Srf ChIP-DNA and Srf-Avi[Bio] ChIP-DNA, respectively (fold enrichment ≥ 2.5; p-value ≤ $1x10^{-5}$). Srf peaks were visualized using the Overview software, which was developed in our lab (M. Werber). As an example, the genomic loci of *Egr1* and *Actb* and the corresponding Srf- and Srf-Avi[Bio] peaks are depicted in Figure 15. The peaks in both loci are very similar showing that, in principal, both ChIP-Seq experiments were performed

successfully.

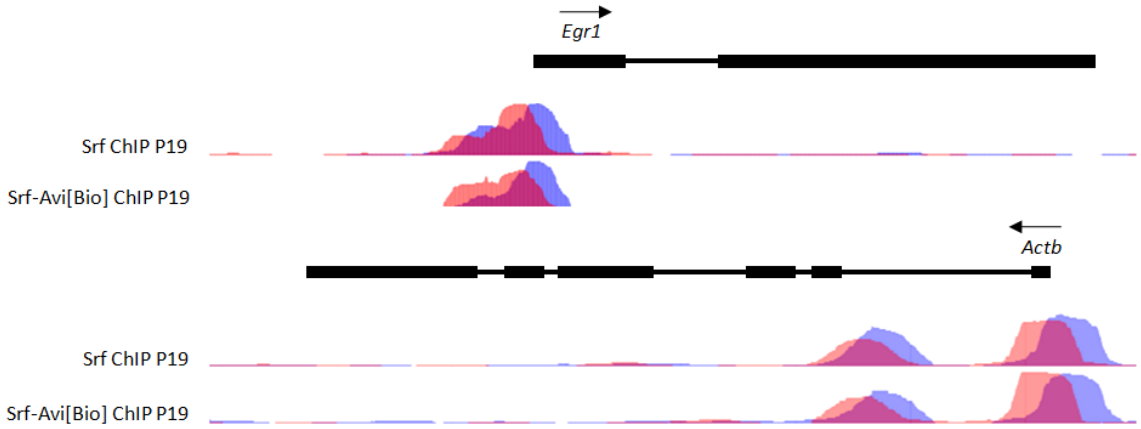

Figure 15: *Linear schematic of the genomic loci of* Egr1 *and* Actb*. Exons are depicted as black boxes, introns as black lines. Orientations of the genes are indicated by arrows. Peaks generated by Srf ChIP and Srf-Avi[Bio] ChIP are depicted in blue (+strand) and red (- strand). Srf and Srf-Avi[Bio] bind close to the transcriptional start site (TSS) of both genes, and there is a second binding site in intron 1 of the Actb gene. However, no significant differences can be seen between the peaks detected by the two methods.*

As previously mentioned, Srf specifically binds to CArG boxes in the regulatory regions of its target genes. Hence, *de novo* sequence analysis can be used to examine the quality of the Srf ChIP and Srf-Avi[Bio] ChIP data. Motif analysis for all enriched sequences was performed using the Meme software (http://meme.sdsc.edu/meme/intro.html). In both data sets, the most abundant motif was very similar to the known CArG box (Fig.16), demonstrating that Srf target genes are able to be identified on a genome-wide level using both methods.

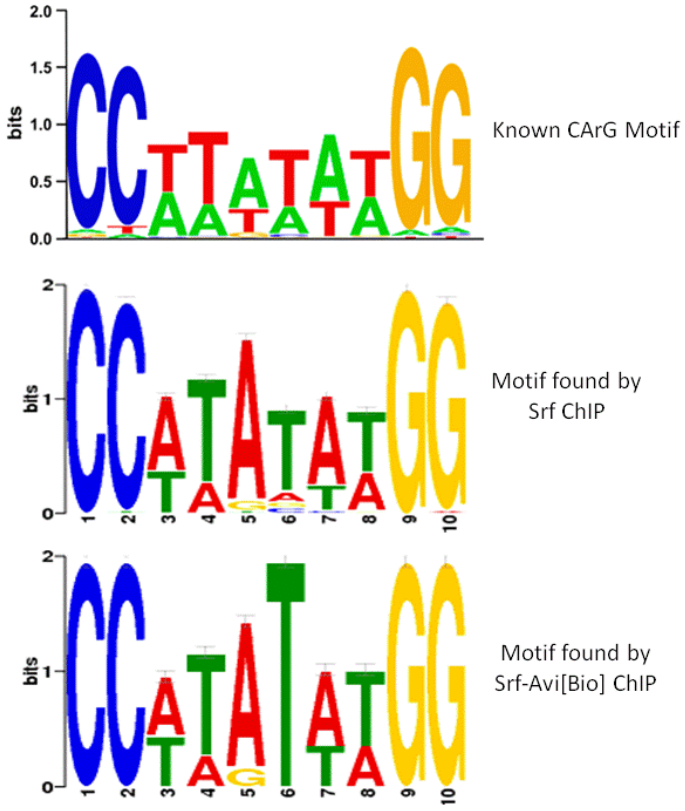

Figure 16: *Sequence similarity of the known CArG box (Miano, 2010) and the top hits found in a* de novo *sequence analysis of Srf ChIP-Seq data and Srf-Avi[Bio] ChIP-Seq data.*

To compare all peaks identified with the Srf ChIP and the Srf-Avi[Bio] ChIP, each peak was linked to the gene in its closest proximity. It is important to note that the peaks do not necessarily belong to this gene, since there was no limitation in distance to the transcriptional start site (TSS). The linkage was used instead to technically define the peaks for easier comparison of the data sets. For the subsequent biological analysis, I used only peaks with a maximal distance of 4kb to the TSS of a gene (see 2.1.3.3.).

In total, 3425 and 2095 genes were linked to one or more peaks identified by Srf ChIP and Srf-Avi[Bio] ChIP, respectively. Comparison of these genes revealed a relatively low overlap (823 genes, representing 24.02% of Srf targets and 39.28% of Srf-Avi[Bio] targets) (Fig.17,A), which would question the reliability of the data or the suitability of the Biotin-Streptavidin system. However, this comparison used peaks that were identified with a relatively low threshold (fold enrichment ≥ 2.5; p-value ≤ $1x10^{-5}$), and therefore, they are likely to contain many false positives, which could explain the low overlap.

In order to reduce the likelihood of false positives, I increased the stringency by which the enriched DNA sequences are defined as peaks (fold enrichment ≥ 10; p-value ≤ $1x10^{-10}$). This step drastically reduced the number of identified targets to 183 Srf targets (identified by using the antibody) and 120 Srf-Avi[Bio] targets (appendix, tables 10 and 11). However, it also increased the overlap significantly to 54.6% of Srf targets and 83.3% of Srf-Avi[Bio] targets (100 genes) (Fig.17,B).

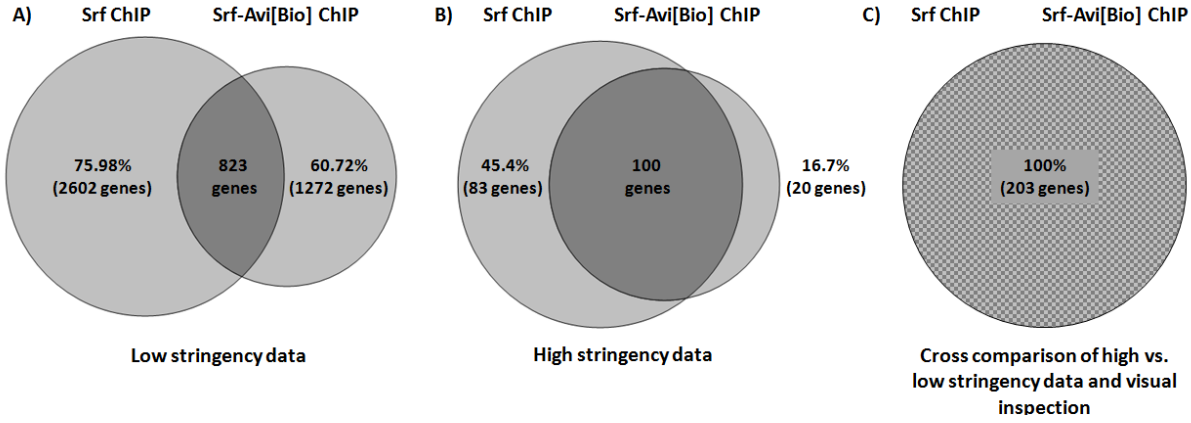

Figure 17: *Comparison of Srf targets identified by Srf-Avi[Bio] ChIP or Srf ChIP reveals that the Srf-Avi[Bio] ChIP can be used for reliable identification of Srf targets. A) Peak identification with low stringency (fold enrichment ≥ 2.5; p-value ≤ $1x10^{-5}$) results in a high number of Srf targets, but the overlap between the targets identified by Srf ChIP and by Srf-Avi[Bio] ChIP is low. B) The low overlap is probably caused by a high number of false positives. Peak identification with high stringency (fold enrichment>10; p-value<$1x10^{10}$) drastically decreased the number of Srf targets, many of which were likely to be false positives. The overlap of high stringency Srf targets and Srf-Avi[Bio] targets is amplified. C) Visual inspection of the peaks that seem to be unique for the Srf-Avi[Bio] ChIP or Srf ChIP method after the high stringency thresholds were applied reveals that they are actually enriched by both methods, but in one method the enrichment is below the detection level.*

This indicates that the low overlap seen before is indeed caused by a high number of false positives, which can be drastically reduced by integration of a more stringent threshold for peak calling. However, there are still 83 and 20 genes that are only identified as Srf or Srf-Avi[Bio] targets, respectively, even after the more stringent threshold was applied. To analyze whether this is owing to the presence of the biotin tag, I wanted to know whether these peaks were really specific to the Srf ChIP or the Srf-Avi[Bio] ChIP method, or if these DNA sequences are bound by Srf and by Srf-Avi[Bio], but filtered out in one of the data sets due to the high stringency threshold.

Indeed, from the 83 genes that are identified as Srf targets but not Srf-Avi[Bio] targets when the high stringency threshold was applied, 60 can be identified also as Srf-Avi[Bio] targets when the

lower stringency threshold is used (data not shown). And *vice versa*, from the 20 genes that were identified as Srf-Avi[Bio] targets but not as Srf targets with high stringency, 15 were also identified as Srf targets with low stringency thresholds (data not shown), demonstrating that only very few genes (23 for Srf ChIP and five for Srf-Avi[Bio] ChIP) are targets only identified by one of the two methods. Moreover, focusing on the overview-visualization of these peaks reveals that they are actually enriched by both methods, but in one method the enrichment is below the detection level of the peak-finding software (Fig.17,C).

Together, the detailed comparison of Srf targets and Srf-Avi[Bio] targets reveals that Srf-Avi[Bio] ChIP can be used for reliable identification of Srf targets. The Biotin-Streptavidin system does not create Srf-independent enrichment of DNA sequences caused by endogenously biotinylated proteins, nor does it inhibit the binding to a subset of Srf target genes, at least in P19 cells. However, this analysis also showed that the usage of a low stringency threshold for peak identification results in many false positives. The number of false positives can be drastically decreased by employment of a higher threshold, therein also reducing the number of identified targets.

2.1.3.3. *Identification of novel Srf target genes in P19 cells*

Application of a high stringency filter for peak identification from the ChIP-Seq data results in a relatively low number of identified peaks, most likely less than actual peaks present. However, it also reduces the number of false positives and results in a high overlap between Srf targets and Srf-Avi[Bio] targets (Fig.17,B). Thus, peaks that are identified by both methods are highly reliable and can be used to identify novel Srf target genes.

The 100 genes that are linked to peaks identified by both methods when applying the high stringency threshold (Fig.17,B) were further examined. In the previous analysis, the peaks were linked to the gene in closest proximity, regardless of the distance to the TSS. To improve the chance that the peaks are really biologically related to the gene they are linked to, I analyzed those 47 of the 100 genes that have a peak within 4kb to the TSS. This distance was based on observations by Sun and colleagues, who estimated that nearly all Srf binding sites are within 4kb from the TSS of a gene (Sun et al. 2006).

Of the 48 genes analyzed, 24 were previously described as Srf target genes in the literature. Consequently, the remaining 24 were regarded as novel putative Srf target genes (Table 1). Among

these genes were actin binding proteins like *Arc* (activity regulated cytoskeletal-associated protein), and *Dixdc1* (DIX domain containing 1), which is also implicated in canonical Wnt signaling, the putative transcription factor *Zswim6* (zinc finger, SWIM domain containing 6), the micro RNA *mmu-mir-143* and the apoptosis associated gene *Bcl2l12* (B-cell leukemia/lymphoma 2 like 12). *Zswim6* as well as other novel putative Srf target genes are described in more detail in Chapter 2.2.5.

Table 1: *Srf target genes bound by Srf in P19 cells. Targets were identified by Srf ChIP-Seq as well as by Srf-Avi[Bio] ChIP-Seq when applying a high stringency threshold (fold enrichment ≥ 10; p-value ≤ 1x10^{-10}). The distance between the Srf binding site and the transcriptional start site of the closest gene is indicated*

	Gene	***Gene ID***	***Distance (bp)***	***GO term***
Known Srf targets	*Actb*	*ENSMUSG00000029580*	*330*	*cytoskeleton*
	Cfl1	*ENSMUSG00000056201*	*1568*	*regulation of actin*
	Cnn1	*ENSMUSG00000001349*	*465*	*actin cytoskeleton*
	Cnn2	*ENSMUSG00000004665*	*380*	*actin cytoskeleton*
	Cyr61	*ENSMUSG00000028195*	*2970*	*regulation of cell adhesion*
	Dstn	*ENSMUSG00000015932*	*435*	*actin cytoskeleton*
	Dusp5	*ENSMUSG00000034765*	*271*	*MAP kinase phosphatase*
	Egr1	*ENSMUSG00000038418*	*1000*	*transcription factor activity*
	Egr2	*ENSMUSG00000037868*	*812*	*transcription factor activity*
	Egr3	*ENSMUSG00000033730*	*2874*	*transcription factor activity*
	Egr4	*ENSMUSG00000071341*	*322*	*transcription factor activity*
	Fhl2	*ENSMUSG00000008136*	*243*	*transcription repressor activity*
	Fosb	*ENSMUSG00000003545*	*40*	*transcription factor activity*
	Fosl1	*ENSMUSG00000024912*	*433*	*transcription factor activity*
	Fosl2	*ENSMUSG00000029135*	*844*	*transcription factor activity*
	Ier2	*ENSMUSG00000053560*	*712*	*transcription factor activity*
	Junb	*ENSMUSG00000052837*	*2321*	*transcription factor activity*
	Map3k14	*ENSMUSG00000020941*	*1311*	*protein kinase activity*
	Mcl1	*ENSMUSG00000038612*	*558*	*BCL2-like apoptosis inhibitor*
	mmu-mir-143	*ENSMUSG00000065445*	*2703*	*Smooth muscle cell*
	Smtn	*ENSMUSG00000020439*	*1930*	*Calponin-like actin-binding*
	Srf	*ENSMUSG00000015605*	*364*	*transcription factor activity*
	Vcl	*ENSMUSG00000021823*	*607*	*cell adhesion*
	Wdr1	*ENSMUSG00000005103*	*268*	*actin cytoskeleton*
Novel putative Srf targets	*Zswim6*	*ENSMUSG00000032846*	*177*	*zinc ion binding, transcription*
	Mpdu1	*ENSMUSG00000018761*	*502*	*oligosaccharide biosynthetic*
	Aida	*REF00000124*	*1910*	*dorsal/ventral pattern*
	Bcl2l12	*ENSMUSG00000003190*	*421*	*cell survival*
	Ptms	*ENSMUSG00000030122*	*1777*	*zinc ion binding*
	Sertad1	*ENSMUSG00000008384*	*3532*	*transcription factor*
	Eif3h	*ENSMUSG00000022312*	*601*	*translation initiation factor*
	Homer1	*ENSMUSG00000007617*	*293*	*skeletal muscle contraction*
	Dixdc1	*ENSMUSG00000032064*	*875*	*Actin binding, Wnt receptor*
	Vmn2r29	*ENSMUSG00000074394*	*714*	*G-protein coupled receptor*
	Arc	*ENSMUSG00000022602*	*631*	*Actin binding, cell migration*
	Gga1	*ENSMUSG00000033128*	*598*	*intracellular protein transport*
	Plekho2	*ENSMUSG00000074239*	*1140*	*Unknown*
	Hnmt	*ENSMUSG00000026986*	*3050*	*N-methyltransferase activity*

Gene	*Gene ID*	*Distance (bp)*	*GO term*
Nr2f2	*ENSMUSG00000030551*	*980*	*ligand-regulated transcription*
Fam100a	*ENSMUSG00000039568*	*527*	*Unknown*
3110003A17RIK	*ENSMUSG00000078453*	*1251*	*Unknown*
AC069469.5	*ENSMUSG00000081564*	*2632*	*Not found in database*
AC175246.2	*ENSMUSG00000026361*	*337*	*Not found in database*
AC131780.5-202	*ENSMUSG00000074558*	*1564*	*Not found in database*
AC165092.3	*ENSMUSG00000076291*	*383*	*Not found in database*
AC131780.5-209	*ENSMUSG00000074566*	*1207*	*Not found in database*
AL645600.11	*ENSMUSG00000040838*	*2961*	*Not found in database*
AL732612.28	*ENSMUSG00000081392*	*3519*	*Not found in database*

2.1.4. Srf-Avi[Bio] ChIP in vivo

2.1.4.1. *Generation of transgenic ES cells and mouse embryos for* in vivo *Srf-Avi[Bio] ChIP*

The suitability of the Streptavidin-Biotin to identify Srf targets as well as the integrity of the expression cassette were demonstrated in the experiments shown above. Therefore, the biotin system could be transferred *in vivo* to identify Srf targets during the process of axis extension. Instead of using the constitutively active CAGGS promoter, which was driving the Srf-Avi[Bio] expression in P19 cells, I chose the tissue-specific T_{streak} promoter to drive the expression specifically in the caudal embryo. This would allow for the identification of Srf target genes specifically in the ps and nascent mesoderm. Additionally, I chose to express Srf-Avi[Bio] in a Doxycycline-inducible fashion for time-specific induction, since I did not know whether the exogenous expression of a functional *Srf* gene during all stages of development would harm the normal developmental process.

The *Srf-Avi[Bio]* expression cassette was transferred into an exchange vector system which was established in our laboratory. First, it was cloned into a donor vector which contains an optimized *tetS* binding site (TRE-tight, Clontech) for Doxycycline inducibility, the T_{streak} promoter and the *Srf-Avi[Bio]* expression cassette. This is followed by the rabbit *β-globin* polyadenylation signal, the chicken *β-globin* core insulator, and a *PGK* promoter. The DNA fragment carrying these elements is flanked by *loxP* and *lox5171* sites for site-specific recombination and depicted in Figure 18,A.

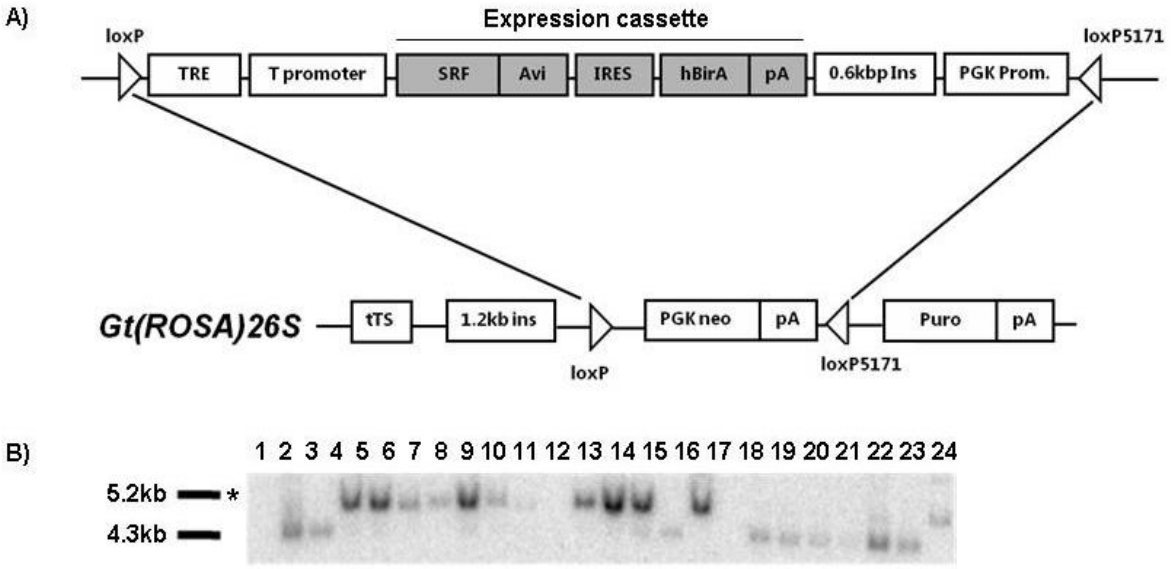

Figure 18: *Generation of transgenic ES cells carrying the Srf-Avi[Bio] expression cassette. A) The expression cassette including the tetracycline responsive element (TRE), the tissue-specific* T_{streak} *promoter, the* Srf-Avidin *fusion sequence and the* biotin ligase (hBirA) *encoded in the same transcript. Upon integration into the* Gt(ROSA)26S *locus, the expression cassette is flanked by a 1.2kb insulator on the 5'side and a 0.6kb insulator on the 3' side. B) Southern blot analysis. The blot depicts a typical recombination event. Following puromycin selection, genomic DNA is isolated from ES cell clones, digested with* HindIII *and labeled with a probe binding to the Puromycin resistance gene. Appearance of the larger 5.2kb band (*) indicates correct integration of the expression construct into the* Gt(ROSA)26S *locus, while the 4.3kb band shows digestion of the unchanged recipient transgene.*

Next, recombinase mediated cassette exchange was used to transfer the whole fragment into a recipient transgene, located in the *Gt(ROSA)26S* locus of the F1G4 ES cell line (Vidigal et al. 2010). The *Gt(ROSA)26S* locus is transcriptionally active ubiquitously throughout mouse embryogenesis, so that expression of genes integrated into the locus is not silenced by chromatin configurations (Zambrowicz et al. 1997). The recipient transgene contains the following elements (5' to 3'): a SV40 splice acceptor fused to codon optimized (mouse) version of *TetS* (Clontech) with an SV40 polyadenylation (pA) signal, activated by the endogenous *ROSA26* promoter; the 1.2-kb chicken *β-globin* insulator (*5'HS4*), a *PGK-neomycin* selection cassette flanked by opposing *loxP* and *lox5171* sites, and a promoter-less puromycin resistance gene with bidirectional polyadenylation signal. The splice acceptor, the pA signals and the insulator function in inhibition of endogenous promoter activity in the *Gt(ROSA)26S* locus to achieve T_{streak}-specific transgene expression. Upon integration, the *PGK* promoter of the exchange vector complements the promoter-less *puromycin* resistance

gene of the recipient locus to allow for puromycin selection. Additionally, proper integration of the transgene was confirmed by Southern blot (Fig.18,B).

Positively tested clones were assayed in an *in vitro* differentiation ChIP experiment. Aggregates of transgenic ES cells are cultivated in medium that contains recombinant bone morphogenetic protein 4 (BMP4) (10ng/ml), which induces differentiation into the mesodermal lineage and activation of the T_{streak} promoter (Nostro et al. 2008). After three days of differentiation, the cells express the mesodermal marker gene *T,* and were used for Srf-Avi[Bio] ChIP (Fig.19,A). The ChIP resulted in enrichment of the Srf target genes *Egr1, Actb, Tpm4, Fos* and *Fhl2* (Fig.19,B), which indicates proper functionality of the expression cassette under control of the T_{streak} promoter.

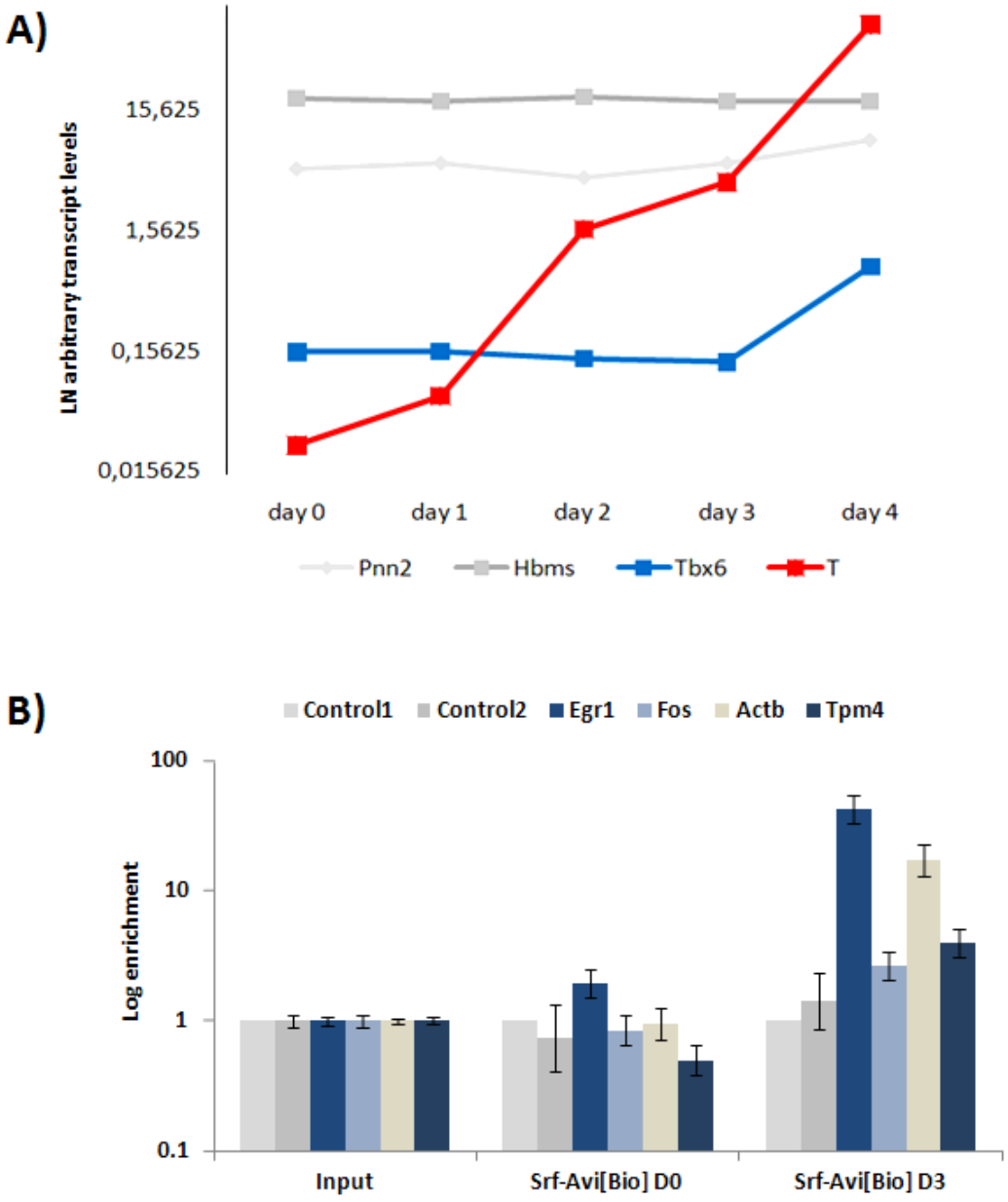

***Figure 19**: Srf-Avi[Bio] ChIP from transgenic ES cells that were differentiated into the mesodermal lineage. A) ES cell aggregates from transgenic ES cells activate the expression of* T *and* Tbx6*, which are marker genes for the mesodermal lineage. B) Since the* T *promoter also drives the transgene expression, enrichment of the Srf target genes* Egr1, Fos, Actb *and* Tpm4 *was detected after three days of differentiation (Day 3, right) but not at Day 0.*

In the next step, a positively tested ES cell line (*T_{streak} Srf-Avi[Bio] #8*) was used to generate fully ES-cell-derived embryos using the method of tetraploid complementation (Eakin & Hadjantonakis 2006). Head and tail samples were analyzed by western blot analysis for the expression of the

biotinylated version of Srf (Fig.20,A). Only embryos derived from foster mice that were fed with doxycycline during pregnancy expressed Srf-Bio, proving the functionality of the Tet responsive element.

The concentration of Srf-Avi[Bio] was much higher in the tail of transgenic embryos than in the head (Fig.20,A). The high expression in the tail resembles the T_{streak} promoter activity, however, the detection of Srf-Avi[Bio] in the head sample indicates additional endogenous promoter activity from the *Gt(ROSA)26S* locus. Moreover, whole mount *in situ* hybridization with a *hBirA*-specific probe exposed rather ubiquitous expression of the transgene (Fig.20,B), which is in disagreement with the western blot results. One reason for the unspecific expression pattern could be the generation of a transcript which is not translated into the Srf-Avi[Bio] protein, from endogenous activity of the *Gt(ROSA)26S* locus. It is conceivable that in addition to its tissue-specific expression, the transgene is expressed together with an endogenous transcript, for example due to incorrect splicing. This would result in a transcript which would also be detected by the *in situ* probe, but might not result in proper translation into the protein, e. g. due to a frame shift or incorrect splicing.

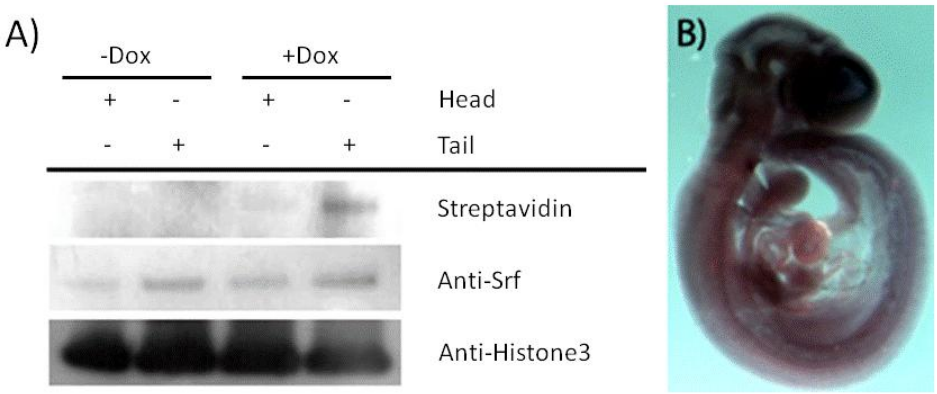

Figure 20: *Generation of transgenic embryos expressing Srf-Avi[Bio]. A) Head and tail samples from Dox induced and non-induced embryos were analyzed by western blot. Streptavidin detects the biotinylated version of Srf, while the anti-Srf antibody detects both the biotinylated and the endogenous version. Histone3 (anti-H3) is used as loading control. B)* In situ *hybridization of transgenic mouse embryos with an* hBirA-*specific probe reveals that expression of the transgene is not restricted to the activity domain of the* T_{streak} *promoter, but instead is ubiquitously detected.*

2.1.4.2. *Srf-Avi[Bio] ChIP improves enrichment of Srf targets from embryonic samples*

Western blot analysis showed strong *Srf-Avi[Bio]* expression in tails of Dox-induced transgenic T_{streak}*Srf-Avi[Bio]*#8 embryos. Thus, caudal ends were dissected at E9.5 as shown in Figure 9 and used for an Srf-Avi[Bio] ChIP experiment. 13 tailbuds were collected, processed as described above and the ChIP protocol that was established with chromatin from P19 cells was applied. For comparison, I also performed Srf-ChIP, using the Srf antibody as described in 2.1.2.

We applied the streptavidin-biotin system with the aim to increase the enrichment of Srf target genes from embryonic samples. Indeed, Srf-Avi[Bio] ChIP enrichment was 40 fold with standard washing and up to 120 fold when stringency and time of washings were increased (Fig.21). The increase of enrichment is not caused by the exogenous expression of *Srf* in T_{streak} *Srf-Avi[Bio]*#8 embryos, because this would also cause an increase of *Egr1* enrichment obtained by Srf ChIP, which was performed with lysate from those embryos. However, as depicted in Figure 21,A, Srf ChIP performance does not differ between lysate derived from transgenic or wild-type caudal ends. Moreover, enrichment of the weaker Srf target gene *Fos* is increased up to tenfold with Srf-Avi[Bio] ChIP, while it is not enriched with Srf ChIP (Fig.21), showing that the Streptavidin-Biotin system significantly improved Srf target gene enrichment from embryonic tissue.

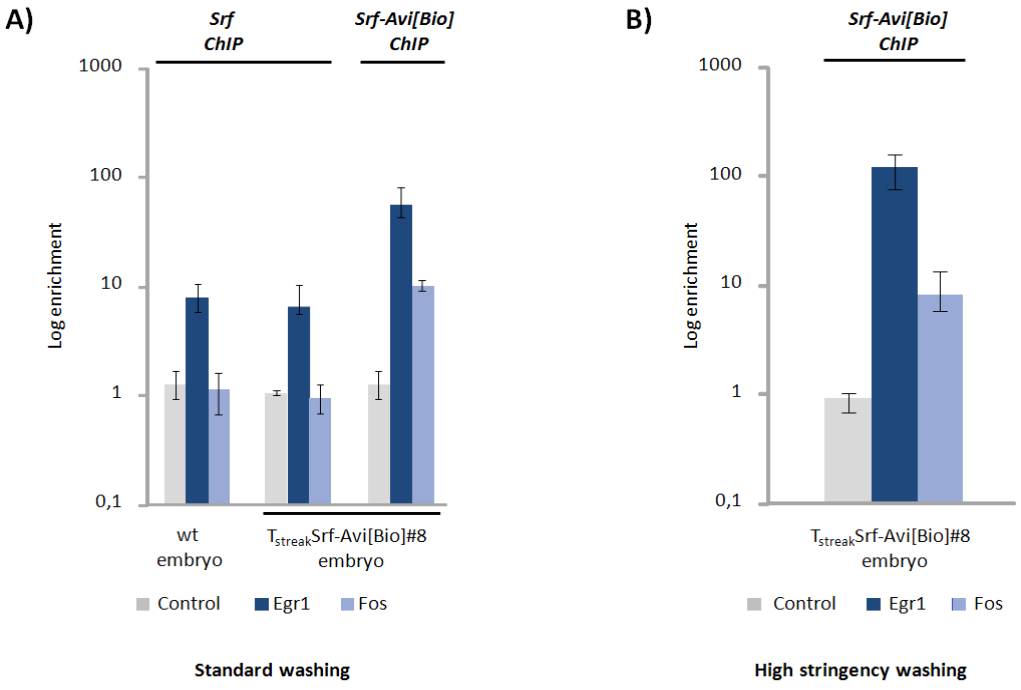

Figure 21: *Srf-Avi[Bio] ChIP increases enrichment of Srf targets from embryonic tissue. A) qPCR analysis of Srf ChIP and Srf-Avi[Bio] ChIP performed with caudal end lysate from wt embryos or Dox-*

treated T_{streak}Srf-Avi[Bio] #8 *embryos; standard washing procedure. The exogenous expression of Srf-Avi[Bio] does not result in an increase of Srf target gene enrichment in the Srf ChIP, but enrichment of* Egr1 *and* Fos *is significantly higher in Srf-Avi[Bio] ChIP. B) High stringency washing increases enrichment of targets genes form* T_{streak}Srf-Avi[Bio] #8 *embryo caudal ends.*

2.1.4.3. *ChIP-Seq analysis of embryonic Srf-Avi[Bio] ChIP*

Employment of the Biotin-Streptavidin system for Srf-Avi[Bio] ChIP increased the enrichment of Srf target sites from *in vivo* tissue. Next, I wanted to use this method to identify Srf targets in embryo caudal ends on a genome-wide level. To obtain enough material for ChIP-Seq, DNA from several ChIP experiments was pooled, yielding 7.3ng of DNA, slightly less than was used for P19 Srf-Avi[Bio] ChIP-Seq and Srf ChIP-Seq experiments (8 and 9 ng, respectively.). *Egr1* enrichment in the *in vivo* experiments varied between 40 and 120 fold, which means that the average enrichment of *Egr1* was lower than in the P19 samples used for ChIP-Seq, where *Egr1* enrichment was always higher than 100 fold.

Sequencing and bioinformatic analysis were performed as described earlier (2.1.3.2.) and identified 8,648 peaks, when a low stringency threshold was used for peak detection (fold enrichment ≥ 2.5; p-value ≤ $1x10^{-5}$). Again, the peaks were linked to the gene in closest proximity, revealing 5,884 genes corresponding to one or several peaks. Among the identified genes were the known Srf target genes *Egr1* and *Actb*, and the overview visualization of their genomic loci is depicted in Figure 22,A. The analysis of the P19 ChIP-Seq data showed that the employment of a higher stringency threshold for peak detection (fold enrichment ≥ 10; p-value ≤ $1x10^{-10}$) reduced the number of identified targets, but also the number of false positives. However, integrating the high stringency threshold for the P19 data had a more significant impact than for the embryo data. In detail, it reduced the number of Srf and Srf-Avi[Bio] targets in P19 cells to 183 genes (5.3%) and 120 genes (5.6%) of those identified with the low stringency thresholds. When the same high stringency threshold was used to filter the embryo ChIP-Seq data, the number of Srf-Avi[Bio] targets was only reduced to 28.7% (1,692 genes) of the targets identified with low stringency thresholds. In principal, this could be the result of some biological factors, i.e. Srf occupies many more targets in the caudal end than it does in P19 cells, but a detailed analysis of the peaks indicates that the embryo ChIP-Seq data comprises a high amount of false positives.

It was recently estimated that most of the Srf binding sites are located within 4kb distance to the TSS of a gene (Sun et al. 2006). Unlike Srf binding sites, false positives should be equally distributed

throughout the genome. Based on these assumptions, the ratio of peaks that are within 4kb distance to a TSS versus those that are more distant provides information about the reliability of the data. Strikingly, 39.3% of Srf-Avi[Bio] peaks and 37.1% of Srf peaks in P19 cells, but only 12.2 % of Srf-Avi[Bio] peaks from embryonic caudal ends are located within 4kb distance to the TSS (Fig.22,B).

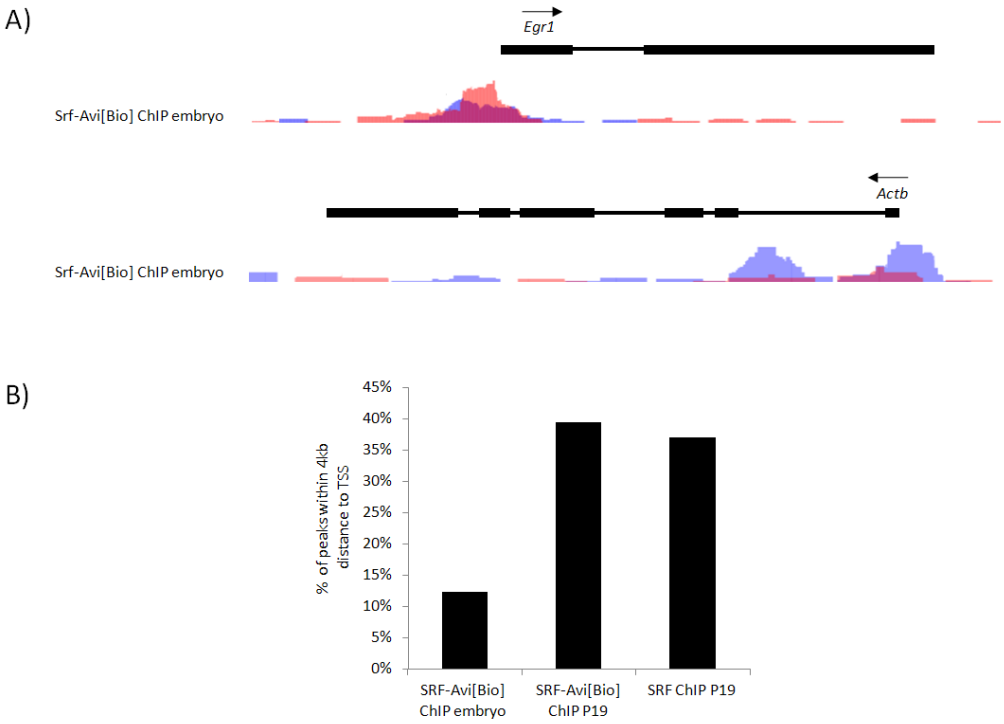

Figure 22: *A) Scheme of the genomic loci of Egr1 and Actb, Exons are depicted as black boxes, introns as black lines. Orientations of the genes are indicated (arrows). Peaks generated by embryo Srf-Avi[Bio] ChIP are depicted in blue (+strand) and red (- strand). B) Srf is known to bind primarily at sites proximal to the TSS of a gene. Hence, assessment of the percentage of peaks that are within 4kb to the TSS of a gene can be used to evaluate the quality of the ChIP-Seq data. Only 12.2% of the peaks identified from Srf-Avi[Bio] ChIP with embryonic tissue are found within 4kb of a gene TSS, indicating that this data set may contain a higher number of false positive peaks than the data obtained from the Srf-Avi[Bio] ChIP and Srf ChIP from P19 cells, where the ratio of proximal peaks is around three times higher.*

In general, these numbers are lower than expected when considering the estimation of Sun and colleagues (Sun et al. 2006). In part, the high number of peaks that are not within 4kb distance to a TSS in both P19 cells and embryonic caudal ends can be explained by the existence of genes that are not yet annotated in the mouse genome and in fact, there are Srf binding sites close to unknown transcripts (see Fig.23 and 30). However, the percentage of genes within 4kb distance is much lower for the embryo ChIP data than for the P19 data, indicating that the embryo data may contain a high

number of false positives. This assumption is supported by the fact that several peaks identified by embryo ChIP-Seq could not be confirmed by qPCR (data not shown). Also, different from the analysis of the P19 ChIP data (Fig.16), *de novo* sequence analysis with the Meme software did not identify the CArG sequence from the Srf-Avi[Bio] ChIP-Seq data. As mentioned before, the minimal *Egr1* enrichment in the P19 samples was 100 fold, while the *Egr1* enrichment in the embryo samples varied between 40 and 120 fold. The resulting higher background in the embryo data may be the reason for the presumed high ratio of false positives, and this might be caused for instance by incorrect mapping of low complexity DNA such as repeat regions.

Among the 1,692 genes identified as putative Srf-Avi[Bio] targets in embryonic tissue were also known Srf target genes like *Egr1, Actb, Junb, Map3k14,* and *Fhl1*. Furthermore, genes that are novel putative Srf target genes identified by P19 Srf ChIP are found as targets in the embryo, such as *Zswim6, Chl1*, the microRNA *AC165092.1*, the predicted genes *Gm11668* (*AL732612.28*) and *13717* (*AC069469.5*), and *Vmn2r29*, which encodes for a protein involved in the G-protein coupled receptor protein signaling pathway. More known Srf target genes such as *Cfl1* and *Vcl* are identified as Srf-Avi[Bio] targets in the embryo only when the low stringency threshold is used. On the other hand, genes that are known Srf targets such as *Fos, Tpm1 or Fhl2* are missing from the embryo Srf-Avi[Bio] ChIP-Seq data. In general, this could mean that in caudal ends Srf does not bind to these targets, however, qPCR analysis of Srf-Avi[Bio] ChIP demonstrates that the *Fos* promoter is bound by Srf in embryo tails (Fig.21), showing that at least some of the genes that are actually bound by Srf in the embryo are not identified as such by the ChIP-Seq analysis, probably also due to high background. Nevertheless, it is well known that binding of Srf to its target genes differs depending on the cellular context, such as the presence of a specific Srf cofactor or extracellular signals triggering pathways affecting cofactor binding (Cooper et al. 2007). Hence, certainly not all genes bound by Srf in P19 cells are also bound in caudal ends, and vice versa.

In further analysis of these data I attempted to identify Srf-Avi[Bio] peaks that are found in the embryo but not in P19 cells and can be confirmed by qPCR. As mentioned earlier, many peaks in the embryo data set are false positives and therefore qPCR analysis did not reveal any enrichment for these peaks. However, a few peaks could be confirmed, and two of them are shown in Figure 23 (peak 1: chromosome 14; 57,312,668 - 57,313,429; peak 2: chromosome 14; 105,114,583 - 105,115,460). Interestingly, both peaks are in close proximity to putative microRNAs (miRNA). Peak 1 is associated with the novel miRNA AC154731.1-201 (ENSMUST00000157771), while peak 2 is located close to the known miRNA Mmu-mir-2135-4.1-201 (ENSMUSG00000089022). Currently, the

biological relevance of these peaks and the biological significance of the association with the microRNAs remains unclear.

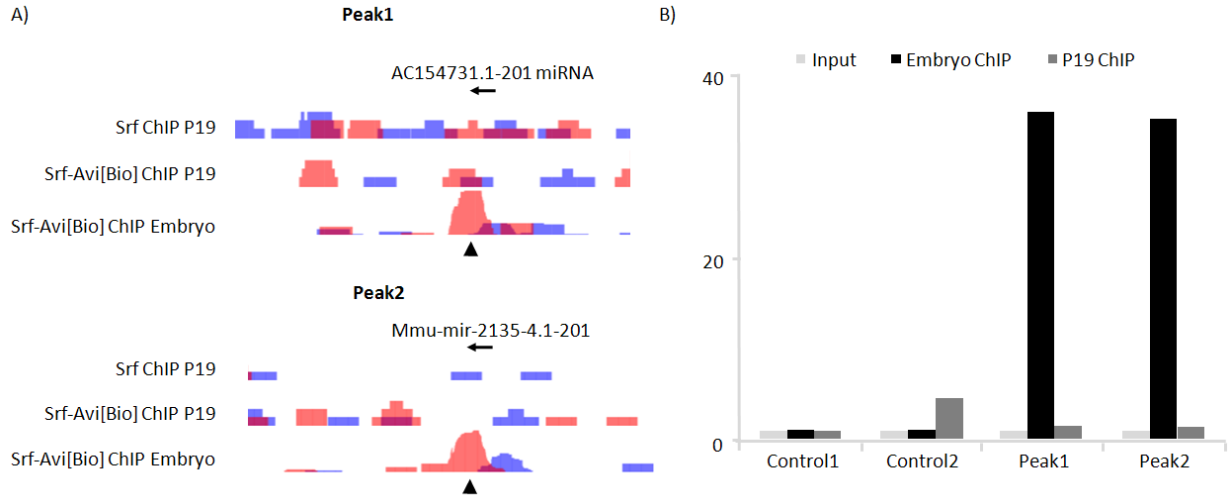

Figure 23: In vivo *Srf-Avi[Bio] ChIP allows the identification of regions that are bound by Srf in the embryo but not in P19 cells. A) Schematic of two loci which were identified to have a peak in the Srf-Avi[Bio] ChIP in embryonic caudal ends (marked by an arrowhead), but not in P19 cells and the corresponding sequence reads, depicted in blue (+strand) and red (- strand). Both peaks are in close proximity to microRNAs, depicted as black arrows. B) Confirmation of peaks by qPCR analysis of independent ChIP experiments.*

To summarize, the informative value of the Srf-Avi[Bio] ChIP-Seq data from embryo caudal ends is somewhat limited. The data are likely to contain many false positives, and not all known Srf targets were identified. However, they contain known Srf targets as well as novel putative Srf target genes, some of which are also bound by Srf in P19 cells. Others are bound in caudal ends of mouse embryos but not in P19 cells, demonstrating the differential binding of Srf to certain target sites depending on the cellular context.

2.1.5. Chapter Summary

In this part of the study I aimed to identify direct Srf target genes in caudal ends from mouse embryos by ChIP. The Srf antibody used during this work was successfully applied for ChIP with chromatin from P19 cells (Srf ChIP) (Fig.11) and the enrichment was sufficient to identify Srf targets by ChIP-Seq (Srf ChIP-Seq) on a genome-wide level (for example Fig.15), which led to the identification of both known and thus far unknown putative Srf target genes (Table 1). However, the

antibody was not sufficient for ChIP with chromatin from embryonic caudal ends (Fig.11), a problem also observed by others (P. Tam, T. Jenuwein, personal communication). The streptavidin-biotin system was employed to improve *in vivo* ChIP performance. First, I confirmed the suitability of the Srf-Avi[Bio] ChIP method using chromatin from stably transfected P19 cells (Fig.13). Next, I demonstrated that the application of the streptavidin-biotin system improves target gene enrichment as compared to the conventional antibody method when low amounts of chromatin are used for the ChIP (Fig.14). The genome wide identification of targets by Srf-Avi[Bio] ChIP-Seq essentially identified the same targets as the Srf ChIP-Seq analysis (Fig.17), resulting in the confirmation of known Srf targets and in the identification of novel putative Srf target genes by two independent ChIP methods (Table 1). Additionally, comparison of both, Srf ChIP-Seq data and Srf-Avi[Bio] ChIP-Seq data revealed that the biotinylation of Srf does not generate artifacts. Next, ES cells and transgenic embryos expressing biotinylated Srf were successfully generated (Fig.18, 19, and 20). Srf-Avi[Bio] ChIP using chromatin from these cells and embryos resulted in enrichment of known Srf targets, as seen by qPCR analysis (Fig.19 & 21). Genome wide analysis of the Srf-Avi[Bio] ChIP DNA by ChIP-Seq led to the identification of novel putative Srf target genes that were also found to be bound by Srf in P19 cells (Fig.29 & 30) or that are bound by Srf in the embryo but not in P19 cells (Fig.23).

2.2. Identification and analysis of Srf*-dependent genes*

ChIP-Seq analysis identified Srf bound regions in the genome of P19 cells and cells in embryo caudal ends. Not all genes that are bound by Srf are necessarily regulated by Srf in all cellular contexts. Instead, Srf is a transcription factor with a relatively low intrinsic activity, and cofactor binding to the Srf-DNA complex is required for transcriptional regulation (reviewed in Miano 2010). To examine to what extend Srf bound genes require Srf activity in the ps and nascent mesoderm, I aimed to compare Srf ChIP-Seq data with gene expression changes in embryo caudal ends that are deficient for *Srf*.

2.2.1. Generation of conditional Srf knock out embryos

Srf KO embryos do not generate a ps or any detectable mesoderm and die shortly after the onset of gastrulation (Arsenian et al. 1998). To overcome the early lethality, I conditionally targeted the *Srf* gene in a tissue-specific manner, using the *Cre-loxP* system. The $Srf^{flex1/flex1}$ line, which I used in this experiment, was generated by Wiebel and colleagues (Wiebel et al. 2002) and carries *loxP* sites flanking the first exon of *Srf*, which also includes the translation start site. Consequently, Cre-mediated recombination results in a non-functional *Srf* gene (Wiebel et al. 2002). I crossed the $Srf^{flex1/flex1}$ line with the tissue-specific T_{streak} *Cre* driver line, in order to obtain loss of *Srf* activity only in those cells that are in the ps and newly generated mesoderm (Perantoni et al. 2005). Embryos that are either heterozygous for the floxed *Srf* allele or that do not express the T_{streak} *Cre* recombinase are viable and do not have any phenotypic abnormalities. For simplicity, they are all regarded as heterozygous littermates.

Embryos that are homozygous for the floxed *Srf* allele and express the T_{streak} *Cre* recombinase ($Srf^{flex1/flex1}$; T_{streak} *Cre* embryos) resemble wt embryos until E8.5. However, the examination of $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos at E9.5 revealed a severe axis truncation. Mutant embryos form the trunk only up until the ~16th somite level, while more posterior structures are missing (Fig.24). In addition, for most embryos at E9.5 the tailbud seems to be “bulky”, with an excess of cells. The axis truncation is first visible in mutant embryos around E9.0. At this stage, most mutants did not complete turning, a process that embryos usually accomplish by E8.75. This result demonstrated that *Srf* is not only expressed in the ps and in nascent mesoderm (Fig.10), but that it is essential for axial elongation.

Additionally to arrested axis extension, mutant embryos exhibit a malformed heart and a thickened allantoic stalk. Both defects can be detected before the onset of the axis truncation, by E8.75. Mutant embryos die before E10.5, which is likely the result of impaired perfusion from the heart defect, or lack of nutrients caused by the malformed allantois.

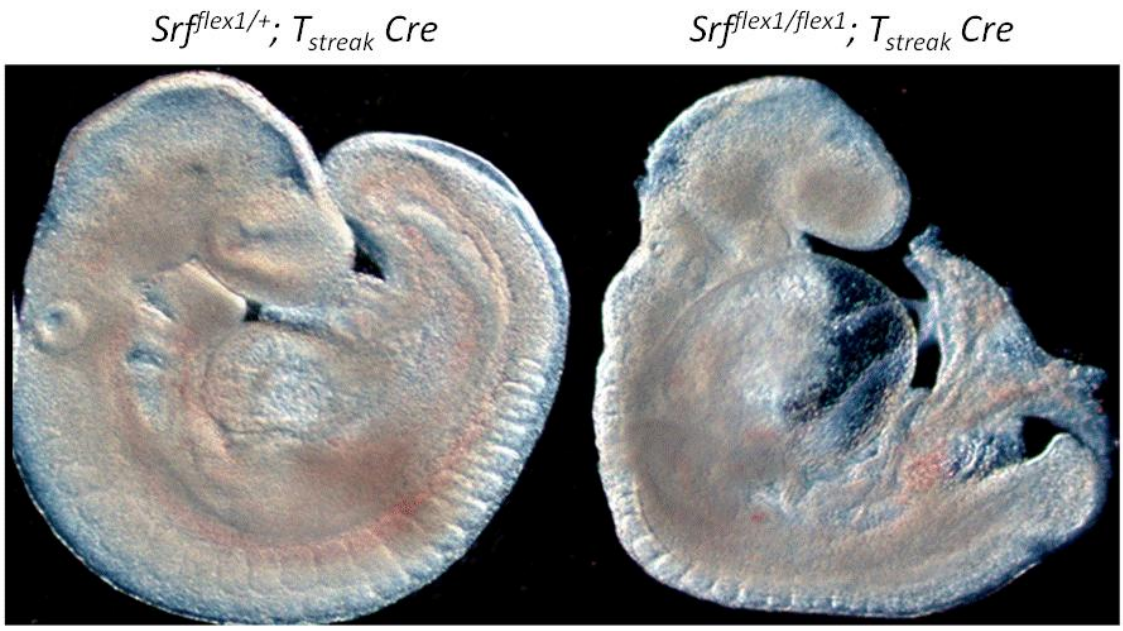

Figure 24: *Conditional* Srf *KO embryos display impaired axial elongation. At E9.5,* $Srf^{flex1/flex1}$; T_{streak} Cre *embryos (right) have a malformed heart, a thickened allantois and a shortened AP axis, as compared to their phenotypically normal heterozygous littermates (left).*

2.2.2. Identification of dysregulated genes in conditional Srf knock out embryos

Since Srf is a transcription factor, the impaired axis extension must result from misexpression of direct Srf target genes, and the dysregulation of those genes should occur before the axis truncation becomes apparent. Hence, I decided to examine a time series of differentially expressed genes in caudal ends of $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos. I chose to assay the gene expression in mutant embryos shortly before they display any abnormalities (E8.5; 3-9 somite pairs), and from embryos that exhibit malformations in the heart and allantois, but do not have an elongation defect (E8.75; 11-13 somite pairs). Additionally I assessed gene expression in mutant embryos that already displayed impaired axis elongation (E9.0; max. 16 somite pairs in mutants, 16-20 somite pairs in heterozygous littermates), which would allow me to analyze the temporal order of gene dysregulation.

In order to accomplish this, ps and newly generated mesoderm were dissected from mutant embryos and littermates, and gene expression profiles were evaluated by Illumina microarray analysis. Importantly, *Srf* transcript levels were significantly decreased at all three examined stages, indicating that recombination of the floxed *Srf* allele had occurred (Fig.25,A). Also, the total numbers of significantly dysregulated genes (E8.5 = 10, E8.75 = 139 and E9.0 = 439; fold change ≥ $\log2^{0.4}$, $p \leq 0.05$; n=4; appendix, table 12) correlates with the severity of the phenotype at the consecutive stages (Fig.25,B), which indicates an increasing trend towards secondary dysregulated genes. An increase in secondary dysregulated genes in older mutants was expected. On the one hand, the Srf deficiency is likely to result in a biased composition of tissues in the caudal end, which would cause a shift in the expression level of tissue-specifically expressed genes. On the other hand, some of the Srf target genes also encode transcription factors (such as *Egr1* and *Fos*), and an early misregulation of those factors would cause misexpression of their target genes, resulting in a compounding effect.

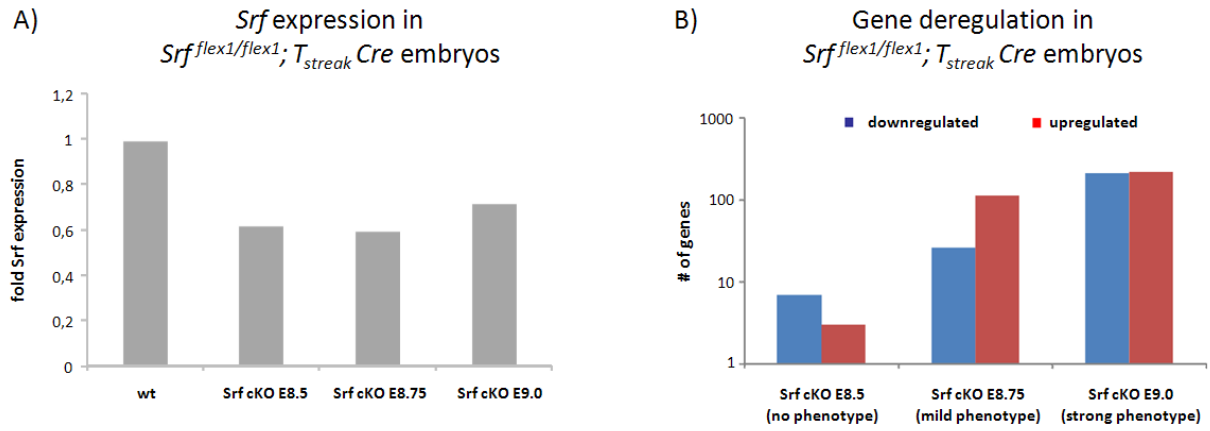

Figure 25: *A)* Srf *transcript levels are significantly decreased in caudal ends of* $Srf^{flex1/flex1}$; T_{streak} Cre *embryo at all three stages examined B) The number of dysregulated genes at E8.5, E8.75, and E9.0 increases with the severity of the phenotype. cko: conditional Knockout*

At E8.5, six genes are downregulated in addition to the targeted *Srf* gene and three genes are upregulated. Five of the six downregulated genes are known Srf targets (*Tpm1, Tpm4, Wdr1, Pdlim7* and *Acta2*). At E8.75, the number of deregulated genes is increased. Twenty-six genes are downregulated, including all of the genes found to be downregulated at E8.5 (Fig.26,B) and additional known Srf target genes (*Actb, Fhl2, Zyx* and *Cnn2*). From 113 upregulated genes, only one is a known Srf target gene (*Egr1)*, indicating an increase in secondarily regulated genes. At E9.0, the

stage when the axis truncation has become apparent, 216 genes are downregulated and 223 genes are upregulated (appendix, table 12). Twelve of the genes that are downregulated are also downregulated at E8.5 and/or at E8.75 (Fig.26,B). However, the majority of downregulated genes were not found to be deregulated before and are also not known Srf targets. This indicates a high number of secondarily regulated genes, most likely caused by the compositional bias at this stage, owing to the severity of the phenotype.

Altogether, the overlap between the various stages is much higher for the downregulated genes than for the upregulated genes. Most strikingly, no common genes are upregulated at all three examined stages. Furthermore, among the upregulated genes only very few are known Srf target genes (*Egr1*, *Fhl1* and *Anxa2*), although in total more genes are upregulated than downregulated (Fig.26,B). Both findings are in accordance with the fact that Srf is usually described as a transcriptional activator, and suggest that most of the upregulated genes are not directly regulated by Srf. In contrast, the amount of known Srf targets among the downregulated genes, particularly at E8.5 and E8.75, is much higher. The genes downregulated at these stages are more likely to cause the phenotype observed in *Srf-deficient* embryo tails, and their molecular roles are analyzed in detail in section 2.2.4.

As mentioned earlier, only 10 genes (including *Srf*) were found to be significantly dysregulated at E8.5 when applying the described filter (fold change $\geq \log 2^{0.4}$, $p \leq 0.05$). However, it is notable that the dysregulation that is apparent at E8.75 and/or E9.0 already begins by E8.5 for most genes (compare unfiltered versus filtered in Fig.26,B). This argues that lack of Srf protein already affects the transcriptional activity of the target genes during this stage of development.

To validate the microarray data, I performed whole mount *in situ* hybridization of Srf mutant embryos and heterozygous littermates, using probes against the genes Tropomyosin 1 *(Tpm1*) and Four and a half LIM domains 2 *(Fhl2).* According to the expression data, these genes are downregulated at E9.0 and E8.75, respectively, and this downregulation was clearly confirmed (Fig.26,C).

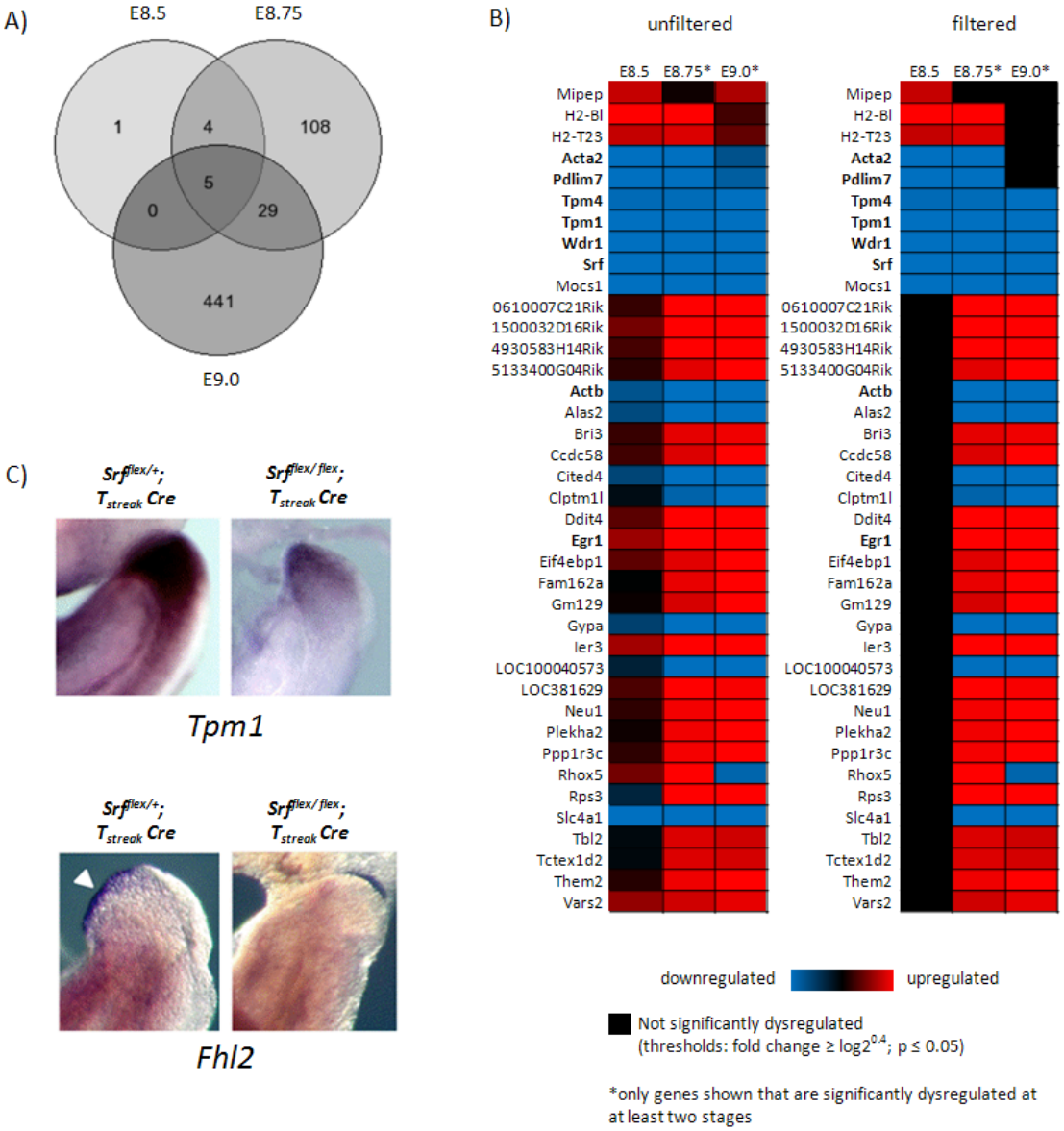

***Figure 26**: A) Venn diagram showing the overlap between the genes that are dysregulated in caudal ends of* Srf$^{flex1/flex1}$; T_{streak} Cre *embryos* at *E8.5, E8.75 and E9.0. B) The heat map shows the degree of dysregulation. Known Srf target genes are depicted in bold. In the columns E8.75 and E9.0, only genes that are dysregulated in mutants of at least two stages are shown. The degree of dysregulation is encoded by the color (see legend). The left panel shows colour encoded expression values without filtering for significantly dysregulated genes. The right panel shows gene expression that is filtered with the depicted thresholds; expression which is below these thresholds is represented in black. C) Whole mount* in situ *hybridization of* Srf$^{flex1/flex1}$; T_{streak} Cre *embryos and heterozygous littermates at E9.0 with* Tpm1- *and* Fhl2*-specific probes confirm the downregulation of both genes.*

2.2.3. Identification of Srf-dependent genes during axis extension

Genes that are both bound by Srf and dysregulated if Srf activity is abrogated are very likely to be directly regulated by Srf. Therefore, combination of the information about Srf target sites (Chapter 2.1.) and gene dysregulation in *Srf*-deficient caudal ends (Chapter 2.2.2) allows the identification of genes that are most likely directly depend on Srf in the context of mesoderm development and axial elongation. Considering the non-sufficient quality of the embryo ChIP-Seq data, genes that are identified as Srf targets in P19 cells or that were shown to be bound by Srf elsewhere are included in this analysis, and the source of information is indicated (Table 2).

At E8.5, only ten genes were significantly dysregulated in *Srf*-deficient caudal ends and six of these genes (60%) are also bound by Srf (Fig.27). Strikingly, these six genes were found to be downregulated in the absence of *Srf*. This demonstrates that the genes that are upregulated upon *Srf* deficiency are most likely indirectly repressed by Srf in the wild type embryo, in accordance with the fact that Srf mainly acts as a transcriptional activator (Schlesinger et al. 2011). At E8.75, the number of dysregulated genes was increased to 139, and only 15 of them (12.2%) were identified to be direct Srf targets. Again, the ratio of bound versus unbound genes was much higher for the downregulated genes (11 out of 26, 42.3%) than for the upregulated genes (4 out of 113, 3.5%) (Fig.27), supporting the assumption that the upregulation of genes in *Srf*-deficient tails occurs by an indirect mechanism. At E9.0, the amount of secondarily dysregulated genes is drastically increased. Only 4.6% (20 out of 439) of all genes dysregulated at this stage are found to be bound by Srf, and also when considering only the downregulated genes, the fraction of Srf bound genes is still low (15 out of 216, 6.9%). This clearly demonstrates the high impact of the compositional bias in the mutant embryos and/or the misexpression of transcription factors, which are regulated by Srf, at this stage.

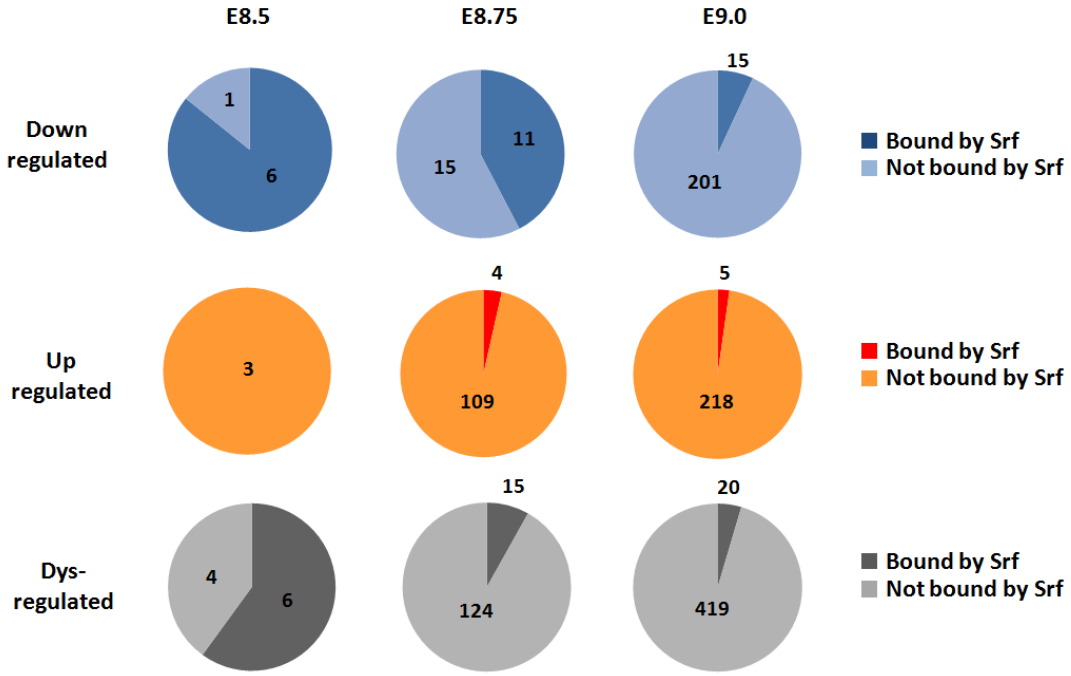

***Figure 27**: Gene expression data from Srf$^{flex1/flex1}$; T_{streak} Cre embryos at three different stages (E8.5, E8.75 and E9.0) are combined with Srf ChIP-Seq data to identify genes that are dependent on Srf in the process of mesoderm development and axial elongation. 85.7% of the genes downregulated at E8.5 are also bound by Srf. While the total number of downregulated genes increases with age of the embryos, the percentage of Srf-dependent targets drops to 42.3% (E8.75) and 6.9% (E9.0). The fraction of genes that are bound by Srf and upregulated in its absence is significantly smaller (E8.5=0%, E8.75= 3.5%, E9.0= 2.2%), which is in line with Srf being mainly a transcriptional activator.*

In general, studies that analyze gene expression following the abrogation of a transcription factor reveal that most of the dysregulated genes are not directly regulated by the transcription factor. Qin and colleagues estimate that in recent studies only 6-17% of differentially deregulated genes could be characterized as direct targets of transcription factors and that the majority of differentially regulated genes are indirectly affected (Qin et al. 2011). In comparison to these findings, the ratio of genes that were bound by Srf and dysregulated in mutants at E8.5 and E8.75 versus those that were not bound by Srf but still dysregulated is relatively high (60% at E8.5 and 14.1% at E8.75), especially when focusing on downregulated genes (60% at E8.5 and 44.8% at E8.75). This demonstrates that E8.5 and E8.75 are adequate time points for the assessment of differentially regulated genes in *Srf$^{flex1/flex1}$; T_{streak}* Cre embryos. Therefore, the genes that are deregulated at these time points are likely to be causal for the observed phenotype.

Altogether, I identified 27 genes that are differentially expressed in *Srf*-deficient tails and also bound by Srf (Tables 2 & 3). In comparison to the total number of genes that are bound by Srf (see chapter 2.1.), this is only a small fraction. However, it is in agreement with previous data that only a subset of Srf-bound genes are actually regulated by Srf in a specific context. For example, less than 10% of the genes that are bound by Srf in cardiac muscle cells were found to be deregulated upon Srf knock down in these cells (Schlesinger et al. 2011). It is thought that Srf generally binds to a large group of target genes, while only a subset of targets is regulated by Srf in each tissue or cell. Which subset of genes is being regulated depends on the cellular context, such as the presence and activity of Srf cofactors. Besides this biological explanation for the small number of Srf-dependent genes, some technical aspects have to be considered as well. First of all, the limited informative value of the *in vivo* Srf-Avi[Bio] ChIP seq does not allow one to identify all Srf target genes in embryonic caudal ends. Thus, some of the genes that are dysregulated in *Srf*-deficient caudal ends could also be bound by Srf in this tissue, but are not bound in P19 cells and are not published as Srf targets by others. Moreover, Srf binding sites can be located in enhancer regions (Sullivan et al. 2011; Aurore et al. 2003) which can be located at great distances from the corresponding gene. Since the analysis presented here only included peaks that are close to the TSS of a gene, direct Srf targets whose regulation occurs over a longer distance would not be identified. Another aspect that has to be considered is that the gene expression data used in this analysis only involve genes that are represented on the microarray (MouseRef-8 v2.0 Expression BeadChip from Illumina). In total, these are 18138 unique genes. Since 37681 mouse genes are annotated (http://www.ensembl.org/index.html, May 19th, 2011), the microarray data miss a substantial amount of the mouse transcriptome. This includes genes which were identified as Srf targets in embryo caudal ends or P19 cells (for example *Zswim6*, see Chapter 2.2.5., Fig.29). Furthermore, the microarray data do not contain non-coding RNAs, which are also targeted by Srf (e.g. *mmu-mir-143,* see Table1) and all not annotated genes (see Chapter 2.2.5., Fig.30). This being said, 27 genes were identified whose expression directly depends on Srf in the ps and nascent mesoderm, and examination of what is known about their molecular functions will reveal the major role of Srf in these tissues.

2.2.4. Srf-dependent genes are involved in cell migration

The majority of the genes that are both bound by Srf and dysregulated in *Srf*-deficient caudal ends are associated with cell migration (Fig.28; Tables 2 & 3). This is particularly true for Srf-

dependent genes that are dysregulated in *Srf*$^{flex1/flex1}$; T_{streak} Cre embryos at E8.5 and E8.75, which are likely to be the main cause for the axis truncation. At these stages, 100% (E8.5) and 71.4% (E8.75) of the dysregulated genes are implicated in cell migration.

Gene symbol	Gene name	Main function (ref.)	Dereg. at E			*Peak Position in embryo(e),P19 (p)*	*Known target?*
			8.5	8.75	9.0		
Srf	*Serum Response Factor*	Transcription factor	+	+	+	*0.5kb; e*** 3kb; p*	*Yes (*ref.1)*
Tpm1	*Tropomysin 1*	Actin filament capping, stabilizes F-actin	+	+	+	*0.2kb; e***, p** 3kb; p**	*Yes (*ref.2)*
Tpm4	*Tropomyosin 4*	Binds to actin filaments, stabilizes F-actin	+	+	+	*1kb; e***, p**	*Yes (*ref.3)*
Pdlim7/ Enigma	*PDZ and LIM domain 7*	Actin cytoskeleton organization, sequestering of nuclear factors (T-box proteins), regulation of cell survival	+	+	+	*No*	*Yes* (*ref.4)
Wdr1/ Aip1	*WD repeat domain 1 /actin interacting protein 1*	Actin binding, promotes cofilin-mediated actin filament disassembly	+	+	+	*0.4kb; p***	*Yes* (*ref.5)
Acta2	*Actin, alpha 2, smooth muscle, aorta*	Actin cytoskeleton component	+	+		*1.8kb; p*	*Yes (*ref.2)*
Actb	*Actin beta*	Actin cytoskeleton component		+	+	*0.3kb; e**, p**	*Yes (*ref.2)*
Cnn2	*Calponin 2*	Stabilizes F-actin		+		*0.3kb; p*	*Yes (*ref.3)*
Zyx	*Zyxin*	Cell adhesion, cell-cell adherens junction		+		*0.5kb; p*	*Yes (*ref.3)*
Fhl2	*Four and a half LIM domains 2*	Located in focal adhesions and actin cytoskeleton; adherens junction signaling; Srf cofactor		+		*0.5kb; e***, p**	*Yes (*ref.6)*
Rasa1	*RAS p21 protein activator 1*	Regulation of actin filament polymerization, cell survival, negative regulation of cell adhesion		+		*0.3kb; e, p*	*No*
Fscn1	*Fascin homolog 1, actin bundling protein*	Actin cytoskeleton organization, actin filament bundle assembly			+	*0.3 kb; p*	*No*
Dstn	*Destrin, aka actin depolymerizing factor, ADF*	Positive regulation of actin filament depolymerization			+	*0.7kb; p*	*Yes (*ref.6)*
Vcl	*Vinculin*	Regulation of cell migration, lamellipodia assembly, Rho GTPase binding, part of focal adhesions			+	*0.7kb; e, p*	*Yes (*ref.2)*
Mrps23	*Mitochondrial ribosomal protein S23*	Found in the mitochondrion, in a nuclear membrane location, and in the intermediate filament cytoskeleton			+	*0.4kb; p*	*No*
Olfm1	*Olfactomedin 1*	Protein oligomerization			+	*2.1kb; e*	*No*
Noc3l	*Nucleolar complex associated*	Fat cell differentiation			+	*3.3kb; e*	*No*

	3 homolog				
Cdv3	*Carnitine deficiency associated gene expressed in ventricle3*	Unknown, located in Cytoplasm	+	*0.1kb; e*	*No*
Crk	*V-crk sarcoma virus CT10 oncogene homolog*	Focal adhesion formation and turnover, cell migration, FGFR1 induced proliferation	+	*TSS; p*	*No*
Immt	*Inner membrane protein, mitochondrial*	Integral to mitochondrial inner membrane	+	*0.1kb; e*	*No*

Table 2: Srf-dependent genes that are downregulated in the absence of* Srf.** *Shown are the genes that were both downregulated in $Srf^{flex1/flex1}$; T_{streak} Cre embryos at E8.5, E8.75 or E9.0 and identified to be bound by Srf in P19 cells and/or in embryonic caudal ends, and/or that were previously described as Srf targets. Position of peaks and references for known targets are indicated (*Ref.1:* (Spencer & Ravi P Misra 1999)*; Ref.2: (Olson & Nordheim 2010); Ref.3: (Miano et al. 2007); Ref.4: (Jung et al. 2010); Ref.5: (Luxenburg et al. 2011); Ref.6: (Sun et al. 2006). In some cases ChIP-Seq enrichment was confirmed by qPCR (indicated by**) or enrichment was only detected by qPCR (indicated by) (Fig.S.3)*

Gene symbol	Gene name	Main function (ref.)	Dereg. at E 8.5	8.75	9.0	*Peak Position in embryo (e) ,P19 (p)*	*Known target?*
Egr1	*Early growth response 1*	Regulation of proliferation		+	+	+1kb; e**; p**	Yes (*ref.7)
Naca	*Nascent polypeptide-associated complex alpha polypeptide*	Muscle-specific transcription factor		+		+0.1kb; p	No
Ndufc2	*NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 2*	Mitochondrial electron transport		+		+0.5kb; p	No
Soat2	*Sterol O-acyltransferase2*	Lipid metabolic process		+		-2.7kb; p**	No
Rps6ka1	*Ribosomal protein S6 kinase polypeptide 1*	Transcription factor binding activity, protein phosphorylation			+	1.6kb; e	No
Fhl1	*Four and a half LIM domains 1*	Zinc ion binding, Srf cofactor			+	0.2kb; e	*Yes (*ref.6)*
Nlgn2	*Neuroligin 2*	Neuron cell-cell adhesion			+	1.4kb; p	No
Anxa2	*Annexin2*	Binds to F-actin, implicated in cell cycle control in confluent endothelial cells			+	No	*Yes (*ref.3)*

Table 3: Srf-dependent genes that are upregulated in the absence of Srf. *The table shows the genes that were both upregulated in $Srf^{flex1/flex1}$; T_{streak} Cre embryos at E8.75 or E9.0 and identified to be bound by Srf in P19 cells and/or in embryonic caudal ends. The position of the peaks and references for known targets are indicated (Ref.3: (Miano et al. 2007); Ref.6: (Sun et al. 2006); Ref.7:(Watson et al. 1997). In some cases ChIP-Seq enrichment was confirmed by qPCR (indicated by**) (Fig.S.3)*

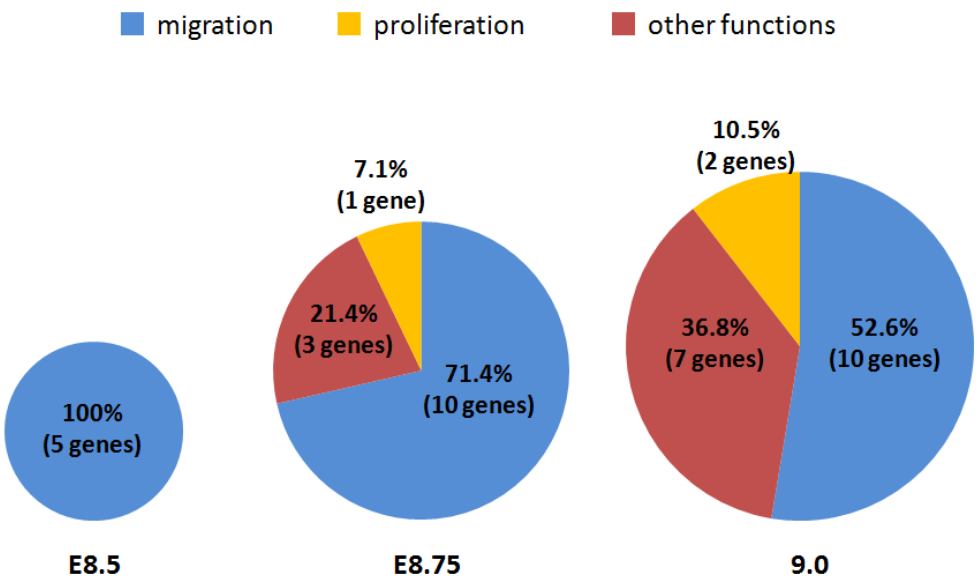

Figure 28: *Clustering of Srf-dependent genes. The pie charts show how many of the Srf-bound genes are dysregulated in* Srf*-deficient caudal ends at the three examined stages, and with which cellular functions the genes are associated.*

A role for Srf in the regulation of cell migration was recently shown *in vitro* (Schratt, Philippar, et al. 2002). Schratt and colleagues demonstrated that *Srf*$^{-/-}$ ES cells display impaired cell spreading, adhesion, and migration and that this correlates with defective formation of actin stress fibers and focal adhesion (FA) plaques. This phenotype was further shown to be caused by the downregulation of Beta-actin (*Actb*) and of several other genes encoding FA proteins such as Talin *(Tln1),* Zyxin *(Zyx)* and *V*inculin *(Vcl). Actb, Zyx* and *Vcl* were also found to be downregulated in *Srf*-deficient caudal ends, suggesting that actin fiber and focal adhesion formation may also be affected in conditional *Srf* KO tails. In addition, many other genes that are *Srf*-dependent in the ps and nascent mesoderm were associated with cell migration (Tables 2&3).

The known Srf target genes Tropomyosin 1 *(Tpm1)* and Tropomyosin 4 *(Tpm4),* Destrin *(Dstn)* and the WD-repeat protein 1 *(Wdr1)* (also called *Aip1,* actin interacting protein 1) were all downregulated in *Srf*-deficient caudal ends. They encode factors crucial for the regulation of actin cytoskeletal flexibility, which is a hallmark of cell migration (Lacayo et al. 2007; Kato et al. 2008). Similarly, Calponin 2 *(Cnn2),* Fascin homolog1 *(Fscn1)*, and v-crk sarcoma virus CT10 oncogene homolog *(Crk)* encode proteins that bind to F-actin, and are implicated in its bundling and/or rearrangement, and thus also contribute to the regulation of cell migration (Huang et al. 2008; Feller,

2001; Antoku, Mayer, 2009; Sedeh, 2010). While *Cnn2* was previously demonstrated to be an Srf target, *Fscn1* and *Crk* are two of the novel putative Srf targets identified in the course of this work.

The genes four and a half LIM domain 2 (*Fhl2*) and LIM domain protein PDZ and LIM domain 7 (*Pdlim7*; also called Enigma) encode factors that bind to integrins at focal adhesion sites (*Fhl2;* Wixler et al. 2000) or to F-actin tropomyosin-complexes (*Pdlim7;* Guy et al. 1999). However, these factors are also involved in the regulation of signaling cascades such as the MAPK pathway and the control of transcription factor activity (Jung et al. 2010; Niu et al. 2007; Guy et al. 1999; Camarata et al. 2010; Camarata et al. 2006). They can therefore be regarded as players in the complex network linking the cell cytoskeleton to the regulation of gene transcription (see 3.4.). While *Fhl2* was shown to be a Srf target gene in many cellular contexts, the gene *Pdlim7* was only shown to be directly bound by Srf in Mouse Embryonic Fibroblast (Jung et al. 2010). It was also shown to be downregulated in the embryonic heart of Srf-knockout mice (Niu et al. 2007). Transcript levels of *Pdlim7* are decreased in caudal ends of Srf mutants at E8.5, E8.75 and E9.0. Its early downregulation indicates that it is a direct Srf targets, since most of the genes dysregulated at this stage are also bound by Srf (Fig.28). However, enrichment of *Pdlim7* could not be detected by ChIP-Seq or by ChIP-qPCR analysis (data not shown). This could indicate that either *Pdlim7* is not a direct Srf target in caudal end cells and P19 cells, or that Srf binds to an enhancer element that is not in the close proximity of the *Pdlim7* gene. The immediate early gene *Egr1* is usually activated by the TCF factor Elk1 in combination with Srf downstream of the Ras-MAPK pathway (Alexandre et al, 1991). However, the gene is one of eight genes that are upregulated in caudal ends of $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos at E8.75 and E9.0, suggesting that in the ps and/or nascent mesoderm, *Egr1* is inhibited either directly by Srf or by another Srf regulated mechanism. *Egr1* encodes a transcription factor that can both act as a repressor and as an activator of transcription (Bahouth et al. 2002; Chapman & Perkins 2000; Lemaire et al. 1990; Wang et al. 2005), and is implicated in the regulation of cell proliferation and apoptosis (Virolle et al. 2003).

2.2.5. Analysis of additional, novel putative Srf target genes

The combination of ChIP-Seq data with gene expression data from Srf-deficient caudal end yielded the identification of genes that are most likely directly dependent on Srf activity (Table 2 and 3). Thirteen of these genes are not yet published as Srf target genes, meaning that they represent novel putative Srf targets. As mentioned earlier, assessment of expression data by microarray has the drawback that not all annotated genes are represented on the array. Moreover, it does not provide

information about microRNAs and not yet annotated genes. In contrast, ChIP-Seq analysis covers the whole genome. As a consequence, many Srf peaks could not be correlated with genes represented on the microarray. Genes that were not represented on the microarray but have an Srf peak in their promoter region can be assayed for differential regulation in *Srf*-deficient caudal ends by qPCR. This analysis extends the identification of genes that are direct Srf targets in the course of axis extension. Here I focused on three genes in particular.

As mentioned above, the majority of Srf binding sites are located within 4kb from the TSS of a gene (Sun et al. 2006). When analyzing the peaks within this distance, the same peak was found to have the highest enrichment in all three Srf ChIP data sets (Srf Chip, P19; Srf-Avi[Bio], P19; Srf-Avi[Bio], embryonic caudal ends). Remarkably, this peak corresponds to the gene *zinc finger, SWIM domain containing 6 (Zswim6)*, which was not previously described as Srf target gene. The peak could be confirmed by qPCR analysis of an independent Srf-ChIP experiment using lysate from P19 cells (Fig.29,B). Bioinformatic analysis of the peak, which is located in close proximity to the transcription start site (approx. 130 bp), revealed that it harbors three imperfect putative Srf binding sites at the positions +109-99bp, +90-80bp and +35-25bp relative to the TSS (Fig.29,A). The protein encoded by the gene contains a Zn-finger-like domain and a Swim domain, predicted to have zinc ion binding, DNA-binding, and protein-protein interaction activity, respectively, however, there are no experimental data available on this protein. All three putative binding sites have two mismatches compared to the known CArG consensus sequence ($CC[AT]_6GG$), but it has been shown that Srf is also able to bind to imperfect binding sites, possibly accompanied by a reduction in Srf-binding affinity (Chang et al. 2001; Reinke et al. 2008). Whole mount i*n situ* hybridization analysis with a *Zswim6*-specific probe revealed that it is similarly expressed to *Srf*, with the strongest expression domains in the psm and in forming somites (Fig.29,C).

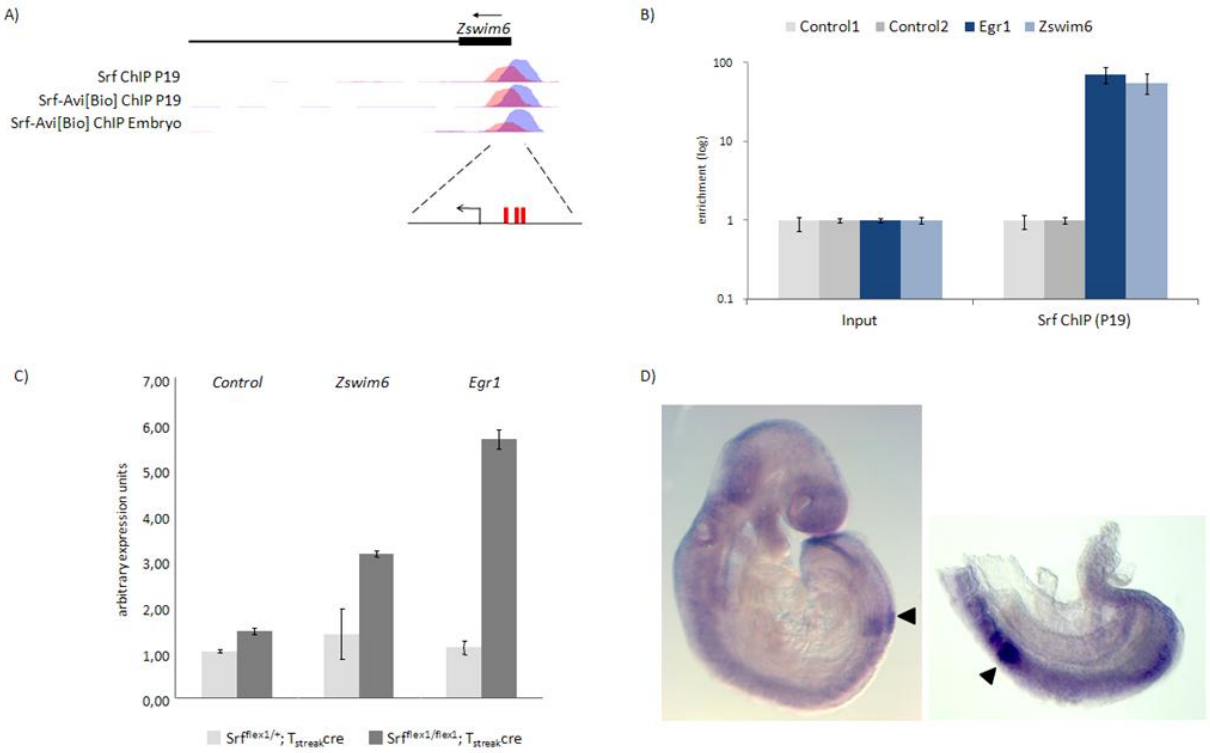

Figure 29: Zswim6 *is a novel putative Srf target gene. A) Schematic of the genomic locus of* Zswim6*. Exon 1 is depicted as black box, intron 1 as a black line. Orientation of the gene is indicated (arrow). Peaks generated by Srf ChIP and Srf-Avi[Bio] ChIP from P19 and embryo ChIPs are depicted in blue (+strand) and red (- strand).Three imperfect Srf binding sites (two mismatches compared to the known CArG box) are located in the middle of the peak (depicted as red bars in the scheme). B) Enrichment of* Zswim6 *could be confirmed by an independent Srf-ChIP experiment using lysate derived from P19 cells C)* Zswim6 *expression is upregulated in tail buds of* $Srf^{flex1/flex1}$*;* T_{streak} *Cre embryos at E8.75, shown by qPCR analysis.* Egr1 *upregulation is shown for comparison and gene expression is normalized to the expression of the housekeeping gene* Pmm2*, control gene is the housekeeping gene* Hbm2*. D) In situ hybridization analysis with a* Zswim6*-specific probe reveals expression in the tailbud and the paraxial mesoderm with highest expression in forming somites (arrowhead).*

Since *Zswim6* is not represented on the microarray used for the gene expression analysis of Srf deficient caudal ends (2.2.2.), I performed qPCR analysis with cDNA from $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos and heterozygous littermates at E8.75 and E9.0 to assess a putative dependency on Srf activity. At both stages, *Zswim6* expression was slightly upregulated in *Srf*-deficient tails, suggesting that it is inhibited by Srf either directly or indirectly, similar to *Egr1* (Fig.29,B). However, *in situ* hybridization on $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos analysis revealed no observable change in expression of *Zswim6* following loss of *Srf* (Fig.S5).

The Srf ChIP-Seq analysis described above focuses on Srf peaks that are found within a 4kb distance from the TSS of an annotated gene. However, the majority of Srf peaks are found in intergenic regions (see 2.1.4.3. and Fig.22). In particular, two such Srf peaks are located downstream of the well-known Srf target genes *Egr1* and *Fos* (Fig.30,A). The distances from the 3' ends of *Egr1* and *Fos* to their corresponding downstream peaks are 22kb and 12kb, respectively. Bioinformatic analysis identified two imperfect CArG boxes (one mismatch) in the Srf peak downstream of *Egr1* and one CArG box in the Srf peak downstream of *Fos,* indicating that both peaks represent *bona fide* Srf target sites. In addition, the peaks can be confirmed by qPCR analysis of Srf-Avi[Bio] ChIP with chromatin from P19 cells (Fig.S.2). However, no transcripts are reported in the genomic regions in which the peaks are located, suggesting that they either correlate with enhancer regions or with weakly expressed, thus far unknown transcripts. Evidence that they are actually associated with active transcription comes from experiments performed by our lab, which analyzed the trimethylation state of lysine tails on histone 3. These histone modifications are located at TSS of genes and correlate with active gene transcription (H3K4 trimethyl ChIP by Grote, Werber, Wittler, and Herrmann; personal communication) (Fig.30,A and B). Therefore, I aimed to identify transcripts that are encoded by these regions. Indeed, qPCR analysis using cDNA from embryonic tissue and several primers located around the Srf peaks resulted in amplification of putative transcripts. To analyze whether their transcription depended on Srf activity, I performed qPCR analysis with cDNA from $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos and littermates, which implied that both transcripts are upregulated upon Srf abrogation. This indicates that they are, similar to *Egr1* and *Zswim6*, repressed by Srf in the process of mesoderm generation and axis extension. The primers that were used for the qPCR amplification span 95bp close to the *Egr1* downstream peak and 67bp near the *Fos* downstream peak. However, the transcripts are likely to be considerably longer, therefore, they could belong to the class of long non-coding transcripts (long ncRNAs), which usually have a minimum length of 100bp and, like protein coding genes, display H3K4 trimethylation marks. It has been shown that long ncRNAs can have an enhancer-like function, specifically activating the transcription of neighboring protein coding genes (Ørom et al. 2010). Assuming a positive effect on the transcription of *Egr1* and *Fos* by the neighboring putative long ncRNAs, it is plausible that they are coregulated by Srf, adding a potential mechanism by which Srf induces the transcription of immediate early genes.

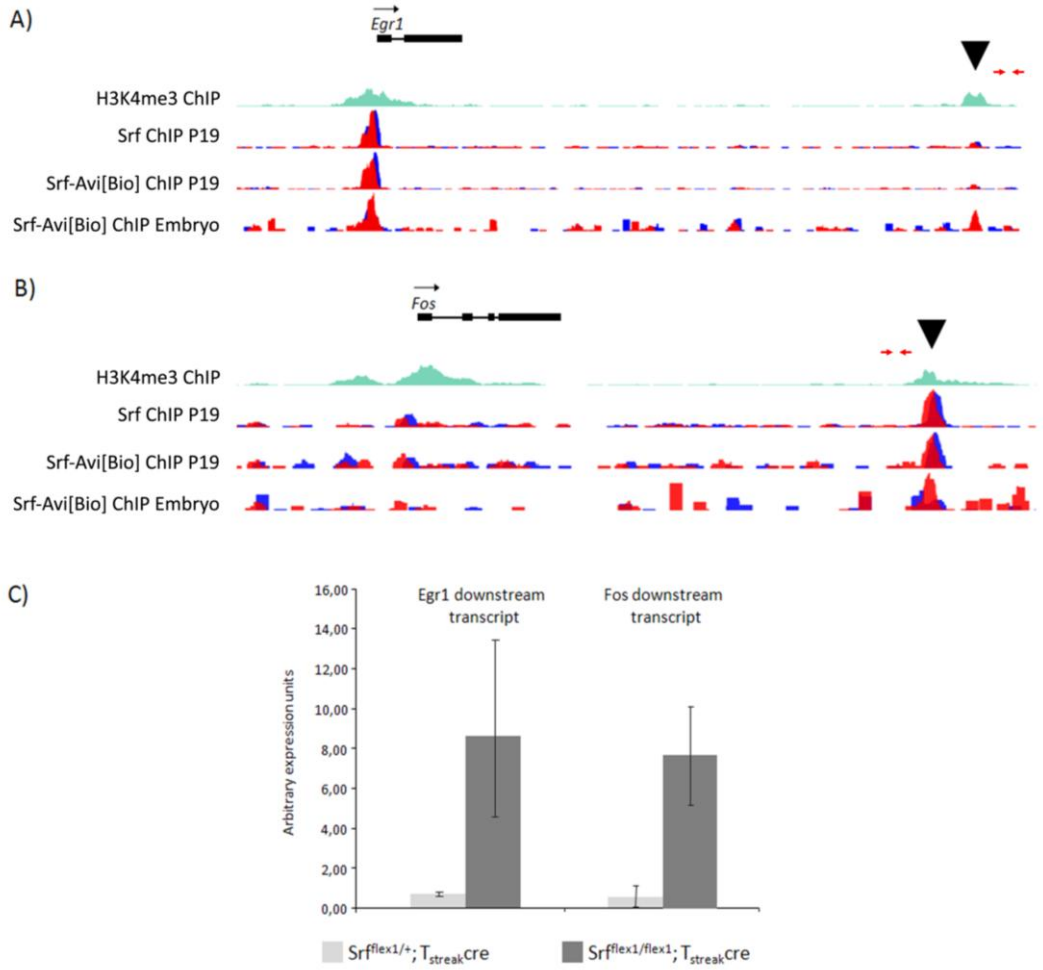

Figure 30: *Identification of Srf-regulated transcripts. Two Srf peaks are located downstream of the Srf target genes* Egr1 *(A) and* Fos *(B) (indicated with an arrow head), both in P19 cells and in embryonic caudal ends. The Srf peaks colocalize with H3K4 trimethyl marks (green lane), indicating transcriptional activity. C) Both transcripts are upregulated in* $Srf^{flex1/flex1}$ *;* T_{streak} *Cre embryos vs.* $Srf^{flex1/+}$ *;* T_{streak} *Cre embryos at E9.0 (n=2), analyzed by qPCR and normalized to the housekeeping genes* Hbms *and* Pmm2*. Localization of the primers for qPCR is indicated by red arrows in A) and B).* Srf *occupation and dysregulation in* Srf*-deficient tailbuds indicates that both transcripts are Srf target genes in the process of mesoderm development and axial elongation.*

2.2.6. Chapter Summary

In this chapter I show that conditional KO of *Srf* in the primitive streak and the presumptive mesoderm results in embryos with a truncated antero-posterior axis (Fig.24). To analyze this defect on a molecular level, I assessed the differential gene expressivity in *Srf*-deficient caudal ends. I generated a series of gene expression profiles, ranging from mutant embryos before the phenotype occurs (E8.5) to mutants with a severe apparent axis truncation (E9.0) (Fig.25 & 26). The data were combined with the Srf ChIP-Seq data described above to identify genes that are both bound by Srf and dysregulated upon Srf abrogation (Fig.27). This analysis revealed a set of previously known Srf target genes, which can be considered bona fide Srf targets in the tissues contributing to axial elongation (Table 2). Additionally, several genes were identified that had not been previously described as Srf targets (Table 2), and these genes are most likely directly dependent on Srf activity in the ps and nascent mesoderm. Together, the vast majority of the genes that are both bound by Srf and deregulated upon *Srf* abrogation display a decreased expression in *Srf*-deficient caudal ends. This indicates that in the wild type embryo, their transcription is induced by Srf, in agreement with the fact that Srf is mainly acting as a transcriptional activator (Schlesinger et al. 2011). Moreover, the ratio of Srf bound versus unbound genes is very high in mutants at E8.5, and decreases drastically with the development of the axis truncation (Fig.27). This demonstrates an increase of indirectly dysregulated genes, probably resulting from a bias in tissue composition in the absence of *Srf*. It also implies that the genes identified to be bound by Srf and dysregulated at E8.5 and also at E8.75 comprise the factors causal for the axis truncation. Nearly all of these genes, including previously described as well as novel putative Srf target genes, are associated with actin cytoskeleton assembly and focal adhesion formation (Fig.28; Table 2). Since both of these processes are hallmarks of cell migration, the identification of genes that are regulated by Srf indicates that the main function of Srf in the process of axis extension is the regulation of cell motility. Additionally, further genes that were not represented on the microarray were identified to most likely be direct Srf target genes in the ps and nascent mesoderm. These genes have not yet been described as Srf targets and comprise protein coding as well as putative long ncRNAs (Fig.29 and Fig.30, respectively). However, their functions remain to be elucidated.

2.3. Loss of Srf impairs migration behavior of nascent mesodermal cells

2.3.1. Cells from Srf-deficient caudal ends display impaired motility ex vivo

The functional analysis of Srf-dependent genes shows that the primary role of Srf in the ps and nascent mesoderm could be the regulation of cell migration. To assess if Srf activity is required for cell migration, I performed an *ex vivo* migration assay. At E9.5, tailbud halves were dissected from mutant and heterozygous control embryos and the explants were cultured on µ-slides (ibidi; Martinsried), which had been coated with the extracellular matrix protein fibronectin (Fn). Fn binds to membrane-spanning proteins such as integrins and thus enables the attachment of cells to the ECM via focal adhesions. Explants from both mutant and control embryos attached within one hour and cells started to migrate in all directions shortly thereafter. One day later, there was already a significant difference in the size of the area that was covered by cells emerging from *Srf*-deficient or control explants. On the second day, this difference was increased, and the maximal migration distance was measured as the length between the original explant and the border of the occupied area in each sample (Fig.31,A). As displayed in Figure 31,B, cells emerging from the five mutant explants displayed a significantly shorter maximal migration distance (604µm +/- 71µm (n=5)) than cells emerging from the control explants (955µm +/-59µm (n=5)) (Students T-test: p= 0.0017). This illustrates a clear and significant difference in the migration ability of *Srf*-deficient cells, demonstrating that Srf not only regulate migration-associated genes, but is essential for the migration of mesodermal cells.

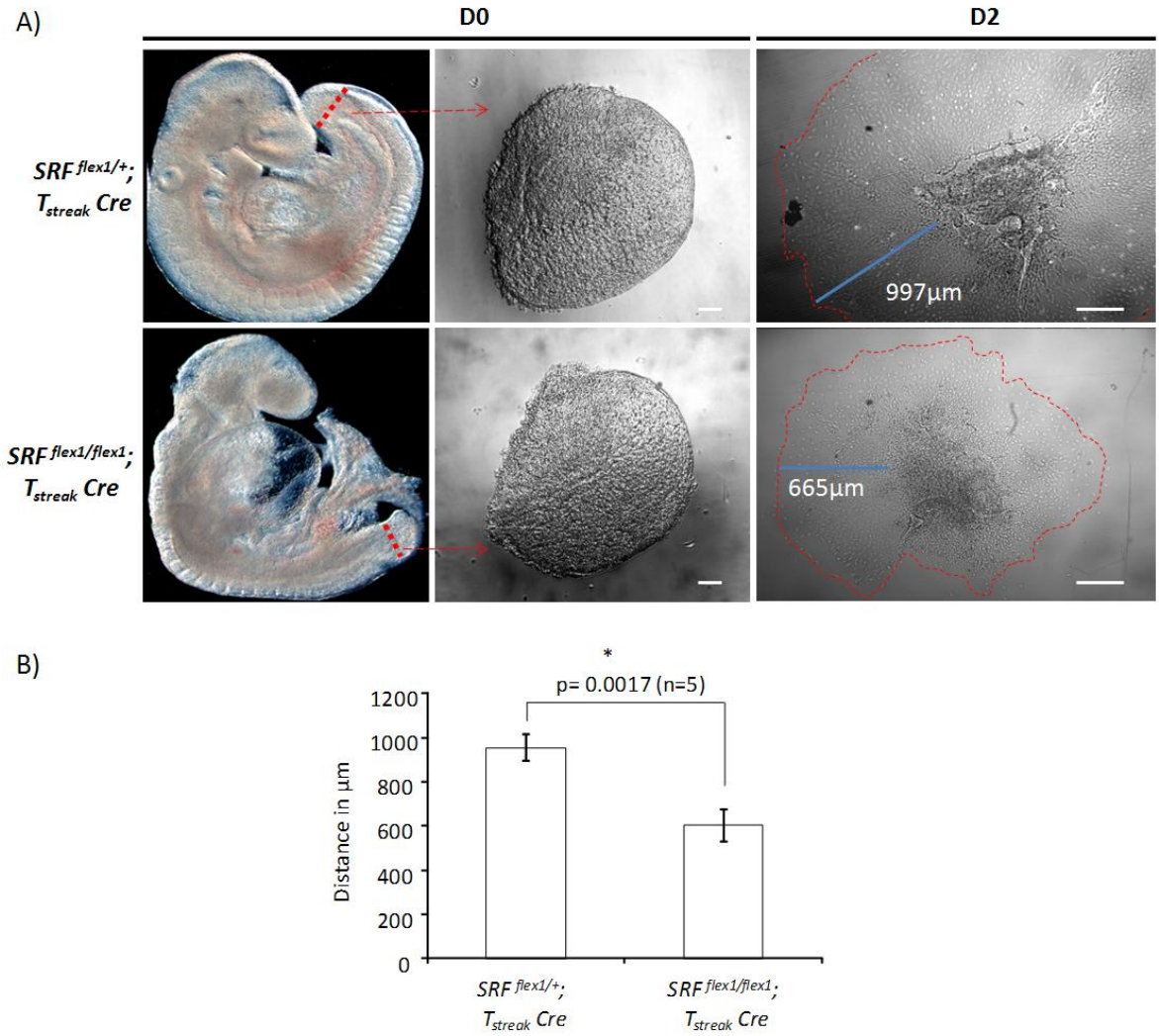

Figure 31: *Cells from conditional Srf KO explants migrate significantly less than control explants.* A*) Tailbud halves from conditional Srf KO embryos (bottom) or heterozygous littermates (top) were cultured on a fibronectin coated surface. Two hours after plating, the explants had attached to the surface and displayed a similar size (middle images). Two days later, cell spreading was measured as the maximal distance (blue line) between the original explant and the border of the occupied area (red dotted line). Scale bars (white) indicate 150µm B) Average maximal migration distances from five mutant and five control explants***.**

In addition to the impaired migration, cells from control samples were long and thin, typical hallmarks of mesenchymal cells, whereas cells from mutant tailbuds were rounded, suggesting differences in the organization of the cytoskeleton (Fig.32,A).

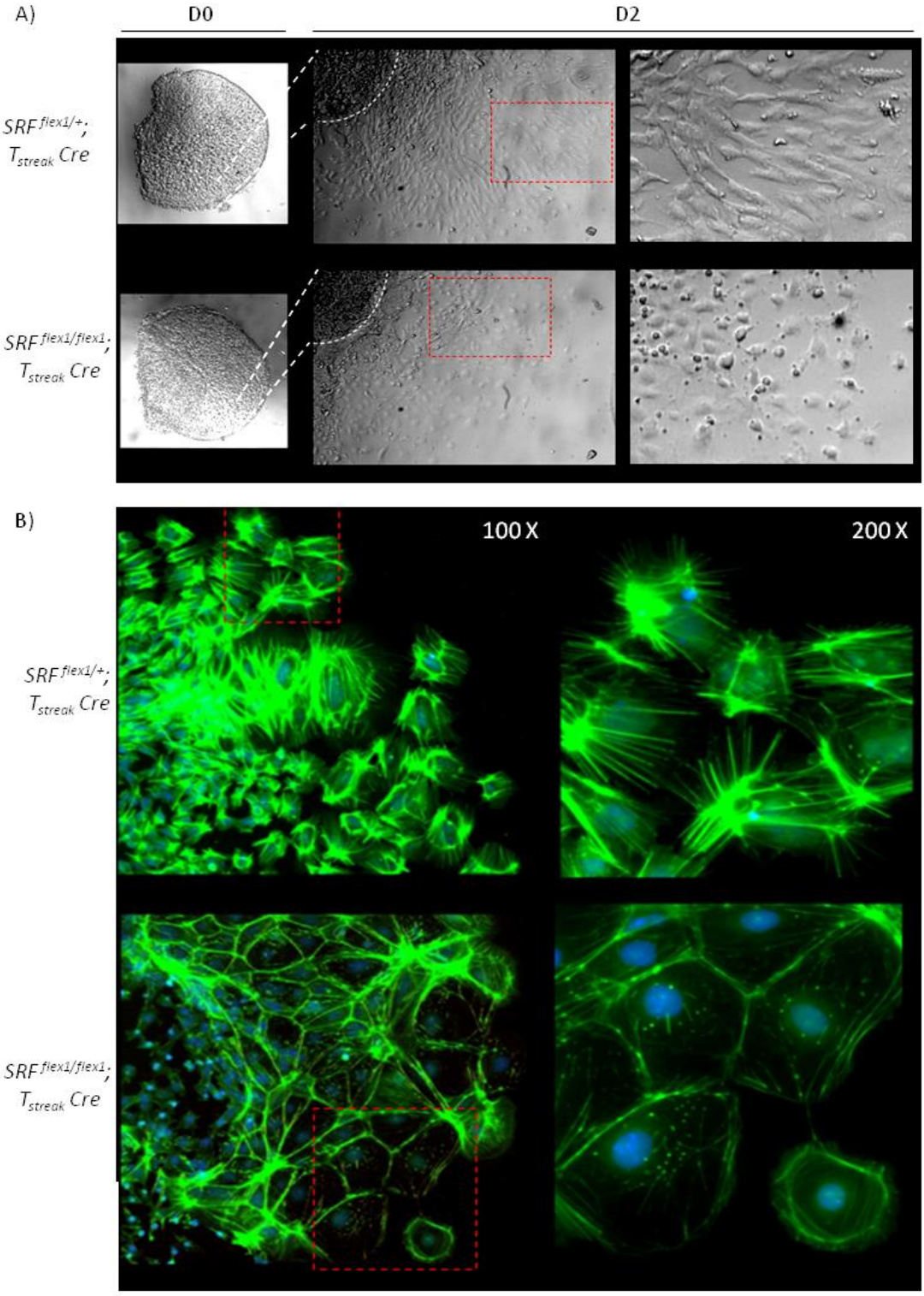

Figure 32: *Cells from Srf$^{flex1/flex1}$; T_{streak} Cre caudal ends display impaired cell migration and drastic differences in the structure of the actin cytoskeleton. A) Bright field images from caudal end explants two hours (D0) and two days (D2) after plating. Magnifications (right) show that control cells are long and thin, resembling mesenchymal cells, while cells from conditional* Srf *KO explants are rather rounded. Dotted squares indicate the position of the area that is shown in the magnifications and in*

B). B) F-actin labeling with Phalloidin-FITC shows structural differences in the composition of the actin cytoskeleton. and an overall decrease in F-actin in Srf-*deficient cells. Left: 100x magnification, right: zoom in of dotted squares in images on the left.*

To analyze differences in the cytoskeletal composition between both samples, I stained for F-actin in the tissue explants using Phalloidin-FITC conjugate. The staining exposed an overall lower F-actin level in cells originating from *Srf*-deficient caudal ends (Fig.32,B), which is in agreement with the observed decrease in expression of the actin encoding genes *Actb* and *Acta1* (Table 2). Furthermore, it revealed a drastic difference in the assembly of the cytoskeleton. In cells from control samples, F-actin is mainly incorporated into parallel bundles of stress fibers, which traverse the whole cell body, resembling the cytoskeletal organization of a migrating cell. In sharp contrast, actin stress fibers are nearly absent in cells from Srf-deficient explants. Instead, theses cells display circumferential bundles along the cell borders, which is characteristic of epithelial cells (Gloushankova et al. 1997)(Fig.32,B).

2.3.2. In vivo *analysis of impaired motility in* Srf-*deficient mesoderm*

The genes identified to be regulated by Srf in the Ps and nascent mesoderm indicate that, in the ps and nascent mesoderm, Srf regulates genes that are associated with cell migration (Fig.28, Table2), and an *ex vivo* migration assay demonstrated that *Srf*-deficient mesodermal cells display impaired migration. All these data were based on the activity of the T_{streak} promoter. To investigate the migration defect *in vivo*, I performed immunofluorescence staining of mutant embryos using an antibody against Brachyury (T) to stain those cells, in which the T_{streak} promoter is active. This allowed for the detection of cells in which Srf usually should activate genes required for stress fiber and focal adhesion plaque formation.

At E9.5, wt embryos and embryos heterozygous for a non-functional *Srf* allele display T-positive (T^+) cells in the tailbud, the psm, the notochord, and the allantois (Fig.33). In contrast, in conditional *Srf* KO tails, T^+ cells seem to accumulate at the posterior end of the embryo, while no T^+ cells are found in the more anterior side of the psm, and staining in the notochord is much weaker (Fig.33). Absence of T^+ cells in anterior regions of the psm can also be observed in sagittal sections of embryo tails (Fig.33,B, top). The generation of new mesodermal cells occurs in the posterior end of the tail, and spreading of these cells throughout the psm requires migration of individual cells (Bénazéraf et al. 2010), therefore, the data presented here indicate that *Srf* deficiency in mesodermal cells also causes a migration defect *in vivo*.

In order to gain motility, cells in the tailbud undergo EMT. A fundamental early step in EMT is the repression of the cell-adhesion molecule E-cadherin (Cdh1). *Cdh1* expression is inhibited by the transcriptional repressor Snai1, and failure in expression of *Snai1* or in downregulation of *Cdh1* results in a truncated axis (Ciruna & Rossant, 2001). The microarray data from $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos did not show a differential expression of *Cdh1* or *Snai1*. However, it was previously shown that during gastrulation E-cadherin is also regulated at the protein level (Zohn et al. 2006). To test whether Srf activity is required for the degradation of the E-cadherin protein, which could also be an explanation for the observed migration impairment, I performed immunofluorescence staining with an E-cadherin-specific antibody. Figure 33,B shows consecutive sections of an *Srf*-deficient mouse tail, stained against T or E-cadherin. It reveals that T^+ cells, which accumulated in the posterior end of mutant embryos are negative for E-cadherin, indicating that their impaired motility is not due to impaired degradation of E-cadherin.

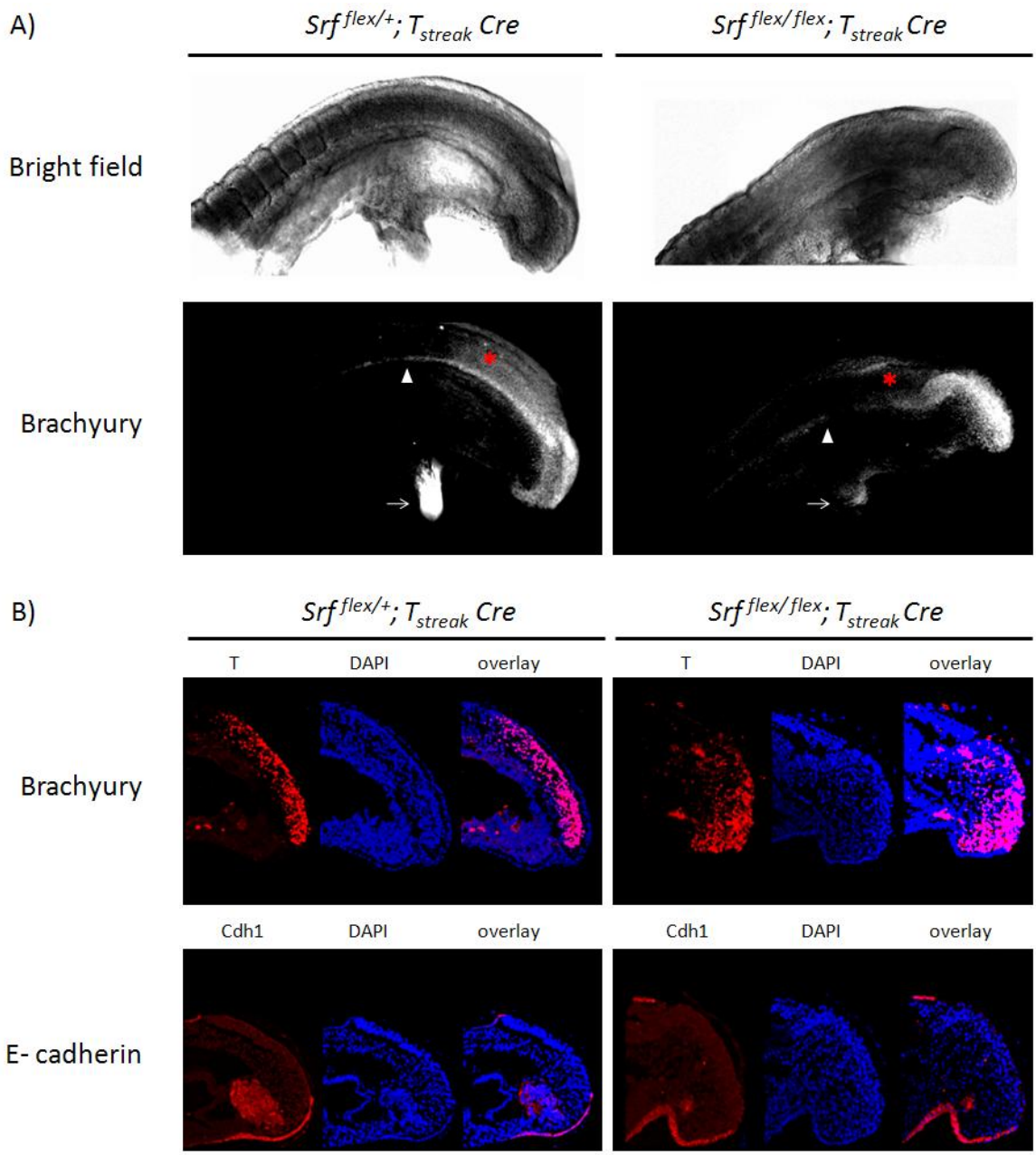

Figure 33: *Immunofluorescence staining indicates a migration defect in conditional* Srf *KO embryo tails. A) Bright field (top) and fluorescence images (bottom) of whole embryo tails (E9.5). In conditional* Srf *KO embryos, Brachyury positive cells (white staining) are located in the tip of the tail but are not found in the psm (marked by *). In contrast, in heterozygous littermates the Brachyury signal can also be detected in the unsegmented psm (*). Staining in the notochord (arrowhead) and the allantois (arrow) is much weaker in mutant embryos. B) Staining of paraffin sections of conditional* Srf *KO embryo tails (E9.5) (top) confirms the absence of Brachyury positive cells (red) in the anterior psm, while they are spread throughout the unsegmented psm in heterozygous littermates. Brachyury positive cells are negative for E-cadherin in control and mutant embryos, while cells of the epithelial layer at the ventral side of the tails are E-cadherin positive (bottom).*

Several of the genes that are regulated by Srf in embryonic caudal ends are directly involved in the regulation of proliferation (*Egr1, Crk; Fhl2*) and apoptosis (*Pdlim7, Rasa1*). To test for putative effects of *Srf*-deficiency on apoptosis or proliferation in embryonic caudal ends, I stained for the apoptosis marker cleaved caspase3 and the proliferation marker Histone3-Serine10-phosphate (phosphoH3(S10)). While there was no effect seen on apoptosis, the proliferation rate seems to be slightly reduced in *Srf*-deficient tails (Fig.34). However, the microarray data from $Srf^{flex1/flex1}; T_{streak}\,Cre$ embryos showed that the cell cycle-dependent genes *cyclin D1* and *cyclin E1* were not deregulated upon loss of *Srf* in any of the examined stages, arguing against an altered proliferation rate. Whether there is a decrease in cell proliferation and if this has an impact on the observed phenotype remains to be tested.

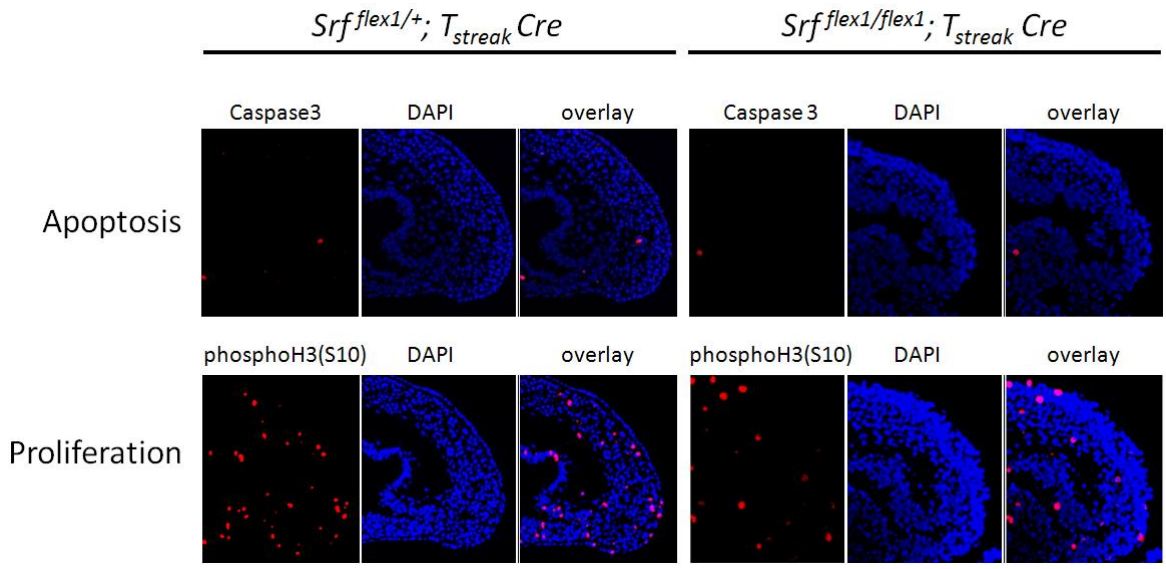

Figure 34: *Immunofluorescence staining of tailbuds from conditional* Srf *KO embryos and heterozygous littermates reveals no differences in apoptosis (caspase3, red) (top row), while proliferation seems to be slightly reduced (phosphoH3(S10), red) (bottom row). Nuclear counterstain with DAPI (blue).*

2.3.3. Analysis of $Srf^{flex1/flex1}$; Msd Cre embryos

The conditional *Srf* KO embryos analyzed thus far were generated by using the T_{streak} Cre recombinase, which promotes recombination in all descendants of the ps. The axis truncation observed in the $Srf^{flex1/flex1}; T_{streak}\,Cre$ embryos results from impaired cell motility, directly owing to the lack of *Srf*. Recent publications confirm a connection between the migration of mesodermal cells and axis extension. Work in chick embryos suggests that movement of psm cells along the antero-

posterior axis is essential for axis extension, with a high cell motility in the posterior psm and decreased motility more anteriorly (Bénazéraf et al. 2010).

In order to narrow down the region in which Srf-mediated cell migration is crucial for axis extension in the mouse, I generated conditional *Srf* KO embryos using the *Dll1-msd Cre* driver line (When et al. 2009) (*Srf*$^{flex1/flex1}$*; Msd Cre* embryo). In this line, the Cre recombinase is driven by the Dll1 mesoderm enhancer, located upstream of a minimal promoter. It differs from the T_{streak} promoter driven recombinase in two main aspects. First, it induces recombination specifically in the psm instead of effecting all mesodermal derivatives. However, this is only true from E7.5 to E8.5, since at E9.5, the activity is more widespread and also lateral plate and intermediate mesoderm are affected (Wehn et al. 2009). Second, the tip of the tail, comprising the Ps and later on the tailbud, does not display recombinase activity throughout axial elongation (Wehn et al. 2009), resulting in loss of the targeted gene only in cells that left the ps/tailbud.

In agreement with the reported *Dll1-msd Cre* activity, *Srf*$^{flex1/flex1}$*; Msd Cre* embryos at E11.5 lacked *Srf* transcripts in the psm and its derivatives, in the fore limb bud, which emerges from lateral plate mesoderm (reviewed in Duboc & Logan 2011), and in intermediate mesoderm. Faint *Srf* staining could be detected in the tip of the embryo (Fig.35), in accordance with the lack of recombinase activity in the tailbud.

Figure 35: *Whole mount* in situ *hybridization of* Srf$^{flex1/flex1}$; Msd Cre *embryos and heterozygous littermates at E11.5 using an* Srf*-specific probe. The right panel shows the dissected tails of the embryos shown on the left. While* the *heterozygous littermates display* Srf *expression in the tail, psm, early somites, branchial arches, forebrain, and limb bud mesenchyme, the expression in the mutants is missing in the psm and lpm, and their derivatives. However, a faint signal can be detected in the tip of the tail (red arrow).*

However, phenotypically the mutant embryos resembled wt embryos and did not display axis truncation (Fig.35), demonstrating that KO of *Srf* only affects axial elongation when the gene is already lost in cells that are in the ps or later on in the tailbud. This could indicate that Srf activity is dispensable in the psm, and that the arrested axial elongation observed in $Srf^{flex1/flex1}$; T_{streak} *Cre* mutants only results from *Srf* deficiency in the ps and the corresponding impaired migration away from the ps. However, it cannot be ruled out that axial elongation also requires Srf mediated cell migration within the psm. Since the Msd Cre recombinase is only active in cells that left the ps or the tailbud (Wehn et al. 2009) (Fig.35), there might still be sufficient Srf protein remaining in psm cells to regulate the cytoskeletal rearrangements that are required for migration. To distinguish between these possibilities, further analyses are required, including immunostaining using an Srf-specific antibody to examine for the presence of remaining Srf protein in $Srf^{flex1/flex1}$; *Msd Cre* embryos.

2.3.4. Chapter Summary

In this chapter I demonstrate that cells originating from *Srf*-deficient caudal ends display an impaired migratory capacity when plated and cultivated on a fibronectin-coated surface (Fig.31). Furthermore, they display dramatic differences in cellular morphology as compared to cells from phenotypically normal littermates (Fig.32). Control cells have a spindle shaped morphology and generate actin stress fibers, which are hallmarks of migrating cells. In contrast, cells from *Srf*-deficient caudal ends have a rounded morphology and fail to form stress fibers, but have circumferential bundles of actin fibers at the cell edge instead, which is typical for epithelial cells. *In vivo*, Brachyury-positive cells are confined to the posterior end of mutant tails, instead of spreading throughout the psm, indicating that *Srf* is required for mesodermal cell migration during axial elongation (Fig.33). In contrast to the T_{streak} *Cre*-mediated KO of *Srf*, a *Msd Cre*-mediated KO of *Srf* does not result in axis truncation (Fig.35). The observations indicate that KO of *Srf* only affects axial elongation when the gene is already lost in cells that are in the ps or in the tailbud, where the *Msd*-driven recombinase was shown to be not active (Wehn et al. 2009). This is either because Srf is dispensable in the psm or because there is still sufficient Srf protein remained in psm cells to enable cell migration.

2.4. Involvement of Srf in FGF signaling during axial elongation

Srf has been shown to be a potent mediator of FGF signaling, mainly by interaction with the FGF/MAPK regulated TCF family of Srf cofactors (Richard Treisman 1996), but also the MRTF family of Srf cofactors (reviewed in Olson and Nordheim, 2010). During the formation of the tracheal system of *D.melanogaster*, tracheal cell migration and the branching pattern of tracheae are dependent on the function of the FGF homologue *branchless* and *DSrf* (Affolter et al., 1994; Guillemin et al., 1996; Sutherland et al., 1996).

FGF signaling in the ps and in mesoderm controls EMT, cell migration, and differentiation (Ciruna & Rossant, 2001), and it is speculated that Srf represents an essential mediator of converging FGF signals during mouse gastrulation (Arsenian et al. 1998). Thus, we aimed to analyze a putative role for Srf in mediating the FGF signal in the ps and newly formed mesoderm during axial elongation. Tail halves from wild type embryos at E9.5 were treated with the chemical FGF receptor inhibitor SU5402 for four hours, and gene expression was assessed by microarray analysis. Compared to the non-treated tail halves, 105 genes were upregulated and 168 genes were downregulated (fold change ≥ $\log2^{0.4}$, $p \leq 0.05$; n=4) (Appendix, table 13). Among the downregulated genes are known FGF targets such *as Etv2, Etv5, Dusp2,* and *Dusp6,* demonstrating that FGF signaling was effected. Moreover, Ingenuity Systems Pathway Analysis revealed that the FGF regulated ERK/MAPK pathway is the top canonical pathway effected by the treatment (p-value=$2.5x10^{-6}$).

Comparison of the FGF-dependent genes with the expression data from *Srf*-deficient caudal ends reveals only a minor overlap. This is especially true for genes that are affected by *Srf* deficiency at E8.5 and E8.75. Only one (*Pdlim7*) at E8.5 or five genes (*Pdlim7, Egr1, Ier3, Neu1, Myo6*) at E8.75 were also found to be deregulated in the absence of FGF signaling (Fig.36, left; Table 4). A comparison with the genes that are deregulated *in* $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos at E9.0, which are at a similar stage as the embryos treated with the FGF inhibitor, reveals an overlap of 19 genes (Fig.36, right; Table 4). The only genes that were shown to be directly regulated by Srf and that are dysregulated in SU5402 treated embryo tails are *Egr1* and *Pdlim7,* suggesting that Srf does not primarily act as a mediator of FGF signaling in the ps and newly formed mesoderm.

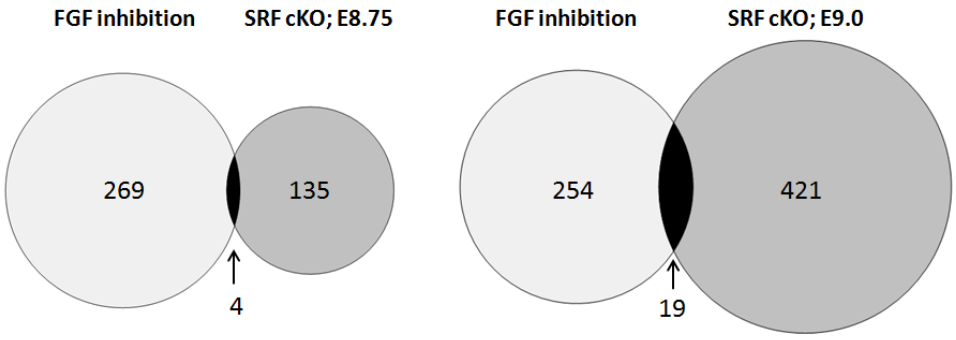

Figure 36: *Comparison of genes that are differentially expressed in caudal ends treated with te FGF inhibitor SU5402 and genes that are dysregulated in Srf-deficient caudal ends, either at E8.75 (left) or at E9.0 (right) reveals only a small overlap (fold change ≥ log2 $^{0.4}$; p ≤ 0.05; n=4)*

Table 4: *Genes that are dysregulated in embryonic caudal ends upon FGF inhibition as well as on Srf abrogation. For easier comparison, downregulation is indicated in blue and upregulation in red.*

Gene	***Gene ID***	***Dysregulation upon FGF inhibition***	***Dysregulation upon Srf abrogation (at embryonic day)***
Pdlim7	*ENSMUSG00000021493*	*down*	*down (8.5; 8.75, 9.0)*
Egr1	*ENSMUSG00000038418*	*down*	*up (8.75; 9.0)*
Ier3	*ENSMUSG00000003541*	*down*	*up (8.75; 9.0)*
Neu1	*ENSMUSG00000007038*	*down*	*up (8.75; 9.0)*
Myo6	*ENSMUSG00000033577*	*up*	*up (8.75)*
Cdr2	*ENSMUSG00000030878*	*down*	*down (9.0)*
Rbm38	*ENSMUSG00000027510*	*down*	*down (9.0)*
Mat2a	*ENSMUSG00000053907*	*down*	*down (9.0)*
Nedd9	*ENSMUSG00000021365*	*down*	*down (9.0)*
LOC100047261	*ENSMUSG00000025283*	*down*	*down (9.0)*
Flrt3	*ENSMUSG00000051379*	*down*	*down (9.0)*
LOC100048721	*ENSMUSG00000051379*	*down*	*down (9.0)*
Nptx2	*ENSMUSG00000059991*	*down*	*down (9.0)*
Pop5	*ENSMUSG00000060152*	*down*	*down (9.0)*
Zmiz1	*ENSMUSG00000007817*	*up*	*down (9.0)*
LOC100046690	*ENSMUSG00000031543*	*up*	*down (9.0)*
Ank1	*ENSMUSG00000031543*	*up*	*down (9.0)*
Dusp4	*ENSMUSG00000031530*	*down*	*up (9.0)*
Vegfa	*ENSMUSG00000023951*	*down*	*up (9.0)*
Bbc3	*ENSMUSG00000002083*	*down*	*up (9.0)*
Jmjd1a	*ENSMUSG00000053470*	*down*	*up (9.0)*

2.5. Embryos with a mesoderm-specific knock out of Snai1 *do not display axis truncation*

Members of the Snail gene family are central regulators of the epithelial to mesenchymal transition, the process when cells from a polarized, adhesive epithelium transform into motile, mesenchymal cells. The gene *Snai1* encodes a transcription factor that has been shown to induce EMT by binding to specific binding sites, called E-boxes, and repressing the transcription of epithelial genes, such as E-cadherin (*Cdh1),* Occludin *(Ocln),* and Claudins *(Cldn)* (Cano et al. 2000; Ikenouchi et al. 2003). During embryogenesis, *Snai1* is expressed in the ps and nascent mesoderm at E7.5, and its transcription has been shown to be induced by FGF signaling (Ciruna & Rossant 2001). *Snai1* KO embryos display gastrulation defects including ectopic expression of E-cadherin in newly formed mesoderm, and die at approximately E8.0 (Carver et al. 2001). In the psm, *Snai1* expression occurs in a cyclic manner (Fig.37,A). This expression pattern correlates with an additional function for *Snai1* in the coordinated segmental patterning of the psm (Dale et al. 2006).

To analyze whether *Snai1* is essential during somitogenesis and/or axial elongation, I generated embryos with a conditional KO of *Snai1* in the ps and nascent mesoderm. Gridley and colleagues previously generated a mouse line with floxed *Snai1* (*Snai1*$^{flox/flox}$) (Murray et al. 2006) and demonstrated that deletion of the *Snai1*^{flox} allele by the *Meox2* promoter-driven recombinase (Tallquist & Soriano 2000), which is highly active in the germ cell line, results in embryos with a *Snai1*-null allele. I crossed these *Snai1*$^{flox/flox}$ embryos with the T_{streak} *Cre* line to obtain *Snai1*$^{flox/flox}$*;* T_{streak} *Cre* embryos. No *Snai1* transcripts could be detected in mesodermal tissues of *Snai1*$^{flox/flox}$*;* T_{streak} *Cre* embryos, as examined by *in situ* hybridization using a *Snai1*-specific probe (Fig.37,B). However, the lack of *Snai1* did not result in any abnormalities in embryos at E9.5, E10.5 or E11.5. Moreover, one *Snai1*$^{flox/flox}$*;* T_{streak} *Cre* embryo survived to birth, indicating that *Snai1* is dispensable for the elongation of the postcranial AP-axis and patterning of the somites. These results could be due to compensation by other members of the Snail gene family (*Snai2, Snai3*) or by other factors like Zeb1, Zeb2, which also bind to E-boxes and inhibit the expression of genes such as *Cdh1* (Guaita et al. 2002).

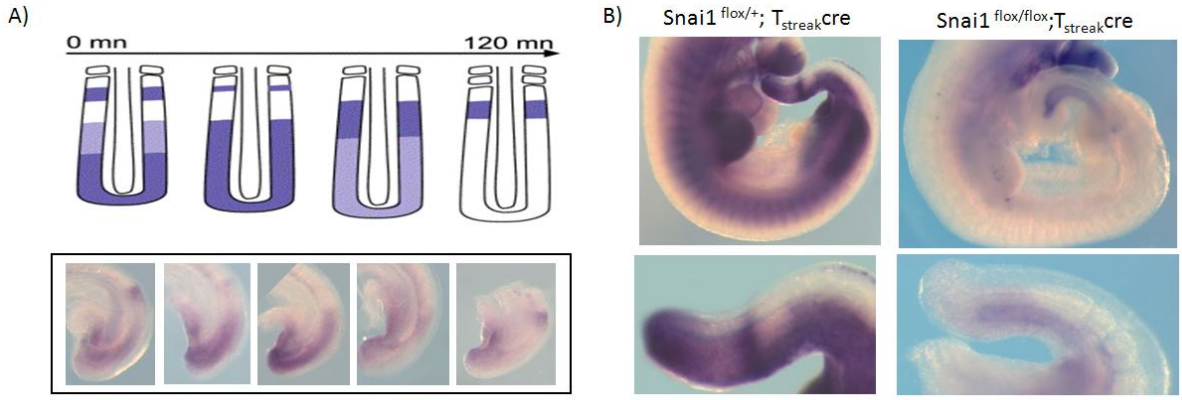

Figure 37: *A)* Snai1 *is expressed in the psm in a cyclic manner. Top: Schematic representation of the dynamics of* Snai1 *expression over the course of the formation of one new somite pair, which takes approximately 120 min* (Dale et al. 2006). *Bottom: Tails of E9.5 embryos hybridized with a* Snai1*-specific probe, exhibiting the different patterns of Snai1 expression. B)* In situ *hybridization of* $Snai1^{flox/flox}$*;* T_{streak} *Cre embryos (right) and heterozygous littermates (left) at E9.5 reveals a complete lack of* Snai1 *transcripts. However, the* Snai1*-deficient embryos do not display any abnormalities, indicating that Snai1 activity in the psm is not required for axis extension or subsequent EMT events in mesodermal derivatives.*

3. *Discussion*

3.1. Srf activity is essential for cell migration in nascent mesoderm cells

The transcription factor Srf is a versatile factor involved in various developmental processes such as cell differentiation, migration, proliferation, apoptosis, and myogenesis. Strong *Srf* expression can be detected in the primitive streak and in the presomitic mesoderm of mouse, chick, and frog embryos (Fig.10 and Mohun et al. 1991; Croissant et al. 1996), which implies that it plays a role in mesoderm development and/or axis extension. In the course of this work, I generated conditional KO embryos that lack Srf activity specifically in the ps and nascent mesoderm. These embryos display a severe axis truncation (Fig.24), demonstrating that Srf is essential for axial elongation. Since Srf is a transcription factor, the assessment of its molecular role in the ps and nascent mesoderm requires the identification of the genes regulated by Srf. However, Srf not only binds DNA but also provides docking surfaces for interaction with a wide variety of cofactors (reviewed in Majesky, 2003). Those cofactors enable differential transcriptional activity of Srf target genes, as well as integration of diverse signaling cues. This means that not all genes that are bound by Srf in a certain tissue are necessarily regulated by Srf in this tissue. Instead, only a subset of Srf bound genes is likely to be regulated, depending on the presence of the appropriate Srf cofactors and/or the required cellular signals that induce the activity of the Srf-cofactor complex. Consequently, I aimed to identify genes that are both bound and regulated by Srf, presumably in combination with a cofactor, during mesoderm development and axis extension, on a genome-wide level. Therefore, I combined Srf ChIP-Seq data with differential gene expression data from conditional *Srf* KO caudal ends. This revealed 27 genes that directly depend on Srf during the process of axis extension (Tables 2 and 3); the majority of which are implicated in the reorganization of the actin cytoskeleton and the formation of focal adhesions (Fig.28), both hallmarks of cell migration and cell adhesion. These findings indicate that the major role for Srf in the ps and nascent mesoderm is the regulation of cell motility.

Among the identified Srf-dependent genes are *Actb* and *Acta1*, both of which encode actin monomers, which are the building blocks of the actin cytoskeleton. The *de novo* synthesis of actin monomers has been shown to be essential for cell motility (Kislauskis et al. 1997). In addition, proper cell migration requires the flexible reorganization of the actin cytoskeleton, depending on both

stabilization of actin stress fibers and disassembly of actin filaments. Accordingly, Srf directly activates genes encoding proteins that stabilize actin filaments, such as members of the Tropomyosin family, as well as proteins implicated in the disassembly of F-actin, such as Destrin. Additionally, Srf activates other genes that are implicated in the stability and composition of F-actin, thereby regulating cell migration, such as Calponin 2 (*Cnn2*) and Fascin homolog1 (*Fscn1*) (Huang et al. 2008; Sedeh et al. 2010). Cell migration also requires the formation of focal adhesion (FA) complexes through which the cytoskeleton of a cell connects to the extracellular matrix. Downregulation of the Srf-regulated FA complex components Zyxin and Vinculin indicates that FA complex formation might be also affected in *Srf*-deficient mesodermal cells. In line with this assumption, the novel putative Srf target gene v-crk sarcoma virus CT10 oncogene homolog (avian) *(Crk)* encodes a protein that is involved in FA formation and in remodeling of the cytoskeleton at focal adhesions (Feller 2001; Antoku & Mayer 2009).

An essential role for Srf in the regulation of cell migration has been recently demonstrated both *in vitro* and *in vivo* (Schratt, Philippar, et al. 2002; Alberti et al. 2005) (see 1.2.2.). However, both studies focused on the analysis of only a few transcriptional targets of Srf, and correlated their misexpression with the observed migratory defects. In agreement with the data provided in this work, Schratt, Alberti and their coworkers found that *Srf* deficiency resulted in the downregulation of *Actb*. Additionally, Schratt and colleagues demonstrate that $Srf^{-/-}$ ES cells failed to form focal adhesion complexes, caused by a lack of the proteins Talin, Zyxin, and Vinculin, strengthening the assumption that the downregulation of *Zyx* and *Vcl* in *Srf*-deficient mesoderm cells correlates with a failure in FA formation. Further analysis of this, e.g. by immunostaining of *Srf*-deficient caudal end explants with antibodies specific for FA proteins such as Paxilin, will allow for the assessment of focal adhesion formation on the protein level .

Further genes that were identified as direct Srf target genes during axis extension and that are associated with the regulation of cell migration, are *Fhl2, Pdlim7, Fhl1* and *Anxa2.* However, they encode proteins that not only bind to the actin cytoskeleton, but are also implicated in the regulation of other processes such as gene transcription and intracellular signaling. Their potential functions in the ps and nascent mesoderm are discussed below.

To analyze, whether *Srf* deficiency affects the migratory capacity of mesodermal cells, as implicated by the molecular data, I assessed cell migration of $Srf^{-/-}$ mesodermal cells in an e*x vivo* migration assay. This assay revealed that Srf is indeed essential for cell motility in embryonic caudal

ends (Fig.31). The migration defect clearly correlates with a reduced overall actin-level and a completely different cytoskeletal organization (Fig.32), both also in agreement with the molecular data. Cells from control samples generated stress fibers that traversed the whole cell body, which is crucial for cell migration. In sharp contrast, cells from *Srf*-deficient caudal ends failed to form any stress fibers. Instead, they displayed convex, circumferential bundles of actin fibers at the cell surface, which is characteristic for epithelial cells (Gloushankova et al. 1997) (Fig.32,B). Hence, lack of *Srf* results in impaired cell motility due to a failure of cytoskeletal rearrangement. Both the deregulation of migration-associated genes and the profound migratory defect seen e*x vivo* imply that *Srf*-deficient mesodermal cells also have a migration defect *in vivo*. Although this is difficult to directly prove, the accumulation of Brachyury positive cells at the posterior end of $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos supports this assumption (Fig.33). In 9.5 day-old wild type embryos, new Brachyury-positive mesodermal cells emerge from progenitor cells that are located in the tailbud. These cells undergo cell movements, thereby moving away from the tailbud and spreading throughout the psm, which is essential for axial elongation (Bénazéraf et al. 2010). In contrast, Brachyury-positive cells in *Srf*-deficient embryo tails accumulate at the posterior end of the embryo, pointing to migration defect *in vivo*.

3.2. Axial elongation requires migration of mesodermal cells

In recent years, several studies have demonstrated a connection between cell motility and elongation of the A-P axis. For example, embryos homozygous null for *Fgfr1* or *Snai1* die during gastrulation and have an accumulation of cells at the ps, at the expense of paraxial mesoderm. Both mutants display an incomplete EMT due to impaired downregulation of *Cdh1*, which results in the retention of adherens junctions and defective detachment from neighboring cells (Yamaguchi et al. 1994; Carver et al. 2001). The defects in both mutants are already visible at the onset of gastrulation (E6.5), where EMT is crucial for the detachment of epithelial cells in the ps. In contrast, the T_{streak} Cre-mediated loss of *Srf* leads to an observable phenotype at E8.75, when most of the newly generated mesoderm arises from mesodermal progenitors in the tailbud. It is not clear whether these cells also undergo EMT. The fact that the T_{streak} Cre-mediated KO of *Snai1* does not result in axis truncation (see 2.5.) might indicate that, at this stage, EMT is less relevant, since *Snai1* was shown to be essential for earlier gastrulation movements (Carver et al. 2001). Alternatively, the lack of *Snai1* could be compensated for by other factors, which can also repress *Cdh1* expression (see 2.5.). Nevertheless,

cell motility remains crucial for axial elongation, also at the stages when mesodermal cells are not derived from the epiblast but from mesodermal progenitors (Bénazéraf et al. 2010; Girós et al. 2011). Recent experiments in chick show that cells which have entered the psm resume migration, and that this migration is essential for axis extension (Bénazéraf et al. 2010). Consequently, embryos that were treated with motility inhibitors show a slower axis extension. In particular, Bénazéraf and coworkers used inhibitors that either interfered with Rho signaling or with stress fiber formation. As mentioned earlier, Rho signaling regulates actin filament assembly and focal adhesion complex formation (see 1.2.1. and Fig.7), therefore treatment with these cell motility inhibitors affects the same functional structures as does the lack of *Srf* in the ps and nascent mesoderm. This indicates that the lack of cell motility caused by *Srf*-deficiency directly correlates with the observed axis truncation.

Studies have also demonstrated a correlation between mesodermal cell migration and axis extension in mouse embryos. Some of these studies focus on mutations that disrupt the extracellular matrix protein fibronectin, or affect members of the integrin family, which connect the FA with fibronectin. These factors are not directly regulated by Srf (as shown in this work), nor are they part of the actin cytoskeleton, but mutations that interfere with the functionality of these proteins disrupt the same migratory machinery as does the abrogation of *Srf* in embryonic caudal ends. As discussed, Srf-dependent stress fiber assembly and focal adhesion formation are essential for the cell to generate a force against the ECM, which is fundamental for cell movement (McHardy et al. 2005) (see 2.1.3.4.). The same process requires the proper connection of the FA to the extracellular matrix, which is impaired by mutations in fibronectin or integrins. Consequently, embryos deficient for fibronectin display impaired mesodermal cell migration and die around E8.5, with a number of mesodermal defects such as a shortened anterior-posterior axis, a disorganized notochord and the absence of somites (George et al. 1993; Georges-Labouesse et al. 1996). A similar phenotype is seen when the two major integrins that are expressed during gastrulation and axis extension, *integrin α5β1 and αv-containing integrin*, are removed (Yang et al. 1999). Together, these studies demonstrate the importance of focal adhesion connection to the ECM for cell migration and axis extension. Similarly to the KOs for *Srf*, *Fgfr1,* and *Snai1*, the KO of *fibronectin* or *integrins* affects the embryos much earlier than seen here for the conditional loss of *Srf* in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos, making a broad comparison of these phenotypes difficult. However, a very recently reported mutation in the RGD motif in fibronectin ($Fn^{RGE/RGE}$ embryos) results in impaired axial elongation, which is observable at the same stage as the axis truncation in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos (Girós et al. 2011). This mutation specifically blocks the binding of α5β1 integrin to fibronectin, and the

mutant embryos exhibit arrested segmentation and impaired axis extension by E9.0. This late onset can be explained by the fact that earlier in axis extension, the ingression of epiblast cells, and the subsequent migration during gastrulation are all mediated by αv integrins. The progenitor-derived mesoderm cells express α5β1 integrin from around E9.0. Therefore, only their migration is severely impaired in $Fn^{RGE/RGE}$ embryos. Interestingly, $Fn^{RGE/RGE}$ embryos phenocopy the $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos in many aspects. They display a shortened posterior trunk and fail to complete the turning process. $Fn^{RGE/RGE}$ embryos develop the first 11–15 somite pairs. This is similar to $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos, where the maximal number of somite pairs was 16. Strikingly, both mutants display the same heart defects (Fig.S.1) and die between E9.5 and E10.5, probably due to the heart defect. Several publications have shown the importance of Srf in heart development, and specifically in differentiation of cardiomyocytes, but these studies describe conditional *Srf* KO embryos which lost Srf activity later than the $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos (see for example Parlakian et al. 2004; Miano et al. 2004). The similar malformation in $Fn^{RGE/RGE}$ embryos suggests that the heart malformation observed in the *Srf* mutants is owing to impaired migration rather than to a defect in differentiation. It is conceivable that heart progenitors, which are among the first mesodermal cells generated during gastrulation in the caudal end of the embryo (Kinder et al. 1999; Parameswaran & Tam 1995) require Srf activity and also α5β1 integrin activity when they migrate to the midline to form the heart tube.

Girós and colleagues demonstrated that the impaired binding of α5β1 integrin to fibronectin not only results in defective cell migration, but also affects cell-cell communication (Girós et al. 2011). In the psm, cell-cell communication is critical for synchronized segmentation and the subsequent formation of somites, and failure of this synchronization can result in arrested somitogenesis. The examination of somitogenesis in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos was not part of this work, but it is tempting to speculate that it is also affected. It would thus be interesting to analyze the synchrony of the expression of genes such as *Axin2,* Lunatic fringe *(Lfng),* and *Hes7,* which are representative cyclic genes of the somitogenic clock regulated by Wnt, Notch, and FGF signaling, respectively, in conditional *Srf* mutants.

Axis extension not only depends on cell migration but also on the continuous generation of new mesodermal cells by the posterior growth zone. Whether there is a mechanism that correlates axis extension and the activity of the growth zone is not clear. Such a mechanism could inhibit the formation of new tissue in the growth zone in the case of arrested axis extension caused by impaired mesodermal cell migration. Recently it has been demonstrated how, in general, the activity of mesodermal progenitors in the growth zone is correlated to axis length, in particular to the length of

the psm. Briefly, the vitamin A derivative retinoic acid (RA), which can inhibit the activity of the posterior growth zone, is generated in somites and in the anterior psm. It exhibits a graded distribution along the A-P axis, with decreasing amounts towards the posterior end of the tail. In the posterior growth zone, RA is actively degraded by the enzyme Cyp26a1 (Sakai et al. 2001). Either administration or high amounts of vitamin A or KO of *Cyp26a1* results in embryos with RA activity that extends in the progenitor zone (Abu-abed et al. 2001). This causes loss of *Fgf8* expression, which is believed to be crucial for the maintenance of the progenitor population (Diez et al. 2003). Consequently, loss of FGF8 activity leads to a premature termination of axis extension and hence to body truncation. Such a mechanism would generally link axis length to mesoderm progenitor activity since a shortened psm, for example caused by impaired migration, would result in an increased RA concentration in the posterior end of the embryo and the premature differentiation of progenitors.

However, in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos, neither *Fgf8*, *Cyp26a1* nor the RA synthesizing enzyme aldehyde dehydrogenase family 1, subfamily A2 *(Aldh1a2)* were deregulated, suggesting that the RA-FGF8 machinery is not affected by the impaired axis extension at this stage. Very recently, Cunningham and coworkers demonstrated that the RA-dependent mechanism for axis extension only applies in embryos before E9.5, but does not play a role during later axis extension (Cunningham et al. 2011). Consequently, RA is not required for the natural occurring termination of axis extension. The mechanism regarding how the progenitor zone in the tailbud stops generating new mesoderm at a particular point in development, therefore determining the final segment number, is not known. Cunningham and coworkers speculate that there might be a mechanism which is similar to the one described above, which inhibits the generation of new mesodermal tissue and constitutes the end of axial elongation. It is possible that such a mechanism can also control the activity of mesodermal progenitors in embryos that lack axis extension due to impaired mesodermal cell migration, such as $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos or $Fn^{RGE/RGE}$ embryos.

3.3. Srf cofactors during axis extension

Srf functions as a platform for various cofactors, allowing for the differential activation of target genes. Consequently, Srf can be bound to the regulatory region of a certain gene, without regulating the gene in that specific cellular context. Only upon binding of a specific cofactor to the Srf-DNA complex, will the target gene be activated or, in rare cases, repressed. This offers an explanation as

to why there are many more genes bound by Srf, as was identified by chromatin immunoprecipitation (see 2.1.), than are actually deregulated in *Srf*-deficient caudal ends. The identification of the cofactors, which regulate Srf activity during axis extension in mouse embryos was not part of this work. However, based on the genes identified to be Srf-dependent in the ps and in nascent mesoderm, one can speculate about the nature of the involved cofactors. By the year 2008 56 Srf cofactors had been identified, most of them expressed in a tissue-specific manner (Miano 2008). Five of them are rather ubiquitously expressed, and their activity is regulated by extracellular signals. Those cofactors can be subdivided into two families, termed the ternary complex factors (TCFs) and the myocardin related transcription factors (MRTFs) (1.2.1.; Fig.6 and 7). During gastrulation and axis extension, cells are transiently exposed to various signals. For example, members of the TGFβ, Wnt, and FGF signaling pathways induce cell fate decisions and regulate morphogenetic changes during the EMT process and during cell migration. As demonstrated in the current work, Srf activity is essential for conducting reorganization of the actin cytoskeleton, which is one of these morphogenetic changes. Therefore, it can be assumed that cofactors that regulate Srf activity during axis extension are likely to be regulated by extracellular signals, and thus belong to the TCF or MRTF family of Srf cofactors.

A previous report demonstrated that a subset of Srf target genes is specifically activated either by members of the TCF family or by MRTF cofactors (Selvaraj & Prywes 2004). The TCF factors, which are regulated by Ras-MAPK-ERK signaling, activate, in combination with Srf, the expression of growth related genes such as *Egr1* and *Fos.* However*, Egr1* and *Fos* expression have been shown to be independent of MRTF factors (Miralles et al. 2003). On the other hand, genes such as *Vcl*, *Zyx*, *Tpm1*, *Acta1,* and *Fhl2*, all associated with cytoskeletal dynamics, have been shown to be MRTF-dependent but TCF-independent (Selvaraj & Prywes 2004). Most of the genes that were identified to be Srf-dependent during axis extension belong to the MRTF-dependent subset of Srf target genes, implying that members of the MRTF family extensively contribute to the regulation of Srf activity in this process. *MRTF-A* and *MRTF-B* are thought to have redundant as well as specific functions. *MRTF-A* KO mice are viable, but exhibit a failure in the development of mammary myoepithelial cells, while *MRTF-B* KO mice die around E13.5 from a spectrum of cardiovascular defects, including abnormal patterning of the branchial arch artery, ventricular septal defects, and thin-walled myocardium (Li et al. 2006; Oh et al. 2005). However, both genes are implicated in the general link between actin cytoskeleton dynamics and correlated gene transcription activities (reviewed in Olson & Nordheim 2010). A double KO of both genes has not yet been published, and it would be interesting to see

whether a mesoderm-specific KO of *MRTF-A* and *MRTF-B* resembles the phenotype of $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos, thus strengthening the Srf/MRTF connection.

MRTF factors are not implicated in the repression of genes, indicating that they are probably not accountable for the observed upregulation of Srf-dependent genes in *Srf*-deficient caudal ends (table 3). This and the fact that *Egr1*, which is one of the genes found here to be upregulated upon *Srf*-deficiency, has been shown not to be bound by the Srf-MRTF complex (Miralles et al. 2003), suggests that additional Srf cofactors are also involved in the regulation of Srf activity during axis extension. For example, the participation of members of the TCF family is possible. In particular, the TCF factor Elk3 could be involved in the inhibition of *Egr1*, since it possesses a unique inhibitory domain and has been shown to inhibit gene expression in mouse embryonic fibroblasts (Maira et al. 1996). Other reports demonstrate that the Srf cofactor Fhl2 can inhibit gene transcription, and its specific expression in the posterior tip of the embryo (Fig.26,C) suggests that it plays a role in the regulation of Srf activity. However, Fhl2 seems to inhibit gene transcription by an indirect mechanism (discussed in 3.6.). Taken together, these observations suggest that the MRTF proteins are the main factors regulating Srf activity during axial elongation. However, whether the MRTF proteins are the only cofactors in this process or if there are additional cofactors involved that also regulate a subset of Srf target genes remains elusive.

3.4. Additional roles for Srf*-dependent genes during axis extension*

Most of the genes that are Srf-dependent during axis extension encode proteins that are associated with the regulation of cell motility. However, some Srf regulated factors are implicated in additional cellular processes (e.g. *Crk, Pdlim7, Fhl2*, and *Rasa1*), or they are not related to cell migration at all (*Egr1*, nucleolar complex associated 3 homolog (*Noc3l*)). Some of these additional functions could be relevant during axis extension, but it is dubious whether genes such as *Noc3l*, which is implicated in fat cell differentiation, can have any significance in embryonic caudal ends.

Egr1, *Crk* and *Fhl2* are all implicated in the control of cell proliferation (Nishihara et al. 2002; Huang et al. 1997; Johannessen et al. 2006), and analysis using the proliferation marker phosphoH3(S10) revealed a slight decrease in cell proliferation in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryo tails (Fig.34). While this indicates that control of proliferation might be a further role for Srf in embryonic caudal ends, the fact that no cell cycle-specific genes are deregulated in the gene expression data of

Srf-deficient tails argues against a direct effect on cell proliferation. However, further experiments that assess the proliferation rate in conditional *Srf* mutants prior to the onset of the axis truncation could clarify this.

Besides having a role in proliferation, Fhl2 has also been shown to act as a mediator between the cytoskeleton and the nucleus, integrating Rho signaling and cytoskeletal dynamics into altered gene expression and cell morphology. Like Zyxin and Vinculin, it is found in focal adhesion complexes, where it binds to integrins (Wixler et al. 2000). Stimulation of the Rho signaling pathway induces its translocation from FA complexes into the nucleus, where it acts as a cofactor for various transcription factors, including Srf and β-catenin, the key mediator of Wnt signaling (Mueller et al. 2002; Wei et al. 2003). As a cofactor of Srf, Fhl2 negatively regulates MRTF-Srf-dependent gene transcription, probably by competing for the same binding site on Srf (Philippar et al. 2004). On the other hand, it positively regulates MRTF protein stability (Hinson et al. 2009). It also inhibits the second family of Srf cofactors, TCF, by interfering with the MAPK signaling pathway (Purcell et al. 2004). In cardiomyocytes, Fhl2 binds to and sequesters activated extracellular signal-regulated kinase2 (ERK2) in the cytoplasm. ERK2 is a component of the MAPK signaling pathway, which phosphorylates and activates the TCF cofactors. Wixler and colleagues also showed that Fhl2 can activate Srf in fibroblasts, illustrating its versatile functionality (Wixler et al. 2007).

The gene *Rasa1* was identified here as a novel Srf target gene in embryo caudal ends. It encodes the protein RasGAP, which is thought to connect the Ras and Rho pathways (Leblanc et al. 1998). On the one hand, it is a well known inhibitor of Ras activity by stimulating the intrinsic Ras GTPase activity, thereby promoting the conversion of Ras from an active, GTP-bound to an inactive, GDP-bound state (Zhang et al. 1990). On the other hand, it was also shown to be necessary for endogenous Rho activation and the triggering of actin stress fiber formation in differentiated PC12 cells and in fibroblasts (Leblanc et al. 1998). Since the Ras and Rho pathways are regulators of the TCF and MRTF families of Srf cofactors, *Rasa1* could represent an interesting factor in the general regulation of Srf cofactor activation and therefore the differential induction of Srf target genes.

The gene *Pdlim7* is a likely Srf target gene, although it was not found to be enriched in the ChIP experiments (see 2.2.4.). It encodes a protein that interacts with T-box transcription factors and regulates their shuttling between the nucleus and cytoplasm (Camarata et al. 2006). In the zebrafish heart, Pdlim7 binds to the actin cytoskeleton and to Tbx5, thereby sequestering Tbx5 in the cytoplasm and inhibiting transcription of Tbx5-dependent target genes (Camarata, 2006). Tbx5 is

shuttled out of the nucleus by the CRM1 export protein that binds to a nuclear export signal (NES) in Tbx5. Interestingly, this site is conserved in all T-box proteins, and an interaction of CRM1 was also demonstrated for Tbx2 and Brachyury (Kulisz & Simon 2008). Moreover, Camarata and colleagues speculate that the interaction of Pdlim proteins with T-box factors and their sequestering in the cytoplasm could be a general mechanism (Camarata et al. 2010). During axis extension, the T-box transcription factors T and Tbx6 are both expressed in the ps and nascent mesoderm. Putative regulation of their localization by Pdlim7 could present a new mechanism for the regulation of T and Tbx6 downstream targets, and possibly a link to cyclic gene expression in the psm during the process of somitogenesis (see 1.1.2.). In this respect, it is interesting to note that cytoplasmic localization of Tbx6 was detected in the caudal end of one Srf mutant embryo (Fig.S.4). However, in other Srf-deficient caudal ends, Tbx6 was found in the nucleus, implying that although localization of Tbx6 is not directly correlated with *Srf* deficiency, it may be a common event during axis extension. It would be interesting to examine whether cytoplasmic localization of Tbx6 can be observed in additional *Srf*-deficient and/or wild type embryos, and whether this can be correlated with cyclic gene expression and/or Pdlim7 activity. Also a closer analysis of the Brachyury localization in psm cells would be interesting, since its cytoplasmic localization was demonstrated for several other tissues (Inman & Downs 2006).

The versatile functions of genes such as *Pdlim7, Fhl2,* and *Rasa1* demonstrate that Srf may not exclusively control components of the migratory machinery during the process of axis extension. Instead, it is conceivable that Srf is implicated in the network connecting signaling pathways, cytoskeletal dynamics, and gene expression. However, the main role of Srf is it to enable cell migration and accordingly, the impaired cell migration is the cause of the observed axis truncation in the absence of *Srf*.

3.5. Connecting FGF signals and Srf activity during axis extension

The Wnt, TGFβ and FGF signaling pathways have central roles during gastrulation and axis extension. While the transcription factors that mediate Wnt and TGFβ at the transcriptional level are well known (Weinstein et al. 2000; Huelsken et al. 2000), the mediator(s) of the FGF signal during gastrulation and axis extension have not been identified. Srf has been shown to be a potent mediator of FGF signaling in numerous processes, mainly by interaction with FGF-regulated cofactors. It was

speculated that Srf mediates FGF signaling during gastrulation (Arsenian et al. 1998). However, this does not seem to be the case, since the overlap between genes that are dependent on FGF signaling and Srf activity, as assessed in this work, is negligible (Fig.36).

In our experimental setup, Srf activity is only lost once cells have activated the Brachyury promoter, which is an initial step in mesoderm specification. Therefore, the gene expression data from *Srf$^{flex1/flex1}$; T_{streak} Cre* embryos are not suitable for drawing conclusions about Srf function in the early specification of mesoderm cells, and a direct correlation between Srf and FGF signaling in processes such as the induction of mesoderm formation cannot be ruled out. FGF signaling is also thought to be able to regulate cell motility throughout gastrulation and axis extension (Bénazéraf et al. 2010). Due to the lack of common target genes, it can be speculated that FGF regulates components of the migratory machinery that are independent of Srf activity. In this context, it is noteworthy that the gene nerve growth factor receptor (TNFR superfamiliy, member 16) *(Ngfr)* (also known as *p75NTR*), which was identified to be downregulated upon FGF inhibition (see Appendix Table 13) but independent of Srf activity, is implicated in the regulation of cell motility in cells that successfully have detached from the epiblast. More precisely, its *Xenopus* homologue *NRH,* which is also regulated by FGF (Nagamune & Ueno 2004), was shown to be essential for gastrulation movements by inducing the formation of cell protrusions such as filopodia, and gain and loss of *NRH* function results in axis truncation (Chung et al. 2005). Hence, FGF-dependent expression of *NRH* is thought of as a mechanism for how FGF regulates cell migration during axis extension. A similar function for the mouse homologue *Ngfr* has not yet been demonstrated, but would explain how FGF signaling could regulate cell migration without affecting the same set of downstream targets as Srf.

3.6. Egr1 *upregulation in* Srf$^{flex1/flex1}$; T_{streak} Cre *embryos*

As discussed earlier, *Egr1* is regulated by the Ras-MAPK-ERK-TCF/Srf signaling cascade, while Rho-MRTF/Srf signaling does not affect *Egr1* expression (Alexandre et al 1991; Selvaraj & Prywes 2004; Gineitis & Richard Treisman 2001). *Egr1* is one of the best studied Srf target genes, and most studies that examined the Srf-*Egr1* interaction demonstrated that Srf induces *Egr1* transcription. Contrary to those studies, *Egr1* was upregulated in *Srf*-deficient caudal ends at E8.75 and E9.0, suggesting that its transcription is usually inhibited in the presence of Srf. Notably, another study also reports on upregulation of *Egr1* as a result of a conditional KO of *Srf*, in this case in macrophages

(Sullivan et al. 2011). The authors explain this by the fact that Srf has been shown to be able to repress transcriptional activity, probably by sequestering a co-activator (Ernst et al. 1995). However, whether *Egr1* inhibition occurs via an indirect mechanism, or Srf inhibits *Egr1* directly, maybe in combination with another cofactor, is not known.

There are several conceivable mechanisms of how Srf could mediate indirect repression of target genes in embryonic caudal ends. Srf can mediate the repression of genes by activating the transcription of microRNAs which, in turn, inhibit the transcription or translation of target genes (Ivey et al. 2008; Yong Zhao et al. 2005). This mechanism could also apply to *Egr1* in the ps and nascent mesoderm. Another mechanism could involve long non-coding RNAs, which were shown to be able to enhance the transcription of neighboring protein coding genes (Ørom et al. 2010). In the course of this work, the analysis of the Srf ChIP-Seq data from embryonic caudal ends and P19 cells lead to the identification of a putative long non-coding RNA downstream of the *Egr1* gene, which, in the absence of *Srf*, is also upregulated (see 2.2.5. and Fig.30). Whether the presence of this putative non-coding RNA is relevant for the regulation of *Egr1* remains to be tested.

An alternative explanation for the upregulation of *Egr1* in *Srf*-deficient caudal ends is linked to the activity of the MAPK signaling pathway. As previously stated, *Egr1* is a high affinity target for the Srf-TCF complex, which is activated by the Ras-MAPK-ERK pathway. Upon activation of the pathway by extracellular signals, activated ERK translocates to the nucleus and phosphorylates TCF factors. In agreement with this, inhibition of FGF signaling with chemical compounds results in downregulation of *Egr1* (see 2.4.; Table 4). Among the Srf-dependent targets in the caudal end are two genes, *Fhl2* and *Rasa1*, which encode negative regulators of MAPK signaling and their deregulation could be causative for temporary upregulation of *Egr1*.

Fhl2 sequesters activated extracellular signal-regulated kinase2 (ERK2) in the cytoplasm (Purcell et al. 2004), and thus inhibits the phosphorylation and activation of TCF factors by ERK2. The *Fhl2* gene is downregulated in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos at E8.75 (Fig.26,C), suggesting that the inhibition of MAPK signaling by Fhl2 might be lost in these embryos. ChIP analysis revealed that *Fhl2* is bound much weaker by Srf than is *Egr1* (Fig.S.3). This is most probably because there is only one Srf binding site in the *Fhl2* promoter, while there are six binding sites in the *Egr1* promoter. It is likely that in conditional *Srf* KO embryos, the Srf protein, which, in fibroblasts, has a half life of greater than 12 hrs (Misra et al. 1991), is only gradually degraded. Thus, there may be a state in which the remaining Srf molecules are more likely to bind to strong Srf targets such as *Egr1* instead of binding

to weaker target genes such as *Fhl2*. Consequently, in *Srf*$^{flex1/flex1}$*; T*$_{streak}$ *Cre* embryos, Srf would still bind to the *Egr1* promoter, while transcription of weaker Srf target genes is already diminished and the encoded proteins are degraded. Without Fhl2 sequestering, activated ERK molecules would translocate to the nucleus and phosphorylate TCF molecules such as Elk1 which, in combination with the remaining Srf molecules, would increase *Egr1* transcription (Fig.38).

A similar scenario can be discussed for the Srf-dependent gene *Rasa1* which is, like *Fhl2*, only weakly enriched in the Srf ChIP-Seq, again suggesting that Srf binding occurs with a relatively low affinity. Since Rasa1 is an inhibitor of Ras signaling, it is upstream of the MAPK-ERK signaling pathway. Downregulation of *Rasa1* could result in enhanced MAPK signaling and, similarly as for Fhl2, the remaining Srf would be bound to the strong Srf target *Egr1* to induce *Egr1* transcription.

In agreement with both hypotheses, the MAPK regulated targets *Ier3* (E8.75, E9.0) and *Dusp4* (E9.0) are also found to be upregulated in *Srf*$^{flex1/flex1}$*; T*$_{streak}$ *Cre embryos*. On the other hand, many known MAPK regulated genes are not differentially expressed between *Srf*-deficient and wild type tails. A putative affect on MAPK activity could be assayed by immunostaining with a phospho-Elk1-specific antibody. If the hypothesis holds true one would expect to see an increase in phosphorylated Elk in response to loss of Srf. Preliminary experiments on paraffin sections from conditional *Srf* KO and control embryos failed in this regard due to non-specificity of the antibody. An alternative approach would be a conventional immunoblot using protein lysate from *Srf*-deficient and control caudal ends, which was not attempted here. However, besides *Egr1*, *Fhl1* and *Anxa2* are also direct Srf target genes that are upregulated in *Srf*-deficient caudal ends, and they are only weakly enriched in Srf ChIP. This indicates that, unlike *Egr1*, they are not bound by Srf with a high affinity and that the mechanism discussed here cannot explain their upregulation.

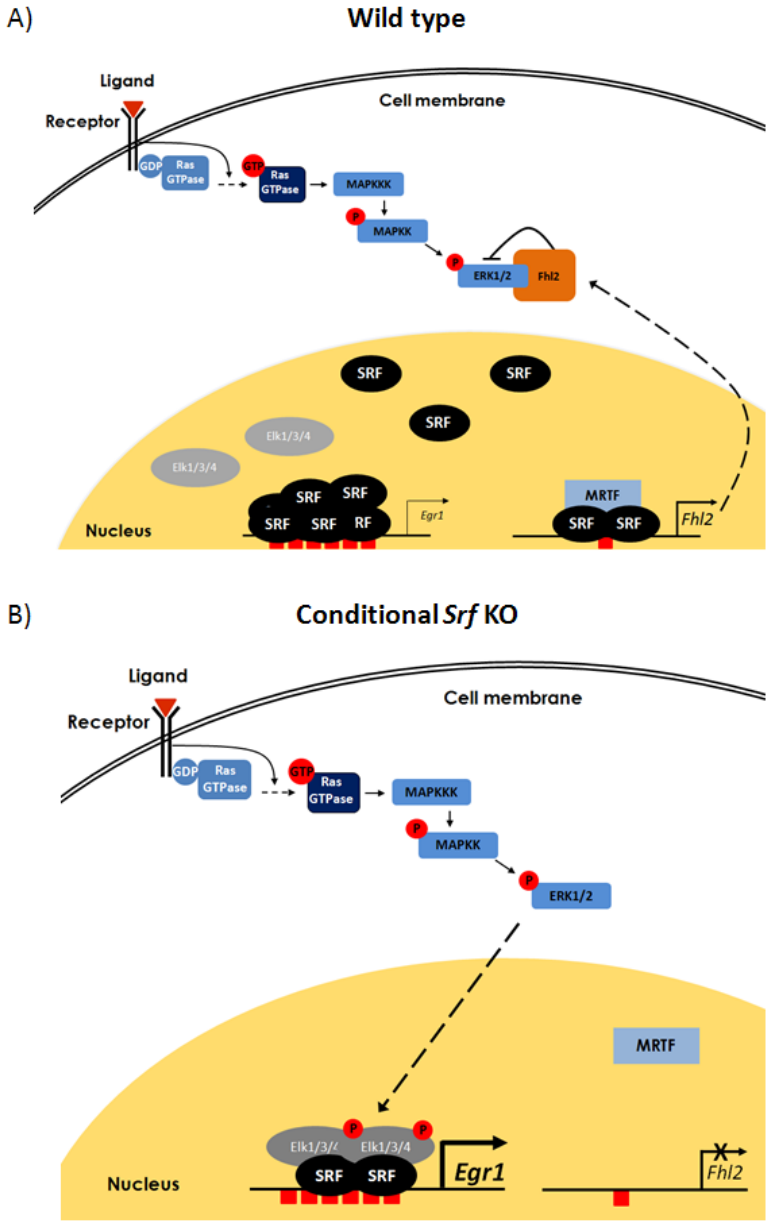

Figure 38: *Model for an indirect involvement of Fhl2 in the regulation of Egr1 transcription. A) In the wild type Srf protein is abundant and binds to Srf binding sites (red squares) in the promoters of* Egr1 *and* Fhl2*. In combination with MRTF, it induces the transcription of* Fhl2*. The Fhl2 protein sequesters the activated ERK1/2 kinase in the cytoplasm, thereby inhibiting the activation of theTCF family of Srf cofactors (Elk1, Elk3, Elk4). Without the activation of the TCF factors, the* Egr1 *gene is transcribed only at a very basal level. B) In the conditional* Srf *KO, the Srf protein is gradually degraded. Remaining Srf protein is more likely to be bound to strong target genes such as* Egr1*, while* Fhl2 *transcription was ceased. Without the Fhl2 protein, ERK1/2 is no longer sequestered within the cytoplasm, translocates to the nucleus and activates the TCF factors. Together with remaining Srf protein, the activated TCF factors positively regulate the transcription of* Egr1*.*

The biological relevance of *Egr1* during mesoderm development and axis extension is not clear. Generally, it encodes a transcription factor that can act as both a repressor and an activator of transcription (Bahouth et al. 2002; Chapman & Perkins 2000; Lemaire et al. 1990; Wang et al. 2005), and it is implicated in the regulation of cell proliferation and apoptosis (Virolle et al. 2003). Additionally, it was shown in *Xenopus*, that *Egr1* overexpression inhibits the expression of the *Xenopus* homolog of *Brachyury* (*xBra*), connecting Egr1 with mesoderm formation (Nentwich et al. 2009). However, there are no data showing that Egr1 is important for mesoderm formation in the mouse. Also, Brachyury expression is not affected by the upregulation of *Egr1* in $Srf^{flex1/flex1}; T_{streak}$ *Cre* embryos, arguing against a direct connection between Egr1 and mesoderm development.

3.7. Srf in conjunction with Brachyury activity

The *Brachyury* gene (T) encodes a transcription factor that is crucial for mesoderm generation and axis elongation (Rashbass et al. 1991). The protein is essential for several events during development, such as the maintenance of mesodermal progenitors as well as the specification and migration of nascent mesodermal cells (Martin & David Kimelman 2010; Beddington et al. 1992; Wilson & Beddington 1997; Wilson & Beddington 1996). Although Brachyury was identified as a transcription factor in 1995 (Kispert et al. 1995), very few of its target genes have been identified to date, and its exact molecular function continues to remain elusive.

There is evidence that Brachyury might directly regulate *Srf* expression. Barron and coworkers have demonstrated that tail and cardiac-specific expression of Srf requires the presence of the 3' untranslated region (3'UTR) of *Srf*, and that this region contains several T-box binding motifs (Barron et al. 2005). They further show that several T-box proteins, including T and Tbx5, can induce tail and cardiac-specific activation of a reporter construct composed of a minimal *Srf* promoter, *LacZ* and the *Srf* 3'UTR. Moreover, mutation of the T-box binding sites in the 3'UTR inhibits this activation (Barron et al. 2005). *Tbx5* KO mice have a reduced *Srf* transcript level in the tail and heart, suggesting that the effect of Tbx5 is either at the transcriptional level or on mRNA stability. Since Srf is able to bind to its own promoter and induce transcription (Belaguli et al. 1997; Kasza et al. 2005), it is difficult to distinguish between an effect on the transcript level or on the protein level. Hence, there might also be an effect on the posttranscriptional level, inhibiting the generation of Srf protein rather than that of *Srf* transcripts.

Current work in our lab demonstrates that Srf expression in differentiating ES cells depends on T-activity. In particular, *T*-knockdown ES cells that are induced to differentiate into the mesodermal lineage by BMP4 treatment (as described in 2.1.4.1.) display a severe downregulation of *Srf* as compared to differentiated control cells (Shin and Herrmann, personal communication). Barron and coworkers speculated that T- and presumably Tbx6-induced expression of *Srf* in nascent mesoderm might be implicated in the specification of paraxial mesoderm. However, the data presented in the current work do not provide evidence that Srf has a role in mesoderm specification. Instead, they clearly demonstrate that Srf is crucial for cell migration. In agreement with this, T activity has also been related to cell motility. Direct measurements of active mesoderm migration on extracellular matrix *in vitro* revealed that T KO cells from E8-E9 embryos display a reduced migration rate compared to wt cells (Hashimoto et al. 1987). Moreover, *in vivo* analysis of *T* KO ES cells reveals that they are compromised in their ability to migrate away from the ps (Wilson et al. 1995; Wilson & Beddington 1997). Together, these findings suggest that one role for Brachyury in the ps and nascent mesoderm could be the activation of Srf expression, either on the transcriptional or the posttranscriptional level, and that Srf acts downstream of T in the regulation of cell migration during axis extension. In this context, it would be interesting to analyze whether the migratory defect of T KO cells correlates with an impaired reorganization of the actin cytoskeleton, as was the case in *Srf*-deficient cells. Additionally, an *ex vivo* migration assay with caudal ends from $T^{-/-}$ embryos would allow further comparison between Srf- and T-dependent cell motility. However, molecular data argue against a direct control of Srf activity by T. Comparison of the differentially regulated genes in $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos with data from *T* KO embryos (Wittler and Herrmann, personal communication) revealed that, among 288 significantly deregulated genes in *T* KO embryos, neither *Srf* itself nor any of the Srf-dependent genes (Tables 2 and 3) are present. Also, the analysis of data from ES cells in which *T* was knocked down and then differentiated into mesoderm (Shin and Herrmann, personal communication) does not reveal dysregulation of Srf-dependent genes during axis extension.

In addition to the idea that *Srf* expression in the ps and nascent mesoderm might be regulated by T, there is also speculation about the possible regulation of Brachyury by Srf. As mentioned earlier, the *Srf* KO mouse does not develop a ps or express any detectable mesodermal markers including *Brachyury*, which led to the assumption that Srf is an essential regulator of mesoderm formation (Arsenian et al. 1998).

The results of this study however do not support the idea of direct regulation of *Brachyury* by Srf. First, *Brachyury* expression is not altered in *Srf*-deficient caudal ends, as detected by gene expression profiling and *in situ* hybridization (Fig.S.5). Furthermore, a direct binding of Srf to the Brachyury promoter could not be detected by any of the assays attempted. Repeated experiments to identify Brachyury promoter enrichment either by ChIP-Seq or by qPCR analyses of Srf ChIPs from embryonic tissue, ES cells, or P19 cells all failed (data not shown). It could be argued that the embryo ChIP might not be sensitive enough for detection of Brachyury enrichment and that undifferentiated P19 cells do not express *Brachyury*; therefore Srf might not bind to the Brachyury promoter in these cells. To bypass the latter issue, I performed Srf ChIP experiments with chromatin from P19 cells after two days of differentiation into the mesodermal lineage. Although P19 cells express high amounts of *Brachyury* after two days of differentiation (Marikawa et al. 2009), binding of Srf to the *Brachyury* promoter could not be detected (data not shown). In agreement with these results are observations by Nordheim and coworkers (Panitz et al. 1998). They analyzed the relationship between the *Xenopus* homologues of *Srf* (*XSrf*), *Elk1* (*Xelk1*), *Egr1* (*Xegr1*) and *Brachyury* (*Xbra*) in animal cap explants, which are explanted regions of the blastula embryo that can be induced to form mesoderm (Nieuwkoop 1973). While a constitutively active version of XSrf (Srf-VP16) could activate *Xegr1* in the absence of any mesoderm inducing factors, the same treatment did not activate *Xbra*. Moreover, both *Xegr1* and *Xbra* could be activated by treatment with mesoderm-inducing factors, namely BMP4, FGF or Activin, but transfection of a dominant-negative version of the *xenopus* ternary complex fator *Xelk1* inhibited only *Xegr1* transcription, while *Xbra* expression was unaffected (Panitz et al. 1998). Together, the data of the current work and from others argue against a direct regulation of *Brachyury* by Srf. However, it cannot be ruled out that there is a subset of cells in which Srf transiently binds to the *Brachyury* promoter. For example, this could happen during the initial activation of *Brachyury* transcription at the onset of gastrulation. In this case, the gene expression data of the conditional *Srf* KO embryos, in which loss of *Srf* occurs only after the initial induction of Brachyury activity, would not reflect *Brachyury* deregulation.

3.8. Identification of novel Srf target genes

The genome-wide identification of transcription factor target genes by ChIP-Seq is a relatively new method, with first results published in 2007 (Johnson, 2011; Robertson et al., 2007). This method significantly increased the number of identified transcription factor target genes, however, it

has the limitation that only genes that are highly enriched can be clearly identified. These genes can be detected when bioinformatic identification of enriched regions is performed using a high stringency filter, but target genes with lower enrichment are lost. In the case of a low stringency filter, the ChIP-Seq data include lower enriched targets, but they are also prone to contain a high number of false positives, making it difficult to distinguish the more weakly bound targets from the background.

In the course of this work, I generated three different data sets for Srf-bound regions. Two data sets were generated by ChIP-Seq using chromatin from cultured P19 cells, and the identification of Srf-bound regions in P19 cells performed equally well with either a conventional Srf antibody-mediated ChIP method or with the alternative Srf-Avi[Bio] ChIP method. The bioinformatic identification of peaks was performed using high stringency criteria, and the data were filtered for peaks with a maximal 4kb distance from the TSS of an annotated gene, resulting in a relatively small number of target genes. The data sets were compared, and overlapping genes were regarded as likely Srf targets, comprising 24 known and 24 novel putative Srf target genes (Table1). These novel putative targets add to a growing number of Srf targets, with more than 200 having been functionally validated to date (Miano, 2010). However, to be considered as Srf target genes, the Srf binding site within the putative target genes has to be validated, for example by luciferase assay. The third data set of Srf-bound genomic regions comes from chromatin isolated from embryonic caudal ends. As demonstrated in section 2.1.4., this data set comprises a large number of false positives and thus, supposedly, only a small subset of true Srf-bound regions. However, using a combination of the embryo ChIP-Seq data, the P19 ChIP-Seq data, and gene expression data allows for the identification of genes that are very likely to be directly regulated by Srf (Tables 2 and 3). In this manner, thirteen novel putative Srf genes were identified as depending on Srf activity in the ps and nascent mesoderm of midgestation embryos. While the roles of Mrp*s23, Olfm1, Noc3l, Cdv3, Immt, Naca, Ndufc2, Soat2, Rps6ka1,* and *Nlgn2* during axial elongation remain elusive, *Rasa1, Fscn1,* and *Crk* are associated with the regulation of the actin cytoskeleton. Together with the other genes that were identified as Srf targets during axial elongation, which had previously been described as Srf target genes and which are also associated with the regulation of the actin cytoskeleton (Table 2), they are likely to contribute to the observed migration defect in $Srf^{flex1/flex1}$; T_{streak} *Cre embryos.*

3.9. Streptavidin-Biotin mediated ChIP with in vivo *tissue*

In the process of the current work, I developed a streptavidin-biotin affinity-based ChIP method, that, for the first time, allows for the enrichment of biotin-tagged transcription factor target genes from *in vivo* tissue (Fig.19; 21; 29,A; 30,A and B). The high streptavidin-biotin affinity has been used for ChIP experiments before, and global analysis of promoter occupation was performed, either by ChIP-Chip or ChIP-Seq analysis (Kim et al. 2009; Soler et al. 2010; Soler, Andrieu-Soler, et al. 2010), but the described experiments were performed using cultured cells, and not *in vivo* tissue. In 2003, de Boer et al. demonstrated that the bacterial biotin ligase hBirA can be used for *in vivo* tagging of the Avi-tagged erythropoietic transcription factor EKLF (De Boer, 2003) and similarly, Driegen and colleagues showed *in vivo* biotinylation of the Avi-tagged transcription factors Gata1 and Oct6, which were tissue-specifically expressed (Driegen et al. 2005). Both reports prove that *in vivo* biotinylation is possible in principle, but protein-DNA interactions were not analyzed.

An additionally achievement of the presented work was to show, *in vitro* and *in vivo*, that application of this method improved Srf target enrichment as compared to enrichment seen with a conventional antibody-based ChIP method. Particularly, it allowed for the enrichment of only weakly-bound Srf targets such as *Fos* from embryonic tissue, which was not possible with the antibody based ChIP method (Fig.21). The streptavidin-biotin based ChIP method was also superior to the antibody ChIP method, when used with low amounts of chromatin, which was derived from *in vitro* tissue (Fig.14).

The establishment of a biotin-mediated ChIP system that works efficiently *in vivo* offers several advantages over using an antibody-mediated ChIP method. The primary advantage is that, due to the high biotin-streptavidin affinity and thus the higher ChIP efficiency, the requirement for the amount of starting material can be decreased (Fig.14). Antibody-mediated ChIP usually requires high amounts of chromatin, which is usually not a problem for *in vitro* experiments, but becomes important when working with embryonic tissue. However, the usage of lower amounts of tissue and extensive washing during the streptavidin-biotin method is accompanied with a relatively low yield in ChIP-DNA, which can be disadvantageous when it comes to the genome-wide analysis of the precipitated DNA by massive parallel sequencing. Currently, the required amount of ChIP-DNA for a ChIP-Seq experiment is around 10ng, independently of the system used for sequencing. However, several reports have recently described a high-fidelity linear DNA amplification method (LinDA) for ChIP-Seq,

which requires only picogram amounts of ChIP-DNA as starting material (for example see Shankaranarayanan et al. 2011)

A second advantage of the method established in the course of this work is that the biotinylated transcription factor can be expressed in a tissue-specific manner. This becomes especially important when the tissue of interest is enclosed by other tissues that also express the transcription factor of interest. It would be nearly impossible to obtain tissue-specific promoter occupancy data from such tissues with an antibody mediated ChIP method. In contrast, tissue-specific expression of a biotinylated transcription factor makes dissection of the embryo, itself not a trivial matter, dispensable. Generally, accurate tissue-specific expression of the transgene depends on integration into the right locus and proper insulation of the transgene from endogenous promoter activity. In the current work, I used the T_{streak} promoter for driving expression of the transgene encoding Srf-Avi[Bio]. The tissue-specificity of this promoter was previously demonstrated (Perantoni et al. 2005), however, upon integration into the *Gt(ROSA)26S locus,* transgene expression was found to not be entirely tissue-specific (Fig.20). It is notable that in subsequent experiments, the expression specificity of the transgene could be greatly improved by changing the design of the integrated expression cassette and the recipient construct. In detail, the Tet responsive element was removed from the donor construct and the reverse Tet silencer from the recipient construct, and a 1.2kb insulator was integrated into the 3'site of the recipient construct to avoid putative endogenous enhancer activity from the *Gt(ROSA)26S* locus (Fig. S.7,A). The ES cell line (T_{streak}*Srf-noTre-Avi[Bio]*#4), which expressed the new transgene, was used for the generation of transgenic embryos by tetraploid complementation. These embryos displayed a transgene expression that, in part, resembled the activity of the T_{streak} promoter (Fig. S.7, B). T_{streak}*Srf-noTre-Avi[Bio]*#4 ES cells that differentiated into the mesodermal lineage were successfully used for Srf-Avi[Bio] ChIP (Fig. S.7, C). The enrichment of Srf target genes was similar to the enrichment seen for the ChIP with Srf-Avi[Bio]#8 ES cells, indicating that, in principle, these cells could be used to tissue-specifically identify Srf target genes *in vivo* using the Biotin-Streptavidin system. For further improvement of the expression specificity other loci might be more suitable, such as the *Col1a* locus, which is currently being used by our lab.

An advantage that specifically comes from the transgene system used to generate the Srf-Avi[Bio] ES cells is the standardization of transgene integration (see 2.1.4.1. and Fig.18). The coding sequence of Srf can be easily exchanged for the sequence of any other transcription factor, and new ES cells expressing a biotinylated transcription factor of choice could be analyzed. In this regard, the

Srf-Avi[Bio] experiments can be taken as a proof-of-principle. This becomes especially interesting when dealing with transcription factors for which no suitable ChIP-graded antibody is available.

Another interesting prospect is the combination of Streptavidin-Biotin mediated ChIP and the employment of *in vitro* differentiation protocols for ES cells, which have been developed in our lab (Schroeder and Herrmann, personal communication) and were briefly introduced in section 2.1.4.1. (Fig.19,A). Streptavidin-Biotin ChIP using ES cells that express a biotinylated version of a transcription factor under control of a tissue-specific promoter such as the T_{streak} promoter, and that were differentiated into any tissue (e.g. mesoderm) *in vitro,* would allow the identification of tissue-specific target genes. Strikingly, this would bypass the requirement for embryo generation, and would be independent of tedious cell sorting experiments.

4. Summary

Following the process of gastrulation, the anterior posterior axis (A-P axis) of the vertebrate embryo progressively elongates. Elongation is based on the continous generation of mesodermal tissue by a growth zone in the caudal end of the embryo, as well as the subsequent migration of newly generated mesodermal cells. The transcription factor Serum Response Factor (Srf) was previously shown to be expressed in the caudal end and nascent mesoderm of mouse, chick, and frog embryos. The goal of this work was to identify its role during axial elongation in the mouse.

Conditional loss of *Srf* specifically in the caudal end and nascent mesoderm resulted in mouse embryos that displayed severe axis truncation. In order to identify Srf binding sites from embryonic caudal ends, I modified a streptavidin-biotin affinity-based ChIP method for *in vivo* application. *In vivo* and *in vitro* ChIP-Seq data were combined with gene expression data, which had been generated using caudal ends from conditional *Srf* KO embryos, to identify 27 genes that are directly regulated by Srf during axial elongation. These genes included both previously known and thus far unknown putative Srf target genes. Nearly all of them are associated with cell migration, mainly with the reorganization of the actin cytoskeleton, indicating that the main function of Srf in the process of axial elongation is the regulation of cell motility. In accordance with the molecular data, *ex vivo* migration assays revealed that *Srf*-deficient mesodermal cells displayed impaired migration, accompanied by dramatic differences in the organization of the actin cytoskeleton. In mutant embryos, *Srf*-deficient mesodermal cells accumulated in the caudal end of the embryo instead of being spread along the A-P axis, indicating that the lack of *Srf* results in impaired migration *in vivo*. The accompanied axial truncation is in agreement with previous reports which link impaired cell migration to defective axial elongation. It was previously thought that Srf was involved in mediating FGF signaling during axial elongation, however significant correlation between genes that are directly regulated by Srf and genes that are regulated by FGF signaling was not detected here. Altogether, work presented here strengthens the evidence for the requirement of mesodermal cell migration during axial elongation, and has identified Srf as a major player in executing the genetic program in the process of migration.

5.Zusammenfassung

Nach dem Prozess der Gastrulation verlängert sich die anterior-posteriore Achse (A-P Achse) kontinuierlich. Die Verlängerung basiert auf der fortlaufenden Generierung von mesodermalem Gewebe durch eine Wachstumszone im kaudalen Ende des Embryos, sowie der anschliessenden Migration der neu entstandenen mesodermalen Zellen. Der Transkriptionsfaktor Serum Response Factor (*Srf*) ist im kaudalen Ende sowie im entstehenden Mesoderm von Maus-, Huhn- und Froschembryonen exprimiert. Ziel dieser Arbeit war es, die Rolle von Srf während der Achsenverlängerung in der Maus zu identifizieren.

Die konditionale Inaktivierung von Srf spezifisch im kaudalen Ende und dem entstehenden Mesoderm resultierte in Maus-Embryonen, die einen schwerwiegenden Achsenabbruch aufwiesen. Für die Identifizierung von Srf-Bindestellen im embryonalen kaudalen Ende habe ich eine auf Streptavidin-Biotin-Affinität beruhende ChIP-Methode modifiziert, um sie *in vivo* anwenden zu können. *In vivo* und *in vitro* ChIP-Seq-Daten wurden mit Expressionsdaten kombiniert, die aus kaudalen Enden der konditionalen Knockout Embryonen gewonnen wurden. Dies führte zur Identifizierung von 28 Genen, die während der Achsenverlängerung direkt von Srf reguliert werden. Diese Gene beinhalteten sowohl bereits bekannte sowie bisher unbekannte mutmaßliche Srf-Zielgene. Der Großteil dieser Gene wird mit Zellmigration assoziiert, vorallem mit dem Umbau des Aktin-Zytoskeletts, was darauf hindeutete, dass die Hauptfunktion von Srf während der Achsenverlängerung die Regulierung der Zellbeweglichkeit ist. Übereinstimmend mit den molekularen Daten wiesen *Srf*-defiziente mesodermale Zellen eine beeinträchtigte Migration in *ex vivo*-Migrationsanalysen auf, die mit drastischen Abweichungen in der Struktur des Aktin-Zytoskeletts einhergingen. In Mutanten sammelten sich *Srf*-defiziente mesodermale Zellen im kaudalen Ende der Embryonen an, anstatt sich entlang der A-P Achse auszubreiten, was darauf hinweist, dass das Fehlen von *Srf* zu einer Beeinträchtigung der Migration *in vivo* führt. Der damit einhergehende Achsenabbruch ist in Übereinstimmung mit Berichten, die eine beeinträchtigte Zellmigration mit ausbleibender Achsenverlängerung verbinden. Es wurde vermutet, dass Srf an der Vermittlung von FGF-Signalen während der Achsenverlängerung beteiligt ist, jedoch konnte im Rahmen dieser Arbeit keine signifikante Korrelation zwischen Srf-regulierten und FGF-regulierten Genen festgestellt werden. Diese Arbeit bestätigte, dass Migration mesodermaler Zellen notwendig

für die Achsenverlängerung ist, und hat Srf als einen wichtigen Faktor für die Durchführung des dafür verantwortlichen genetischen Programms identifiziert.

6. Contributions to the experimental work

I would like to acknowledge the following people who contributed to specific parts of the presented work:

Lars Wittler performed the embryo dissections for the microarrays and for the *ex vivo* migration assay. He also did most parts of the FGF inhibitor experiment, including dissection of the tail halves and culturing of the halves. Karol Macura performed the tetraploid complementation assays to generate the transgenic embryos used for the Srf-Avi[Bio] ChIP experiments. All other work in the animal facility was conducted by the animal caretaker Eileen Jungnickel. Martin Werber did the bioinformatic analyses of the microarray and ChIP-Seq data.

7. Material and Methods

7.1. Mouse strains and husbandry

All mice used in this study were housed under specific pathogen free conditions at the animal facility of the Max-Planck Institute for Molecular Genetics, Berlin. The animals are kept under a 12 hours cycle of light and dark at 22°C and a relative humidity of 55±10%. They were fed a pelleted, irradiated diet (ssniff M-Z®; Soest, Germany) composed of 22% raw protein, 4.5% fat, 3.9% raw fiber and 6.8% raw ashes. Distilled water was provided *ad libidum*. Timed matings were set up during the day cycle, and noon on the day of finding a copulation plug was assumed to be E0.5. All animal experiments were approved by the Berliner State Office for Safety at Work, Health protection and Technical Safety (Landesamt für Arbeitsschutz, Gesundheitsschutz und technische Sicherheit, LAGetSi) and carried out in accordance with the german animal welfare act (Tierschutzgesetz, TSchG). All work in the animal facility was conducted by the animal caretaker Eileen Jungnickel (Max-Planck Institute for Molecular Genetics, Berlin).

The genetic background strain of all mice that were used in the course of this work was the inbred strain C567BL/6 (Jackson Laboratory). For the conditional KO of *Srf* and *Snai1*, the following mouse strains were bred:

Table 5: *Mouse lines used for the studies of conditional loss of Srf and Snai1*

Mouse line	Provided by	Reference
$Srf^{flex1/flex1}$	A. Nordheim	Wiebel et al. 2002
$Snai1^{flox/flox}$	T. Gridley	Murray et al. 2006
T_{streak} *Cre*	A. Gossler	Perantoni et al. 2005*
Msd Cre	A. Gossler	Wehn et al. 2009*

* The mouse lines used in the current work were generated by A. Gossler et al. using the same promoter elements that were used for the Cre-lines in the cited reports.

Genotyping: Mice and embryos were genotyped for the presence of the floxed *Srf* or *Snai1* alleles as described in (Wiebel et al. 2002) and (Murray et al. 2006) or of the T_{streak} Cre (Perantoni et al. 2005) or Msd Cre (When et al. 2009) recombinase. Genomic DNA was isolated according to the protocol described in (Laird et al. 1991), and PCRs were set up as followed and using the primers shown in table 6.

Genotyping PCR set up:

DNA	1 µl
DMSO	2.5 µl
$MgCl_2$	2 µl
10mM dNTP	1 µl
10pmol/ml Fwd. Primer	2 µl
10pmol/ml Rev. Primer	2 µl
10x PCR Puffer (Invitrogen)	5 µl
H_2O	33.5 µl
Taq polymerase (Invitrogen)	1 µl
Final volume	50 µl

Table 6: *Primers used in Genotyping PCRs*

Mouse line	**Primer pair**	**Forward**	**Reverse**
Srf$^{flex1/flex1}$	E x R (floxed allele)	AGTTCATCGACAACAAGCTGCGG	GAGATTTCCACAGAAAGCAACGG
	L x R (recombined allele)	AGTTCATCGACAACAAGCTGCGG	GCTCGCAGCGGCGGCCAGAT
Snai1$^{flox/flox}$	FloxF2 x FloxR (floxed allele)	CGGGCTTAGGGTGTTTTCAGA	CTTGCTTGGTACCTGCCTTC
	FloxF2 x FloxDelR1(recombined allele)	CGGGCTTAGGGTGTTTTCAGA	TGAAAGCGGCTCTGTTCAGT
T_{streak} Cre	T streak forward x Cre-rev	AATCTTTGGGCTCCGCAGAG	ACGTTCACCGGCATCAACG
Msd Cre	Melta 38 x Cre-rev	ATCCCTGGGTCTTTGAAGAAG	ACGTTCACCGGCATCAACG

7.2. Tissue culture

Standard tissue culture work has been performed in a laminar flow hood (Hera safe clean bench, Heraeus). Before use, unsterile materials were autoclaved and disinfected with 70% ethanol.

7.2.1. ES cell culture

Undifferentiated ES cells were cultivated on a monolayer of mitotically inactivated primary embryonic fibroblasts (feeder cells) ($3x10^6$/plate) in a gelatin-coated 10cm cell culture dish (Corning), and incubated at 37°C in a humidified 7.5% CO_2 incubator (HERAcell 150; Heraeus). Cells were grown in ES cell medium composed of Dulbecco's Modified Eagle's Medium (DMEM) containing 4,500mg/ml glucose, w/o sodium pyruvate (Sigma Aldrich), 10% (v/v) ES cell-qualified, heat inactivated fetal bovine serum (FBS) (Gibco), 2mM L-glutamine (Sigma-Aldrich), 50 U/ml penicillin (Sigma-Aldrich), 50µg/ml streptomycin (Sigma-Aldrich), 1% 100x non-essential amino acids (Sigma-Aldrich), 0.1mM β-mercaptoethanol (Sigma-Aldrich), 1% 100x nucleosides (Sigma-Aldrich). 1000 U/ml murine leukemia inhibitory factor (LIF) (Chemicon) were added to keep the ES cells in an undifferentiated state. The medium was exchanged daily.

7.2.2. Generation of modified ES cells

For the modification of ES cells, transgenic recipient cells were transfected with linearized exchange vector, harboring a transgene flanked by loxP sites, and Cre recombinase to allow for site specific recombination into the *Gt(ROSA)26S* locus.

Recombinase mediated cassette exchange (RMCE): 3 x 10^5 Rosa26A/S ES cells were seeded onto a monolayer of feeder cells in a gelatin-coated 6 well plate (Corning) in 1.5ml ES cell medium. After overnight incubation, fresh medium was added, and ES cells were transfected with 5µg of exchange vector and 0.1µg of Cre recombinase expression vector using Lipofectamine 2000 (Invitrogen). DNAs were mixed with 125µl Opti-MEM, and 25µl of Lipofectamine were mixed with 110µl OPTI-MEM. Both mixtures were incubated for 5min at RT, then mixed and incubated for further 15min. Next, the combined mixture was added to the cells, and incubated for 3-5 hours (37°C incubator) to induce RMCE. In order to split the cells, they were carefully washed with cell-culture grade D-PBS (Lonza), 0.5ml trypsin/EDTA solution (Gibco) was added and the cell were incubated at 37°C for 10min in order to break down the cell-cell contacts. Addition of 2ml ES cell medium stopped the activity of the enzyme, and a single cell suspension was generated by pipetting up and down. Cells were collected by centrifugation (5min, 200 g), and cells were seeded on a monolayer of feeder cells ($1x10^6$/plate), which were resistant to the antibiotic used for the selection of positive clones (puromycin or neomycin), in 3 x 6cm gelatinized cell culture dishes (Corning) in a ratio of 1/6, 2/6, and 3/6. Selection was started 24 hours thereafter, using medium containing Geneticin (in case of Neomycin resistance) (Gibco) in a concentration of 250µg/ml or Puromycin (Sigma) in a concentration of 0.75µg/ml. The

medium was exchanged daily until resistant colonies were visible (approximately 1 week after electroporation).

Picking of ES cell colonies: A gelatinized, flat-bottom 96-well plate (Corning) containing a monolayer of feeder cells ($1x10^6$/plate) was prepared one day before the picking procedure, and fresh ES cell medium was added to the cells 3-4 hours prior to picking. Individual colonies were picked using disposable 10µl pipette tips under a stereo microscope (MZ8; Leica) and transferred to wells of a round-bottom 96-well plate (Corning) prefilled with 20µl Trypsin-EDTA solution. As soon as 96 (or less) colonies had been picked, the cells were placed into the 37°C incubator for 10min. The reaction was stopped by adding 100µl ES cell medium per well and the cell colonies were disaggregated by pipetting the suspension up and down for several times with a multi-channel pipette (Eppendorf). The picked ES cell colonies were transferred to a 96-well plate and grown for two days.

Splitting and Freezing of ES cells: After two days, the cells were washed twice with 200µl PBS, and incubated with 50µl Trypsin-EDTA per well at 37°C for 10min. Meanwhile, 50µl ES cell medium were transferred to a gelatinized flat-bottom 96-well plate, and 200µl ES cell medium were added to a gelatinized flat-bottomed 96-well plate. Trypsinization in the original 96-well plate was stopped by the addition of resuspension medium (100µl/well). 50µl of this cell suspension were transferred to the prepared round-bottom 96-well plate containing 50µl ES cell freeze medium. The plate was sealed using parafilm and stored in a storofoam box at -80°C. 50µl cell suspension were transferred to the gelatinized flat-bottom 96-well plate, containing 200µl ES cell medium, and grown to confluence for DNA extraction (first expansion plate). 200µl ES cell medium were added to the remaining 50µl cell suspension of the original plate and also grown to confluence for DNA extraction (second expansion plate). DNA was isolated and subsequently used for Southern blot analysis.

Southern blot analysis: Genomic DNA isolation for southern blot analysis was performed according to the protocol described in (Ramírez-Solis et al. 1993). Briefly, ES cells were grown to confluence in a 96-well plate, washed twice with 200µl PBS, and lysis buffer containing 1mg/ml proteinase K were added (50µl/well). The cells were incubated over night at 37°C in a humid chamber. The next day, ice cold 75mM NaCl/100% EtOH mix were added (100µl/well) w/o mixing. The plate was incubated at RT to precipitate the DNA at the bottom of the well. The plate was then carefully inverted to discared the solution and excess liquid was blotted on wet paper. The wells were rinsed three times by adding 200µl 70% EtOH per well. After the final wash, the DNA was dried

at RT. The first expansion plate was sealed with parafilm and stored at -20°C. The second expansion plate was used for digestion using restriction enzymes.

A restriction digest master mix was prepared containing 1x Buffer (Promega), 1mM spermidine (Promega), 100mg/ml BSA (Promega), 100µg/ml RNase A (Fermentas) and 30-50µl of the restriction enzyme(s) (Promega). 30µl of the master mix were added per well w/o mixing and incubated in a humidified chamber at 37°C for 4 hours. Next, the content of the wells was mixed by pipetting and incubation was proceeded overnight in the humidified chamber. At the following day, 6µl 6x gel loading buffer were added to each well for electrophoretic analysis. DNA fragments were separated using a 0.7% 1x TAE agarose gel in 1x TAE electrophoresis buffer for 12-18 hours at 30V. DNA fragments were documented with a GelDoc2000 system (Bio-Rad).

Following the documentation, the agarose gel was pretreated to prepare the transfer of large DNA fragments by 30min washes with denaturation buffer, neutralization buffer and 20x SSC. Next, the DNA was transferred to a blotting membrane (Amersham Hybond-N; GE healthcare) overnight using the Whatman® TurboBlotter. The following day, the DNA was UV-crosslinked to the membrane using the UV Stratalinker 2400 (Stratagene; 500µJ/cm^2). The membrane was either air-dried and stored at -20°C in a sealed plastic bag of directly used for hybridization.

For the generation of radioactively labeled probes, 25ng of the template DNA were diluted in 45µl TE buffer and denaturated for 5min at 95°C in a heating block. The solution was cooled down on ice and pipette into a reaction tube containing a dried mix of dATP, dGTP, dTTP, Klenow enzyme, and random primers (Rediprime, Random Prime Labeling System, GE healthcare) . The components were mixed by pipetting before 5µl of α-^{32}P dCTP (Redivue, Amersham) with a specific activity of 3000Ci/mmol were added. The radioactive sample was incubated for 30min - 1hour at 37°C to allow for the Klenow catalyzed labeling reaction. Meanwhile, the blotted membrane was transferred to a glass tube and incubated with 15ml ExpressHyb Hybridization solution (Clontech) under continous rotation at 62°C for 1 hour. The labeled samples were purified using G-25 MicroSpin columns (illustra ProbeQuant, GE Healthcare)according to manufacturer's instructions. Following purification, the labeled DNA was denatured at 95°C (5min) and directly pipette into the glass tube containing the preincubated membrane and hybridized overnight under continous rotation at 62°C.

At the following day the membrane was washed twice with prewarmed 2x SSC (62°C) (62°C, continous rotation,10min). Next, the membrane was transferred to a plastic box containing prewarmed 0.2% SSC (62°C) and incubated for 20min under continous rocking. The radioactivity of

the membrane was controlled using a Geiger counter, and in case of radioactivity higher than 30Bq an additional washing step was performed, using 0.1% SSC (62°C; 30min). The membrane was sealed in a plastic bag and exposed to a phosphoimager (Amersham) to detect the labeled DNA.

Southern blot buffers:

Lysis buffer**:** for 1 liter: 100ml 1M Tris (pH8.3), 10ml 0.5M EDTA, 10ml 20%SDS, 5M NaCl in dH_2O

Denaturation buffer: for 1 liter: 300ml 1.5M NaCl, 100ml 0.5M NaOH in dH_2O

Neutralization buffer: for 1 liter: 300ml 1.5M NaCl, 500ml 0.5M Tris (pH8) in dH_2O

20x SSC: for 1 liter: 88.23g Sodiumcitrate, 175.32 NaCl in dH_2O

7.2.3. In vitro *differentiation of ES cells*

Feeder free ES cell culture: ES cells were grown in a 6cm dish as described in 4.2.1. For feeder depletion, ES cells and feeder cells were trypsinized with Trypsin-EDTA (Gibco) and replated with ES cell medium for 30min to the same 6cm dish. After 30min, the suspension is pipette to a new gelatinized 6cm dish. Since feeder cells attach faster to the gelatin surface than ES cells, this suspension contains mainly ES cells. This step is repeated several times, resulting in ES cell suspension free of feeder cells. About 1 million ES cells were seeded to a 3.5cm dish, cultivated in ES cell medium and used for *in vitro* differentiation experiments at the next day.

***In vitro* differentiation of ES cells into mesodermal tissue:** To differentiate the feeder depleted ES cells into mesodermal fated cells, the cells were singularized using Trypsin/EDTA and counted with a hematocytometer. The suspension was diluted to 40,000 cells/ml and a research pro multichannel pipette (5-100µl, Eppendorf) was used to produce 5µl drops containing approximately 200 ES cells. Drops were pipetted into square petri dishes (120 x 120 x mm; Greiner Bio-one). The dishes were inverted and incubated overnight at 37°C, allowing for the formation of cell aggregates. In parallel, fibronectin coated dishes were prepared. Therefore, fibronectin (Calbiochem #341631) was diluted (1:100) with PBS (Lonza), pipetted into 6cm cell culture dishes and incubated overnight. The following day the ES cell aggregates were collected in a 50ml tube (Corning) with DMEM medium (Gibco, Invitrogen) and centrifuged (1000rmp, 5min). The supernatant was carefully discarded and an additional washing step was performed, again using DMEM medium (to remove as much LIF as

possible). Meanwhile, the prepared fibronectin-coated dishes were washed once with PBS and left covered with PBS until further usage. Following centrifugation, the ES cell aggregates were plated onto the fibronectin-coated dishes with SFM (see 4.2.4.3.) supplemented with fresh recombinant mouse BMP4 (20ng/ml; R&D Biosystems) to induce differentiation and incubated at 37°C. The differentiation medium was renewed daily until the cells were harvested on the desired day.

Serum free medium (SFM):

Composition	Volume for 210ml	Company
DMEM/F-12 media	100ml	Gibco
Neurobasal media	100ml	Gibco
200mM L-glutamine	2ml	Lonza
B-27 supplement minus vitamin A (50x)	2ml	Invitrogen
N-2 supplement (100x)	2ml	Invitrogen
Albumin, bovine fraction V solution (7.5%)	1.33ml	Sigma
1-Thioglycerol	2.6µl	Sigma
Penicillin/Streptomycin (5K/5K)	2ml	Lonza

7.2.4. In vitro *differentiation of P19 cells*

Undifferentiated P19 cells were cultivated in 10cm cell culture dishes (Corning) at 37°C and 7.5% CO_2 humidified atmosphere in 10ml DMEM medium containing 1% Penicilin (250 U/ml) and Streptomycin (250 U/ml) (Sigma-Aldrich), and 10% fetal bovine serum (FBS; Gibco). Cells were passaged at a confluence of 70-80% (approx. 2.5-4x10^7 cells). In order to induce differentiation mainly into the mesodermal lineage, cells were cultivated in non-adherent petri dishes (Greiner) in the presence of 1% DMSO (Sigma) to allow for the formation of cell aggregates. After two days, embryoid bodies were harvested and used for ChIP experiments.

7.2.5. Ex vivo *migration assay*

To analyze the migratory potential of *Srf* deficient mesodermal cells, tailbud halves were dissected from $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos and heterozygous littermates at E9.5, and plated on µ-slides (ibidi; Martinsried). On the previous day, the slides were coated with the extracellular matrix protein fibronectin (Calbiochem #341631). For this purpose, fibronectin was diluted (1:100) with PBS

(Lonza), pipetted onto the µ-slides, and stored at 4°C until the next day. Shortly before usage, the fibronectin was discarded, the slides were washed once with PBS and DMEM medium containing 10% FBS was added. The embryonic tissue was placed in the middle of the slides and attached within one hour. Next, medium was removed until the tissue was just covered. Medium was exchanged daily. After two days, the maximal migration distance was measured and explants were stored for immunofluorescence staining.

7.2.6. FGF inhibition in tail half cultures

The experiment was performed as described in (Pourquié et al. 2007). Briefly, tails were dissected from wild type embryos at E9.5 and divided along the midline. The tail halves were cultured in 10% FBS in DMEM at 37°C in 7.5% CO2 either in 0.1% DMSO or in the presence of the pharmaceutical inhibitor SU5402, 100 µM (Pfizer) in 0.1% DMSO for four hours and than stored in TRIzol® (Invitrogen) until RNA extraction (see 7.6.3.2.).

7.3. Molecular Biology

All standard molecular methods were performed according to Sambrook (2006). These methods include the cultivation and storage of *E.coli* bacteria, the transformation of *E.coli* bacteria, the digestion of DNA using restriction enzymes and the separation of DNA fragments using agarose gels.

The following techniques were performed according to manufacturer's instructions: plasmid DNA preparation of bacterial cultures using QIAprep spin (Qiagen) and extraction of DNA from agarose gels using QIAquick spin (Qiagen). DNA and RNA concentrations were measured with the Qubit® Fluorometer (Invitrogen). Standard primers were ordered from Invitrogen. Sequencing of DNA was done by Agowa (Berlin) or eurofins; mwg operon (Ebersberg).

7.3.1. Chromatin Immunoprecipitation

Tissue preparation: ES cells and P19 cells were cultivated and differentiated as described in 4.2.1.,4.2.4., and 4.2.5. Embryonic caudal ends were dissected from E9.5 embryos in medium containing 10%FCS. Between 10 and 100 tailbuds of wild type embryos (for antibody mediated ChIP) or of transgenic embryos (for Srf-Avi[Bio] ChIP) were collected in 1ml volume. T_{streak}*Srf-Avi[Bio]#8* embryos and T_{streak}*Srf-noTre-Avi[Bio]*#4 embryos were generated by tetraploid complementation (according to Eakin & Hadjantonakis 2006) by Karol Macura. T_{streak}*Srf-Avi[Bio]#8* ES cells were treated with doxycycline (Sigma) (2µg/ml) for transgene activation.

Sample preparation: Samples were crosslinkFed with 0.7% Formaldehyd for 10min at room temperature RT). Crosslinking was stopped with Glycin (0.125M) and incubation for 5min on a shaker. P19 cells were scratched from the dish and collected in a 15ml falcon tube, centrifuged (5min, 1000rpm, 4°C), and transferred to a 1.5 ml eppendorf tube and washed 3x with ice-cold PBS. Embryonic caudal ends were washed 3x with ice-cold PBS in the tube in which they were collected. PBS was removed, and lysis was done with Upstate Lysis buffer at 7°C for 30min in a thermo shaker at 750rpm. The amount of used lysis buffer was 200µl – 400µl, depending on the amount of tissue used. In a next step, the chromatin was fragmented to an average fragment size of 200-400bp by sonication with a Branson 450 sonifier. The sonication was performed in 6 pulses of 10 seconds with 10% power and 50 seconds break. The sample was centrifuged (10,000rcf, 1 min, 8°C) and 10µl of the sample were used as Input DNA. The rest was shock freezed and stored at -80°C until it was used for

ChIP. Following protein digestion, RNA digestion, and phenol-chloroform extraction, the Input DNA was precipitated in H_2O and tested for proper fragment size by gel electrophoration.

Srf-antibody based ChIP: For a ChIP experiment 360ul ChIP Dilution buffer (containing Proteinase inhibitor), 20µl of protein A beads, 20µl of protein G beads and 1µl of Srf antibody (Santa Cruz; sc-335 X) were added to 40 µl of the prepared lysate. The sample was incubated in a non-stick tube at 4°C for 4 hours or overnight (oN). The bead/antibody/chromatin complex was separated from the unbound chromatin with a magnet and washed, using a rotating platform at 4°C. Washing was performed either with standard or high stringency:

Standard washing: washings each 8min: 2x Low Salt Immune Complex Wash Buffer (0.1%SDS) (+Proteinase inhibitor), 1x High Salt Immune Complex Wash Buffer, 1x LiCL Immune Complex Wash Buffer and 1x TE (1min).

High stringency washing: washings each 16min: 2x Low Salt Immune Complex Wash Buffer (2% SDS) (+Proteinase inhibitor), 2x High Salt Immune Complex Wash Buffer, 2x LiCL Immune Complex Wash Buffer and 1x TE (1min).

Next, the antibody/ chromatin complex was eluted from the beads with 100µl Elution buffer (30°C, 15 minutes). Elution was repeated with another 100µl and the eluates were combined to 200µl. 200µl TE and 12µl 5M NaCl were added: Following protein digestion, RNA digestion, and phenol-chloroform extraction, the ChIP-DNA was precipitated in H2O and analyzed for enrichment.

Srf-Avi[Bio] ChIP: For a ChIP experiment 360ul ChIP Dilution buffer (containing Proteinase inhibitor) and 20µl of streptavidin coated beads (Dynabeads® MyOne™ Streptavidin T1; Invitrogen) were added to 40 µl of the prepared lysate. The sample was incubated in a non-stick tube at 4°C for 4 hours. Washing and DNA preparation were according to the antibody-mediated ChIP.

7.3.2. Analysis of ChIP enrichment

Quantative Real time PCR (qPCR): Quantiative PCR analysis was carried out on a StepOnePlus™ Real-Time PCR System using Power SYBR® Green PCR Master Mix in a 10 or 20µl volume. Of the 10µl ChIP-DNA, 1µl was usually used for a qPCR experiment, which was enough for the analysis of up to six specific genomic regions and two control regions (located upstream of the genes *Msgn1* and

Abcb1b). Fragmented, total genomic DNA was used for normalization and calculated versus the non-enriched control regions using the StepOnePlus Software v2.0.2. The sequences of the primers are shown in table 7.

Table 7: *Primers used for the detection of ChIP enrichment by qPCR*

Gene	Forward	Reverse
Msgn1 (control 1)	GGAACTTTGCACAGGAGGAG	TGGGTAACGGCTTCCTAATG
Abcb1b (control 2)	TTGAGTGAGGGCCTGAAAAG	GGCATCAGCTCCAAACATCT
Egr1	GGCCGGTCCTTCCATATTAG	CGAATCGGCCTCTATTTCAA
Fos	CTACACGCGGAAGGTCTAGG	GCGCTCTGTCGTCAACTCTA
Tpm1	ATGGCTCTGAGAGGTGGCTA	CGAGGCCGTCCTATTTCTAA
Tpm4	GTTGGGGTGTGTGTAGCTGA	TACTTGTGGCCAAGCAACAG
Actb	CCGAAAGTTGCCTTTTATGG	AAGGAGCTGCAAAGAAGCTG
Peak1	GACAACCTTGGCTGTCCTTG	AGGGTAAGGCAGGTGGATCT
Peak2	CAGCATCTAAGCAGGCATGA	GCAGCCTTCAGATGACGATA
Srf	GCAGCGAGTTCGGTATGTCT	AGGTATCCCCCAACCCTTC
Wdr1	GGTCGCGTCACGTCACTT	CCCCCAGCTTTGACCATA
Fhl2	AGTCCGCTGGGTTCCCTTAG	GTCCTAGGGCAGCGGTCT
Soat2	ACACACATGCACACATGCAC	CATGTGCATACCCGTGAATC
Zswim6	CTCCCGTGCCCCTACAAG	CCCTATAAACGGCACAAGGA
Egr1 downstream peak	CAGCCCCGCACCATGTAT	CCGGTGACAGAGAAGGAAAC
Fos downstream peak	ACCTAACCCTTCCCAGCTTC	CACCGTGGAAACCAAACAGT

Massive parallel sequencing: Massive parallel sequencing was performed by single read analysis with the Illumina/Solexa Genome Analyzer. Library preparation and sequencing was performed by the group of Bernd Timmerman (Sequencing facility, Max-Planck Institute for Molecular Genetics, Berlin) according to the illumina protocol. Between library preparation and sequencing, I tested the samples again for enrichment of *Egr1* and *Fos* by qPCR. After sequencing, the data were analyzed as described in 7.5.2.

7.3.3. Analysis of gene expression

Quantative Real time PCR (qPCR): Total RNA was isolated from embryos or cells using the Mini Kit (Qiagen). Reverse transcription of the RNA into cDNA was performed using the SuperScript II First-Strand Synthesis System (Invitrogen). All the procedures were performed according to the

manufacturer's instructions. Gene expression was analyzed by quantitative real-time PCR using Power SYBR Green PCR Master Mix (Applied Biosystems) and transcript-specific primers (Table 8) in a 8µl volume. The PCR mix was transferred to the wells of a MicroAmp Fast Optical 96-well Reaction Plate (Applied Biosystems) and 2µl of a 1:20 diluted cDNA sample was added to each well. The reactions were run in triplets on a StepOne Real-Time PCR System, and the results were analyzed using the StepOnePlus Software v2.0.2.

Table 8: *Primers used for the amplification of transcripts by qPCR*

Gene	Forward	Reverse
Hmbs (control 1)	CCTGGGCGGAGTCATGTC	ACTCGAATCACCCTCATCTTTGA
Pmm2 (control 2)	AGGGAAAGGCCTCACGTTCT	AATACCGCTTATCCCATCCTTCA
Tbx6	CTGAAGATCGCAGCCAATC	CCCGAAGTTTCCTCTTCACA
T	CAGCTGTCTGGGAGCCTGG	TGCTGCCTGTGAGTCATAAC
Zswim6	TCTCTGGCTTCTCCGATTGT	AACTTGTGTGCAGGCAGATG
Egr1	AACACTTTGTGGCCTGAACC	GGCAGAGGAAGACGATGAAG
Egr1 downstream transcript	CCCTTGCTGGACATCCTTTA	GAGAAGCAGCCAGATGTTCC
Fos downstream transcript	ACCTGTGTGCTGGTGTGTGT	GCCAGGAAGCATGACAAAAT

Gene Expression Profiling (Illumina Microarrays): The transcriptomes of $Srf^{flex1/flex1}$; T_{streak} *Cre* embryos at E8.5, 8.75 and 9.0 and of the appropriate heterozygous littermates, as well as of tail halves treated with the FGF inhibitor SU5402 or with DMSO for control, were analyzed by Illumina whole-genome expression arrays. In each experiment, four mutant samples or SU5402 treated samples were compared to four control samples. RNA was isolated using the RNeasy Mini Kit (Qiagen), and labeled using the Illumina TotalPrep RNA Amplification Kit (Ambion). All procedures were performed according to the manufacturer's instructions. 1µg of cRNA was provided for hybridization to MouseRef-8 v2.0 Expression BeadChip (Illumina), which was performed by Aydah Sabah (Service Department, Max-Planck Institute for Molecular Genetics, Berlin). ChIPs were scanned on the Illumina BeadArray Reader, and Data were analyzed by M.Werber (see 7.5.1.). Genes were considered to be significantly dysregulated with a fold change $\geq \log 2^{0.4}$, $p \leq 0.05$.

7.4. Histology

7.4.1. Embryo preparation

Mice were euthanized via cervical dislocation and embryos were harvested in ice cold PBS and fixed overnightat 4°C in 4% Paraformaldehyde (PFA; Sigma-Aldrich) in PBS. The next day, embryos were washed twice with PBS (10min, 4°C) and then processed manually through an EtOH series (30%, 50%, 70%; 30min each). Embryos were stored at -20°C in 100% EtOH.

7.4.2. Whole mount in situ hybridization

Whole Mount *in situ* hybridizations were carried out using the standard protocol described on the MAMEP website (http://mamep.molgen.mpg.de/index.php). Probes were generated by PCR on cDNA from whole embryos (E9.5) and subcloned into pBluescript II SK(+) (Stratagene). After verification by sequencing of the generated probe templates, *in situ* probes were generated as described on the MAMEP website using either SP6 or T7 polymerase (Promega).

7.4.3. Immunofluorescence staining

Parrafin sections: For sectioning, the embryos that were stored in 100% EtOH were processed using a MICROM STP 120 processor (MICROM) using the following program:

Solution	Time	Agitation
80% EtOH	120min	2
96% EtOH	120min	2
100% EtOH	60min	2
100% EtOH	60min	2
100% EtOH	60min	2
100% Xylene	90min	2
100% Xylene	90min	1
100% Paraffin	120min	1
100% Paraffin	120min	1

The specimens were embedded in metal molds with paraffin (Histowax; Leica) using an EC350-1 embedding station (MICROM) and placed onto the cooling plate until the paraffin was solidified. 4µm thick were cut using a rotary microtome (HM355 S; MICROM), transferred onto adhesion microscope slides (SuperFrost; Menzel) and dried overnight at 37°C. Slides were stored in a desiccated slide box at 4°C until further usage.

Sections were deparaffinized in xylene (3 x 5min) and rehydrated (100% EtOH, 2 x 5min; 90% EtOH, 5min; 80% EtOH, 5min; ddH_2O, 5min) and washed 2 x with PBS. Next, epitopes were unmasked by boiling the slides in a glass beaker containing a buffer with moderately acidic pH (pH6.0, 10mM sodium citrate) for 20min. The beaker was cooled down on ice to approximately 40°C. The slides were washed 3 x with PBS (5min) and blocked overnight at 4°C with 2.5%. The next day, the sections were incubated with the primary antibody diluted in 2.5% horse serum (see table 9) at RT for 2 hours in a humidified chamber. Unbound antibody was removed by 5 washes with PBS (5min). The sections were incubated with a secondary fluorescence coupled antibody for 1 hour and washed with PBS for 3 x (5min). For imaging the Axio Observer (Zeiss) and the AxioVision Software (Zeiss) was used.

Table 9: *Antibodies used for Immunofluorescence staining*

Antibody	Source	Company	Dilution (sections)	Dilution (whole mount)
anti-cleaved caspase3	rabbit	Cell Signaling	1:200	-
anti-phosphoH3(S10)	rabbit	Cell Signaling	1:200	1:500
anti-Brachyury	rabbit	Kispert & Herrmann, 1993	1:400	1:400
anti-Tbx6	rabbit	Grote & Herrmann	1:200	-
Anti-E-Cadherin	mouse	BD Biosciences	1:200	-
Anti-rabbit (secondary)	goat	Jackson ImmunoResearch	1:500	-
Anti-mouse	goat	Jackson ImmunoResearch	1:500	-

Tissue explants on µ-slides: Following the *ex vivo* migration assay, the cell actin-cytoskeleton was stained using Phalloidin-Fluoresceinisothiocyanat (FITC) conjugate (Sigma) according to the manufacturer's instructions in a dilution of 1:50 (in PBS).

Whole mount embryos: Fixation: Embryos were fixed in 1xPBS/50%MetOH/10%DMSO for 30 min at 4°C and subsequently in 80%MetOH/20%DMSO overnight. The next day, embryos were washed twice with 100% MetOH (10min) and subjected to a reverse MetOH series (70%MetOH/PBS, 50%MetOH/PBS, 30%MetOH/PBS and 2x PBS; each 10 minutes).

Staining: Embryos were washed for 30 min with PBS, for 10 min with 3%H_2O_2/PBS (fresh!) 4x with PBSTB (2x5 minutes and 2x30 minutes) and once for 30 minutes in PBSTBN. The embryos were incubated at 4°C overnight with the primary antibody (Table 9), diluted in PBSTBN. The next day, the embryos were washed with PBSTB (3x5min and 4x30min) and PBSTBN (1x30min) and incubated for 2 hours with a secondary fluorescence coupled antibody. This was followed by washes with PBSTB (2x5 min and 4x30 minutes) before imaging and fixation with 4%PFA overnight.

Buffers: PBSTB: standard PBS, 0.1% Triton X-100, 0.2% BSA fraction V (100mg in 50ml); PBSTBN: PBSTB, 10% heat inactivated FBS (Fetal bovine serum)

7.5. Bioinformatic data analysis

The described bioinformatic analyses were all conducted by Martin Werber, Max-Planck Institute for Molecular Genetics, Berlin.

Expression data analysis: Expression data were processed using the lumiR/Bioconductor package (Du et al. 2008). Raw data were analyzed using background substraction (bgAdjust) and the quantile normalization method. Differentially expressed genes were identified using the limma1 package.

ChIP-Seq data analysis: Sequences were mapped against the mouse genome using the bowtie software (http://bowtie-bio.sourceforge.net/manual.shtml), and peaks were identified with SICER (Zang et al. 2009). Additional analysis were performed using scripts generated by Martin Werber. Motif analysis was performed using meme (http://meme.sdsc.edu/meme/ intro.html).

8.Abbreviations

A-P axis	Anterior posterior axis; head to tail axis
AVE	Anterior visceral endoderm
BirA	Biotin ligase A
BMP	Bone morphogenetic protein
ChIP	Chromatin immuno precipitation
Cdh1	E-Cadherin
E	Day of Embryonic Development
ECM	Extracellular matrix
Egr1	Early growth response 1; transcription factor, Srf target gene
EMT	Epithelial mesenchymal transition
ERK	Extracellular signal-regulated kinase
FGF	Fibroblast growth factor
Fn	Fibronectin
hBirA	humanized version of the Biotin ligase A
IEG	Immediate early gene
IRES	Internal ribosomal entry site
LPM	Lateral plate mesoderm
MAPK	Mitogen activated protein kinase
MMP	Matrix metallo proteinase
MRTF	Myocardin related transcription factor
Msgn1	Mesogenin1
NES	Nuclear export signal
pA	poly-Adenylation signal
PE	Primitive Endoderm
ps	Primitive streak
psm	Presomitic mesoderm
RA	Retinoic acid
Srf	Serum response factor
T	Brachyury; ps and nascent mesoderm marker
Tbx6	T-box 6 transcription factor
TCF	Ternary complex factor
TE	Trophectoderm
TGFβ	Transforming growth factor
TRE	Tet responsive element
TSS	Transcriptional start site
tTS	Tetracycline-controlled Transcriptional Silencer
UTR	Untranslated Region
VE	Visceral endoderm
Wnt	Wingless-type MMTV integration site

9. References

Abu-abed, S. et al., 2001. The retinoic acid-metabolizing enzyme , CYP26A1 , is essential for normal hindbrain patterning , vertebral identity , and development of posterior structures. *Genes & Development*, (Kessel 1992), pp.226-240.

Affolter, M. et al., 1994. The Drosophila SRF homolog is expressed in a subset of tracheal cells and maps within a genomic region required for tracheal development. *Development (Cambridge, England)*, 120(4), pp.743-53.

Alberti, S. et al., 2005. Neuronal migration in the murine rostral migratory stream requires serum response factor. *Proceedings of the National Academy of Sciences of the United States of America*, 102(17), pp.6148-53.

Antoku, S. & Mayer, B.J., 2009. Distinct roles for Crk adaptor isoforms in actin reorganization induced by extracellular signals. *J Cell Sci.*, 122(22), pp.4228–4238.

Arsenian, S. et al., 1998. Serum response factor is essential for mesoderm formation during mouse embryogenesis. *EMBO Journal*, 17(21), pp.6289-6299.

Aulehla, A. & Herrmann, B.G., 2004. Segmentation in vertebrates: clock and gradient finally joined. *Genes & development*, 18(17), pp.2060-7.

Aurore, L. et al., 2003. MyoD Distal Regulatory Region Contains an SRF Binding CArG Element Required for MyoD Expression in Skeletal Myoblasts and during Muscle Regeneration. *Molecular Biology of the Cell*, 14(May), pp.2151-2162.

Bahouth, S.W., Beauchamp, M.J. & Vu, K.N., 2002. Reciprocal regulation of beta(1)-adrenergic receptor gene transcription by Sp1 and early growth response gene 1: induction of EGR-1 inhibits the expression of the beta(1)-adrenergic receptor gene. *Molecular pharmacology*, 61(2), pp.379-90.

Bai, S. et al., 2009. MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib. *The Journal of biological chemistry*, 284(46), pp.32015-27.

Balza, R.O. & Misra, Ravi P, 2006. Role of the serum response factor in regulating contractile apparatus gene expression and sarcomeric integrity in cardiomyocytes. *The Journal of biological chemistry*, 281(10), pp.6498-510.

Barron, M.R. et al., 2005. Serum response factor, an enriched cardiac mesoderm obligatory factor, is a downstream gene target for Tbx genes. *The Journal of biological chemistry*, 280(12), pp.11816-28.

Beckett, D., Kovaleva, E. & Schatz, P.J., 1999. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein science : a publication of the Protein Society*, 8(4), pp.921-9.

Beddington, R.S., 1994. Induction of a second neural axis by the mouse node. *Development (Cambridge, England)*, 120(3), pp.613-20.

Beddington, R.S. & Robertson, E.J., 1999. Axis development and early asymmetry in mammals. *Cell*, 96(2), pp.195-209.

Beddington, R.S.P., Rashbass, P. & Wilson, V., 1992. Brachyury--a gene affecting mouse gastrulation and early organogenesis. *Development (Cambridge, England). Supplement*, 165, pp.157-65.

Belaguli, N.S., Schildmeyer, L. a & Schwartz, R J, 1997. Organization and myogenic restricted expression of the murine serum response factor gene. A role for autoregulation. *The Journal of biological chemistry*, 272(29), pp.18222-31.

Belo, J. a et al., 1997. Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mechanisms of development*, 68(1-2), pp.45-57.

Bertocchini, F. & Stern, C.D., 2002. The hypoblast of the chick embryo positions the primitive streak by antagonizing nodal signaling. *Developmental cell*, 3(5), pp.735-44.

Bertocchini, F. et al., 2004. Determination of embryonic polarity in a regulative system: evidence for endogenous inhibitors acting sequentially during primitive streak formation in the chick embryo. *Development (Cambridge, England)*, 131(14), pp.3381-90.

de Boer, E. et al., 2003. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, 100(13), pp.7480-5.

Buchwalter, G., Gross, C. & Wasylyk, Bohdan, 2004. Ets ternary complex transcription factors. *Gene*, 324, pp.1-14.

Bénazéraf, B. et al., 2010. A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. *Nature*, 466(7303), pp.248-52.

Camarata, T. et al., 2006. LMP4 regulates Tbx5 protein subcellular localization and activity. *The Journal of cell biology*, 174(3), pp.339-48.

Camarata, T. et al., 2010. Pdlim7 (LMP4) regulation of Tbx5 specifies zebrafish heart atrio-ventricular boundary and valve formation. *Developmental biology*, 337(2), pp.233-45.

Candia, a F. et al., 1997. Cellular interpretation of multiple TGF-beta signals: intracellular antagonism between activin/BVg1 and BMP-2/4 signaling mediated by Smads. *Development (Cambridge, England)*, 124(22), pp.4467-80.

Cano, a et al., 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature cell biology*, 2(2), pp.76-83.

Carver, E.A. et al., 2001. The Mouse Snail Gene Encodes a Key Regulator of the Epithelial-Mesenchymal Transition. *Society*, 21(23), pp.8184-8188.

Catala, M., Teillet, M. a & Le Douarin, N.M., 1995. Organization and development of the tail bud analyzed with the quail-chick chimaera system. *Mechanisms of development*, 51(1), pp.51-65.

Chang, J. et al., 2003. Inhibitory cardiac transcription factor, SRF-N, is generated by caspase 3 cleavage in human heart failure and attenuated by ventricular unloading. *Circulation*, 108(4), pp.407-13.

Chang, P.S. et al., 2001. Muscle specificity encoded by specific serum response factor-binding sites. *The Journal of biological chemistry*, 276(20), pp.17206-12.

Chapman, D.L. & Papaioannou, V.E., 1998. Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature*, 391(6668), pp.695-7.

Chapman, N.R. & Perkins, N.D., 2000. Inhibition of the RelA (p65) NF- k B Subunit by Egr-1. *Biochemistry*, 275(7), pp.4719 -4725.

Christ B & Ordahl CP, 1995. Early stages of chick somite development. *Anat Embryol (Berl).*, 191(5), pp.381-96.

Chuai, M. & Weijer, C.J., 2009. Regulation of cell migration during chick gastrulation. *Current Opinion in Genetics & Development*, 19, pp.343-349.

Chung, H. a et al., 2005. FGF signal regulates gastrulation cell movements and morphology through its target NRH. *Developmental biology*, 282(1), pp.95-110.

Ciruna, B. & Rossant, J, 2001. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell*, 1, pp.37-49.

Conlon, F.L. et al., 1994. A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development (Cambridge, England)*, 120(7), pp.1919-28.

Cooper, S.J. et al., 2007. Serum response factor binding sites differ in three human cell types. *Genome Research*, 1, pp.136-144.

Croissant, J.D. et al., 1996. Avian Serum Response Factor Expression Restricted Primarily to Muscle Cell Lineages Is Required for a-Actin Gene Transcription. *Dev. Biol.*, 264(0160), pp.250 -264.

Cunningham, T.J., Zhao, X. & Duester, G., 2011. Uncoupling of retinoic acid signaling from tailbud development before termination of body axis extension. *Genesis*, 49(10), pp.776–783.

Dale, J.K. et al., 2006. Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. *Developmental cell*, 10(3), pp.355-66.

Davis, F.J. et al., 2002. Increased expression of alternatively spliced dominant-negative isoform of SRF in human failing hearts. *American journal of physiology. Heart and circulatory physiology*, 282(4), pp.H1521-33.

Delfini, M.-C. et al., 2005. Control of the segmentation process by graded MAPK/ERK activation in the chick embryo. *Proceedings of the National Academy of Sciences of the United States of America*, 102(32), pp.11343-8.

Diez, R. et al., 2003. Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern , Neuronal Differentiation , and Segmentation during Body Axis Extension. *Control*, 40, pp.65-79.

Diz-Muñoz, A. et al., 2010. Control of directed cell migration in vivo by membrane-to-cortex attachment. *PLoS biology*, 8(11), p.e1000544.

Driegen, S. et al., 2005. A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic Res*, 14(4), pp.477-482.

Du, P., Kibbe, W. a & Lin, S.M., 2008. lumi: a pipeline for processing Illumina microarray. *Bioinformatics (Oxford, England)*, 24(13), pp.1547-8.

Duboc, V. & Logan, M.P.O., 2011. Regulation of limb bud initiation and limb-type morphology. *Developmental dynamics : an official publication of the American Association of Anatomists*, 240(5), pp.1017-27.

Dubrulle, J. & Pourquié, O., 2004. Coupling segmentation to axis formation. *Development (Cambridge, England)*, 131(23), pp.5783-93.

Eakin, G.S. & Hadjantonakis, A.-K., 2006. Production of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. *Nat Protocols*, 1(3), pp.1145-1153.

Ernst, W.H. et al., 1995. Transcriptional repression mediated by the serum response factor. *FEBS letters*, 357(1), pp.45-9.

Feller, S.M., 2001. Crk family adaptors-signalling complex formation and biological roles. *Oncogene*, 20(44), pp.6348-71.

George, E.L. et al., 1993. Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development (Cambridge, England)*, 119(4), pp.1079-91.

Georges-Labouesse, E.N. et al., 1996. Mesodermal development in mouse embryos mutant for fibronectin. *Developmental dynamics : an official publication of the American Association of Anatomists*, 207(2), pp.145-56.

Gineitis, D. & Treisman, Richard, 2001. Differential usage of signal transduction pathways defines two types of serum response factor target gene. *The Journal of biological chemistry*, 276(27), pp.24531-9.

Girós, A. et al., 2011. α5β1 Integrin-Mediated Adhesion to Fibronectin Is Required for Axis Elongation and Somitogenesis in Mice. *PloS one*, 6(7), p.e22002.

Gloushankova, N. a et al., 1997. Cell-cell contact changes the dynamics of lamellar activity in nontransformed epitheliocytes but not in their ras-transformed descendants. *Proceedings of the National Academy of Sciences of the United States of America*, 94(3), pp.879-83.

Guaita, S. et al., 2002. Snail Induction of Epithelial to Mesenchymal Transition in Tumor Cells Is Accompanied by MUC1 Repression and ZEB1 Expression. *Biochemistry*, 277(42), pp.39209 - 39216.

Guettler, S. et al., 2008. RPEL motifs link the serum response factor cofactor MAL but not myocardin to Rho signaling via actin binding. *Molecular and cellular biology*, 28(2), pp.732-42.

Guillemin, K. et al., 1996. The pruned gene encodes the Drosophila serum response factor and regulates cytoplasmic outgrowth during terminal branching of the tracheal system. *Development (Cambridge, England)*, 122(5), pp.1353-62.

Guy, P.M., Kenny, D. a & Gill, G.N., 1999. The PDZ domain of the LIM protein enigma binds to beta-tropomyosin. *Molecular biology of the cell*, 10(6), pp.1973-84.

Hammerschmidt, M. & Wedlich, D., 2008. Regulated adhesion as a driving force of gastrulation movements. *Development*, 3641, pp.3625-3641.

Hashimoto, K., Fujimoto, H. & Nakatsuji, N., 1987. An ECM substratum allows mouse mesodermal cells isolated from the primitive streak to exhibit motility similar to that inside the embryo and reveals a deficiency in the T/T mutant cells. *Development (Cambridge, England)*, 100(4), pp.587-98.

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

Hinson, J.S. et al., 2009. Regulation of myocardin factor protein stability by the LIM-only protein FHL2. *Am J Physiol Heart Circ Physiol*, 295, pp.1067-1075.

Huang et al., 1997. DECREASED Egr-1 EXPRESSION IN HUMAN , MOUSE AND RAT MAMMARY CELLS AND TISSUES CORRELATES WITH TUMOR FORMATION. *Int. J. Cancer*, 109(October 1996), pp.102-109.

Huang, Q.-Q. et al., 2008. Role of H2-calponin in regulating macrophage motility and phagocytosis. *The Journal of biological chemistry*, 283(38), pp.25887-99.

Huelsken, J. et al., 2000. Requirement for beta-Catenin in Anterior-Posterior Axis Formation in Mice. *Cell*, 148(3), pp.567-578.

Ikenouchi, J. et al., 2003. Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *Journal of cell science*, 116(Pt 10), pp.1959-67.

Inman, K.E. & Downs, K.M., 2006. Localization of Brachyury (T) in embryonic and extraembryonic tissues during mouse gastrulation. *Gene expression patterns : GEP*, 6(8), pp.783-93.

Ivey, K. et al., 2008. MicroRNA Regulation of Cell Lineages in Mouse and Human Embryonic Stem Cells. *Cell*, (March), pp.219-229.

Iyer, V.R., 1999. The Transcriptional Program in the Response of Human Fibroblasts to Serum. *Science*, 283(5398), pp.83-87.

Jaffe, A.B. & Hall, A., 2005. Rho GTPases: biochemistry and biology. *Annual review of cell and developmental biology*, 21, pp.247-69.

Johannessen, M. et al., 2006. The multifunctional roles of the four-and-a-half-LIM only protein FHL2. *Cellular and molecular life sciences : CMLS*, 63(3), pp.268-84.

Johnson, D.S., 2011. Genome-Wide Mapping of in Vivo. *Science*, 1497(2007).

Jung, C.-rok et al., 2010. Enigma negatively regulates p53 through MDM2 and promotes tumor cell survival in mice. *Gene*, 120(12).

Kasza, A. et al., 2005. The ETS domain transcription factor Elk-1 regulates the expression of its partner protein, SRF. *The Journal of biological chemistry*, 280(2), pp.1149-55.

Kato, A. et al., 2008. Critical roles of actin-interacting protein 1 in cytokinesis and chemotactic migration of mammalian cells. *The Biochemical journal*, 414(2), pp.261-70.

Kim, J. et al., 2009. Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. *Nature protocols*, 4(4), pp.506-17.

Kinder, S.J. et al., 1999. The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development (Cambridge, England)*, 126(21), pp.4691-701.

Kislauskis, E.H., Zhu, X. & Singer, R.H., 1997. beta-Actin messenger RNA localization and protein synthesis augment cell motility. *The Journal of cell biology*, 136(6), pp.1263-70.

Kispert, A., Koschorz, B. & Herrmann, B.G., 1995. The T protein encoded by Brachyury is specific transcription factor. *EMBO Journal*, 14(19), pp.4763-4772.

Kmita, M. & Duboule, D., 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science (New York, N.Y.)*, 301(5631), pp.331-3.

Knöll, B., 2010. Actin-mediated gene expression in neurons: the MRTF-SRF connection. *Biological chemistry*, 391(6), pp.591-7.

Knöll, B. & Nordheim, A., 2009. Functional versatility of transcription factors in the nervous system: the SRF paradigm. *Cell*.

Kolodziej, K.E. et al., 2009. Optimal use of tandem biotin and V5 tags in ChIP assays. *BMC Molecular Biology*.

Kulisz, A. & Simon, H.-G., 2008. An evolutionarily conserved nuclear export signal facilitates cytoplasmic localization of the Tbx5 transcription factor. *Molecular and cellular biology*, 28(5), pp.1553-64.

Lacayo, C.I. et al., 2007. Emergence of large-scale cell morphology and movement from local actin filament growth dynamics. *PLoS biology*, 5(9), p.e233.

Laird, P.W. et al., 1991. Simplified mammalian DNA isolation procedure. *Nucleic acids research*, 19(15), p.4293.

Latinkic, B.V., Zeremski, M. & Lau, L.F., 1996. Elk-1 can recruit SRF to form a ternary complex upon the serum response element. *Nucleic Acids Research*, 24(7), pp.1345-1350.

Lawson, K. a, Meneses, J.J. & Pedersen, R. a, 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development (Cambridge, England)*, 113(3), pp.891-911.

Leblanc, V., Tocque, B. & Delumeau, I., 1998. Ras-GAP controls Rho-mediated cytoskeletal reorganization through its SH3 domain. *Molecular and cellular biology*, 18(9), pp.5567-78.

Lee, J.M. et al., 2006. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *The Journal of cell biology*, 172(7), pp.973-81.

Lemaire, P. et al., 1990. The Serum-Inducible Mouse Gene Krox-24 Encodes Specific Transcriptional Activator Sequence-. *Gene*, 10(7), pp.3456-3467.

Li, Shijie et al., 2006. Requirement of a myocardin-related transcription factor for development of mammary myoepithelial cells. *Molecular and cellular biology*, 26(15), pp.5797-808.

Liu, P. et al., 1999. Requirement for Wnt3 in vertebrate axis formation. *Nature genetics*, 22(4), pp.361-5.

Luxenburg, C. et al., 2011. Developmental roles for Srf, cortical cytoskeleton and cell shape in epidermal spindle orientation. *Nature cell biology*, 13(3), pp.203-14.

Maira, S.M., Wurtz, J.M. & Wasylyk, B, 1996. Net (ERP/SAP2) one of the Ras-inducible TCFs, has a novel inhibitory domain with resemblance to the helix-loop-helix motif. *The EMBO journal*, 15(21), pp.5849-65.

Majesky, M.W., 2003. Decisions, decisions...SRF coactivators and smooth muscle myogenesis. *Circulation research*, 92(8), pp.824-6.

Marikawa, Y. et al., 2009. Aggregated P19 Mouse Embryonal Carcinoma Cells as a Simple In Vitro Model to Study the Molecular Regulations of Mesoderm Formation and Axial Elongation Morphogenesis. *Genesis*, 47, pp.93-106.

Martin, B.L. & Kimelman, David, 2008. Regulation of Canonical Wnt Signaling by Brachyury Is Essential for Posterior Mesoderm Formation. *Developmental Cell*, (July), pp.121-133.

Martin, B.L. & Kimelman, David, 2010. Brachyury establishes the embryonic mesodermal progenitor niche. *Genes & development*, 24(24), pp.2778-83.

McHardy, L.M. et al., 2005. Strongylophorine-26, a Rho-dependent inhibitor of tumor cell invasion that reduces actin stress fibers and induces nonpolarized lamellipodial extensions. *Molecular cancer therapeutics*, 4(5), pp.772-8.

Mcburney, M., 1993. P19 embryonal carcinoma cells. *Differentiation*, 140, pp.135-140.

Miano, J.M., 2008. Deck of CArGs. *Circulation research*, 103(1), pp.13-5.

Miano, J.M., 2010. Role of serum response factor in the pathogenesis of disease. *Laboratory investigation; a journal of technical methods and pathology*, 90(9), pp.1274-1284.

Miano, J.M., 2003. Serum response factor: toggling between disparate programs of gene expression. *Journal of Molecular and Cellular Cardiology*, 35(6), pp.577-593.

Miano, J.M., Long, X. & Fujiwara, K., 2007. Serum response factor: master regulator of the actin cytoskeleton and contractile apparatus. *American journal of physiology. Cell physiology*, 292(1), pp.C70-81.

Miano, J.M. et al., 2004. Restricted inactivation of serum response factor to the cardiovascular system. *Proceedings of the National Academy of Sciences of the United States of America*, 101(49), pp.17132-7.

Miralles, F. et al., 2003. Actin Dynamics Control SRF Activity by Regulation of Its Coactivator MAL. *Cell*, 113, pp.329-342.

Misra, R P et al., 1991. The serum response factor is extensively modified by phosphorylation following its synthesis in serum-stimulated fibroblasts. *Molecular and cellular biology*, 11(9), pp.4545-54.

Mohun, T.J. et al., 1991. Expression of genes encoding the transcription factor SRF during early development of Xenopus laevis: identification of a CArG box-binding activity as SRF. *The EMBO journal*, 10(4), pp.933-40.

Murray, S.A., Carver, E.A. & Gridley, T., 2006. Generation of a Snail1 (Snai1) Conditional Null Allele. *Genesis*, 44, pp.7-11.

Müller, J.M. et al., 2002. The transcriptional coactivator FHL2 transmits Rho signals from the cell membrane into the nucleus. *The EMBO journal*, 21(4), pp.736-48.

Nagamune, T. & Ueno, N., 2004. Screening of FGF target genes in Xenopus by microarray : temporal dissection of the signalling pathway using a chemical inhibitor. *Genes to Cells*, pp.749-761.

Nakaya, Y. & Sheng, G., 2008. Epithelial to mesenchymal transition during gastrulation: an embryological view. *Development, growth & differentiation*, 50(9), pp.755-66.

Nentwich, O. et al., 2009. Downstream of FGF during mesoderm formation in Xenopus: the roles of Elk-1 and Egr-1. *Developmental biology*, 336(2), pp.313-26.

Nieuwkoop PD., 1973. The organization center of the amphibian embryo: its origin, spatial organization, and morphogenetic action. *Adv. Morphog.*, 10, pp.1-39.

Nishihara, H. et al., 2002. Molecular and immunohistochemical analysis of signaling adaptor protein Crk in human cancers. *Science And Technology*, 180, pp.55-61.

Niu, Z. et al., 2007. Serum Response Factor Micro-Managing Cardiogenesis. *Cell*, 19(6), pp.618-627.

Nostro, M.C. et al., 2008. Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. *Cell stem cell*, 2(1), pp.60-71.

Oh, J., Richardson, J. a & Olson, Eric N, 2005. Requirement of myocardin-related transcription factor-B for remodeling of branchial arch arteries and smooth muscle differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(42), pp.15122-7.

Ohta, S. et al., 2007. Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. *Development (Cambridge, England)*, 134(24), pp.4315-24.

Olson, Eric N & Nordheim, A., 2010. Linking actin dynamics and gene transcription to drive cellular motile functions. *Nature Reviews Molecular Cell Biology*, 11(5), pp.353-365.

Panitz, F. et al., 1998. The Spemann organizer-expressed zinc finger gene Xegr-1 responds to the MAP kinase/Ets-SRF signal transduction pathway. *EMBO Journal*, 17(15), pp.4414-4425.

Parameswaran, M. & Tam, P.P.L., 1995. Regionalisation of cell fate and morphogenetic movement of the mesoderm during mouse gastrulation. *Dev Genet.*, 17(1), pp.16-28.

Parlakian, A. et al., 2005. Temporally controlled onset of dilated cardiomyopathy through disruption of the SRF gene in adult heart. *Circulation*, 112(19), pp.2930-9.

Parlakian, A. et al., 2004. Targeted Inactivation of Serum Response Factor in the Developing Heart Results in Myocardial Defects and Embryonic Lethality. *Society*, 24(12), pp.5281-5289.

Perantoni, A.O. et al., 2005. Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development*, 132, pp.3859-3871.

Perea-Gomez, A. et al., 2002. Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Developmental cell*, 3(5), pp.745-56.

Philippar, U. et al., 2004. The SRF Target Gene Fhl2 Antagonizes RhoA/MAL-Dependent Activation of SRF. *Molecular Cell*, 16, pp.867-880.

Pipes, G.C.T., Creemers, E.E. & Olson, Eric N, 2006. The myocardin family of transcriptional coactivators: versatile regulators of cell growth, migration, and myogenesis. *Genes & development*, 20(12), pp.1545-56.

Poelmann, R.E., 1980. Differential mitosis and degeneration patterns in relation to the alterations in the shape of the embryonic ectoderm of early post-implantation mouse embryos. *Journal of embryology and experimental morphology*, 55, pp.33-51.

Pollard, T.D. & Cooper, J. a, 2009. Actin, a central player in cell shape and movement. *Science (New York, N.Y.)*, 326(5957), pp.1208-12.

Poser, S. et al., 2000. SRF-dependent gene expression is required for PI3-kinase-regulated cell proliferation. *The EMBO journal*, 19(18), pp.4955-66.

Posern, G. & Treisman, Richard, 2006. Actin' together: serum response factor, its cofactors and the link to signal transduction. *Trends in Cell Biology*, 16(11).

Pourquié, O. et al., 2007. FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development*, 4041, pp.4033-4041.

Purcell, N.H. et al., 2004. Extracellular Signal-Regulated Kinase 2 Interacts with and Is Negatively Regulated by the LIM-Only Protein FHL2 in Cardiomyocytes. *Society*, 24(3).

Qin, J. et al., 2011. ChIP-Array: combinatory analysis of ChIP-seq/chip and microarray gene expression data to discover direct/indirect targets of a transcription factor. *Nucleic acids research*, 39 Suppl 2(May), pp.W430-6.

Ramírez-Solis R, Davis, A. & Bradley A, 1993. Gene targeting in embryonic stem cells. *Methods Enzymol*, (225), pp.855-78.

Rashbass, P. et al., 1991. A cell autonomous function of Brachyury in T/T embryonic stem cell chimaeras. *Group*.

Reinke, H. et al., 2008. Differential display of DNA-binding proteins reveals heat-shock factor 1 as a circadian transcription factor. *Genes & development*, 22(3), pp.331-45.

Ridley, A.J., 2011. Life at the Leading Edge. *Cell*, 145(7), pp.1012-1022.

Rigaut, G. et al., 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology*, 17(10), pp.1030-2.

Robertson, G. et al., 2007. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods*, 4(8), pp.651-657.

Sakai, Y. et al., 2001. The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. *Genes & Development*, pp.213-225.

Schlesinger, J. et al., 2011. The cardiac transcription network modulated by Gata4, Mef2a, Nkx2.5, Srf, histone modifications, and microRNAs. *PLoS genetics*, 7(2), p.e1001313.

Schoenwolf, G.C., Garcia-Martinez, V. & Dias, M.S., 1992. Mesoderm movement and fate during avian gastrulation and neurulation. *Developmental dynamics : an official publication of the American Association of Anatomists*, 193(3), pp.235-48.

Schratt, G., Philippar, U., et al., 2002. Serum response factor is crucial for actin cytoskeletal organization and focal adhesion assembly in embryonic stem cells. *The Journal of cell biology*, 156(4), pp.737-50.

Schratt, G. et al., 2004. SRF regulates Bcl-2 expression and promotes cell survival during murine embryonic development. *EMBO Journal*, 23, pp.1834-1844.

Schratt, G., Wang, D.Z., et al., 2002. Potentiation of serum response factor activity by a family of myocardin-related transcription factors. *Proc Natl Acad Sci U S A*, 99, pp.14855-60.

Sedeh, R.S. et al., 2010. Structure, Evolutionary Conservation, and Conformational Dynamics of Homo sapiens Fascin-1, an F-actin Crosslinking Protein. *Journal of Molecular Biology*, 400(3), pp.589-604.

Selvaraj, A & Prywes, R., 2004. Expression profiling of serum inducible genes identifies a subset of SRF target genes that are MKL dependent.

Selvaraj, Ahalya & Prywes, R., 2003. Megakaryoblastic Leukemia-1/2, a Transcriptional Co-activator of Serum Response Factor , Is Required for Skeletal Myogenic Differentiation. *October*, 278(43), pp.41977-41987.

Sepulveda, J.L. et al., 2002. Combinatorial expression of GATA4, Nkx2-5, and serum response factor directs early cardiac gene activity. *The Journal of biological chemistry*, 277(28), pp.25775-82.

Shankaranarayanan, P. et al., 2011. single-tube linear dna amplification (Linda) for robust chiP-seq. *Online*, 8(7), pp.565-568.

Sharrocks, A.D., 1995. ERK2/p42 MAP kinase stimulates both autonomous and SRF-dependent DNA binding by Elk-1. *FEBS letters*, 368(1), pp.77-80.

Shaw, P.E., Schröter, H. & Nordheim, a, 1989. The ability of a ternary complex to form over the serum response element correlates with serum inducibility of the human c-fos promoter. *Cell*, 56(4), pp.563-72.

Snow, M.H. & Bennett, D., 1978. Gastrulation in the mouse: assessment of cell populations in the epiblast of tw18/tw18 embryos. *Journal of embryology and experimental morphology*, 47, pp.39-52.

Soler, E., Andrieu-Soler, C., et al., 2010. A systems approach to analyze transcription factors in mammalian cells. *Methods (San Diego, Calif.)*, 53(2), pp.151-162.

Soler, E., Andrieu-soler, C., et al., 2010. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes & Development*, pp.277-289.

Spencer, J.A. & Misra, Ravi P, 1999. Expression of the SRF gene occurs through a Ras/Sp/SRF-mediated-mechanism in response to serum growth signals. *Oncogene*, 18, pp.7319-7327.

Stemple, D.L., 2005. Structure and function of the notochord: an essential organ for chordate development. *Development (Cambridge, England)*, 132(11), pp.2503-12.

Sullivan, A.L. et al., 2011. Serum response factor utilizes distinct promoter- and enhancer-based mechanisms to regulate cytoskeletal gene expression in macrophages. *Molecular and cellular biology*, 31(4), pp.861-75.

Sun, K. et al., 2009. body growth in mice. , 49(5), pp.1645-1654.

Sun, Q. et al., 2006. Defining the mammalian CArGome. *Genome research*, 16(2), pp.197-207.

Sutherland, D., Samakovlis, C. & Krasnow, M. a, 1996. branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching. *Cell*, 87(6), pp.1091-101.

Takada, S et al., 1994. Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes & Development*, 8(2), pp.174-189.

Tallquist, M.D. & Soriano, P., 2000. Epiblast-restricted Cre expression in MORE mice: a tool to distinguish embryonic vs. extra-embryonic gene function. *Genesis (New York, N.Y.: 2000)* 26(2), pp.113-5.

Tam, P.P.L. & Beddington, R.S.P., 1987. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development (Cambridge, England)*, 99(1), pp.109-26.

Terpe, K., 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology*, 60(5), pp.523-33.

Treisman, Richard, 1986. Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors. *Cell*, 46(4), pp.567-74.

Treisman, Richard, 1996. Regulation of transcription by MAP kinase cascades. *Current opinion in cell biology*, 8(2), pp.205-15.

Treisman, Richard, 1987. Identification and purification of a polypeptide that binds to the c-fos serum response element. *The EMBO journal*, 6(9), pp.2711-7.

Tullai, J.W. et al., 2004. Identification of transcription factor binding sites upstream of human genes regulated by the phosphatidylinositol 3-kinase and MEK/ERK signaling pathways. *The Journal of biological chemistry*, 279(19), pp.20167-77.

Vidigal, J. a et al., 2010. An inducible RNA interference system for the functional dissection of mouse embryogenesis. *Nucleic acids research*, 38(11), p.e122.

Virolle, T. et al., 2003. Egr1 promotes growth and survival of prostate cancer cells. Identification of novel Egr1 target genes. *The Journal of biological chemistry*, 278(14), pp.11802-10.

Wang, C. et al., 2005. Egr-1 negatively regulates expression of the sodium-calcium exchanger-1 in cardiomyocytes in vitro and in vivo. *Cardiovascular research*, 65(1), pp.187-94.

Watson, D.K. et al., 1997. FLI1 and EWS-FLI1 function as ternary complex factors and ELK1 and SAP1a function as ternary and quaternary complex factors on the Egr1 promoter serum response elements. *Oncogene*, 14(2), pp.213-21.

Wehn, A.K., Gallo, P.H. & Chapman, D.L., 2009. Generation of transgenic mice expressing Cre recombinase under the control of the Dll1 mesoderm enhancer element. *Genesis (New York, N.Y.: 2000)*47(5), pp.309-13.

Wei, Y. et al., 2003. Identification of the LIM protein FHL2 as a coactivator of beta-catenin. *The Journal of biological chemistry*, 278(7), pp.5188-94.

Weinhold, B. et al., 2000. Srf(-/ -) ES cells display non-cell autonomous impairment in mesodermal differentiation. *EMBO Journal*, 19, pp.5835-5844.

Weinstein, M., Yang, X. & Deng, C.-xia, 2000. Functions of mammalian Smad genes as revealed by targeted gene disruption in mice. *Cytokine & Growth Factor Reviews*, 11.

Wiebel, F. et al., 2002. Generation of Mice Carrying Conditional Knockout Alleles for the Transcription Factor SRF. *Nucleic Acids Research*, 32, pp.124 -126.

Wilkinson, Bhatt & Herrmann, B.G., 1990. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature*.

Wilson, V. & Beddington, R.S.P., 1996. Cell fate and morphogenetic movement in the late mouse primitive streak. *Mechanisms of development*, 55(1), pp.79-89.

Wilson, V. & Beddington, R.S.P., 1997. Expression of T protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. *Developmental biology*, 192(1), pp.45-58.

Wilson, V. et al., 1995. The T gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development (Cambridge, England)*, 121(3), pp.877-86.

Wilson, V., Rashbass, P. & Beddington, R.S.P., 1993. Chimeric analysis of T (Brachyury) gene function. *Development (Cambridge, England)*, 117(4), pp.1321-31.

Wixler, V. et al., 2000. The LIM-only protein DRAL/FHL2 binds to the cytoplasmic domain of several alpha and beta integrin chains and is recruited to adhesion complexes. *The Journal of biological chemistry*, 275(43), pp.33669-78.

Wycuff, D.R., Yanites, H.L. & Marriott, S.J., 2004. Identification of a functional serum response element in the HTLV-I LTR. *Virology*, 324(2), pp.540-53.

Yamaguchi, T P et al., 1994. Fgfr-1 Is Required for Embryonic Growth and Mesodermal Patterning During Mouse Gastrulation. *Genes & Development*, 8(24), pp.3032-3044.

Yamaguchi, T P, 2001. Heads or tails: Wnts and anterior-posterior patterning. *Current biology : CB*, 11(17), pp.R713-24.

Yamaguchi, Terry P et al., 1999. (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes & Development*, 13, pp.3185-3190.

Yanagisawa, K.O., Fujimoto, H. & Urushihara, H., 1981. Effects of the brachyury (T) mutation on morphogenetic movement in the mouse embryo. *Developmental biology*, 87(2), pp.242-8.

Yang, J.T. et al., 1999. Overlapping and independent functions of fibronectin receptor integrins in early mesodermal development. *Developmental biology*, 215(2), pp.264-77.

Yost, C. et al., 1996. The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. *Genes & Development*, 10(12), pp.1443-1454.

Zambrowicz, B.P. et al., 1997. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8), pp.3789-94.

Zang, C. et al., 2009. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics (Oxford, England)*, 25(15), pp.1952-8.

Zaromytidou, A.-I., Miralles, F. & Treisman, Richard, 2006. MAL and ternary complex factor use different mechanisms to contact a common surface on the serum response factor DNA-binding domain. *Molecular and cellular biology*, 26(11), pp.4134-48.

Zhang et al., 1990. Suppression of c-ras transformation by GTPase-activating protein. *Group*.

Zhang, Shu Xing et al., 2005. Identification of Direct Serum-response Factor Gene Targets during Me 2 SO-induced P19 Cardiac Cell Differentiation. *Biochemistry*, 280(19), pp.19115-19126.

Zhao, Yong, Samal, E. & Srivastava, D., 2005. Serum response factor regulates a muscle-speci c microRNA that targets Hand2 during cardiogenesis. *Nature*, 436(July).

Zohn, I.E. et al., 2006. p38 and a p38-Interacting Protein Are Critical for Downregulation of E-Cadherin during Mouse Gastrulation. *Cell*, 125, pp.957-969.

Ørom, U.A. et al., 2010. Long noncoding RNAs with enhancer-like function in human cells. *Cell*, 143(1), pp.46-58.

10. Appendix

10.1. Supplementary figures

Figure S.1: Fn$^{RGE/RGE}$ *mutant embryos* (left, image from Girós et al. 2011) *and* Srf$^{flex1/flex1}$; T_{streak} Cre *embryos (right) at E8.75. Both mutants display the same heart defect .The mutation in the extracellular matrix protein fibronectin (Fn$^{RGE/RGE}$), specifically blocks binding of α5β1 integrin to fibronectin, resulting in impaired migration of mesodermal cells, similar to what was observed in* Srf$^{flex1/flex1}$; T_{streak} Cre *embryos.*

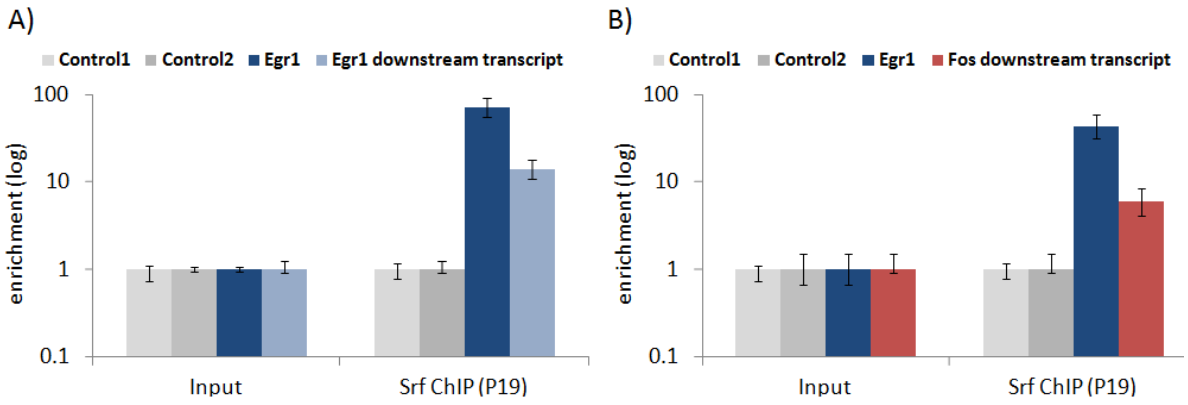

Figure S.2: *qPCR analysis of Srf-Avi[Bio] ChIP using chromatin from P19 cells confirmed binding of Srf to the transcripts that had been identified downstream of* Egr1 *(A) and* Fos *(B.) Srf peaks had been identified by ChIP-Seq experiments from P19 cells and embryonic caudal ends.*

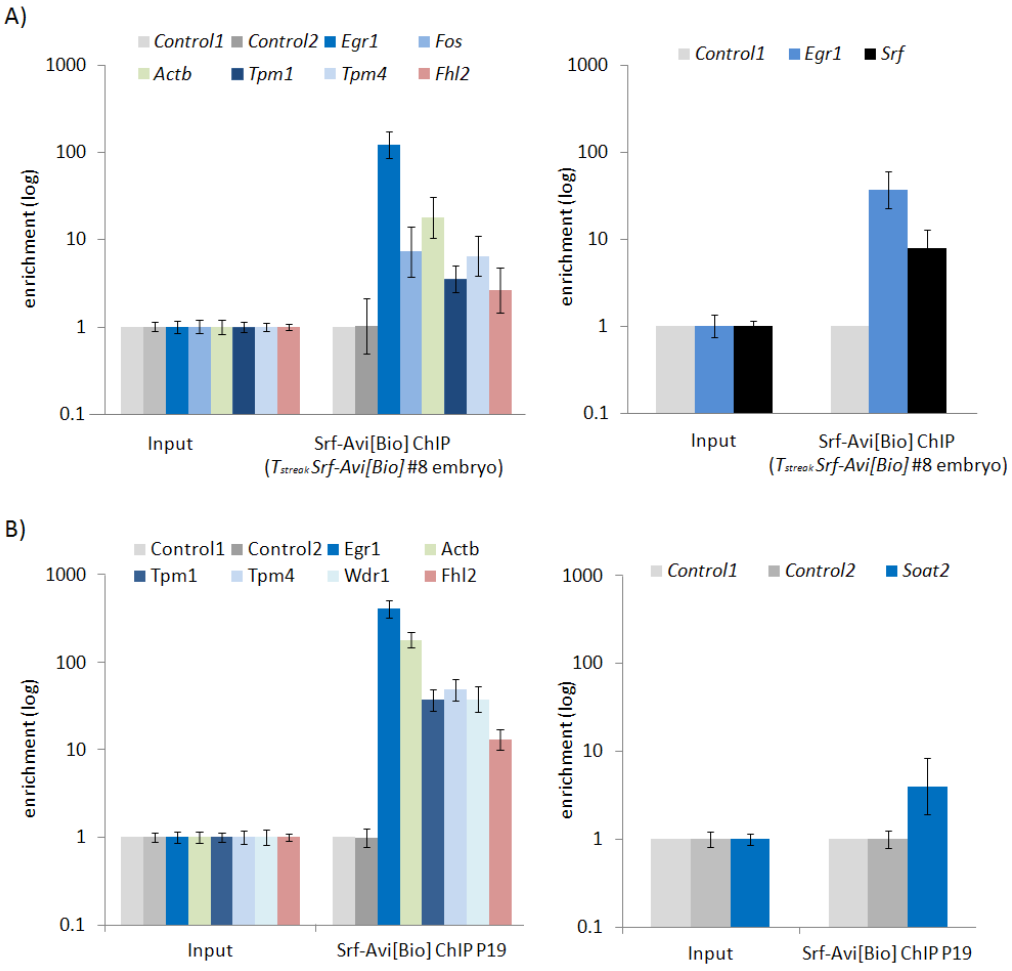

Figure S.3: *qPCR analyses of Srf-Avi[Bio] ChIP using chromatin from caudal ends from* T_{streak}Srf-Avi[Bio]*#8 embryos (A) or from P19 cells (Srf#1) (B) confirmed binding of Srf to the indicated target genes.*

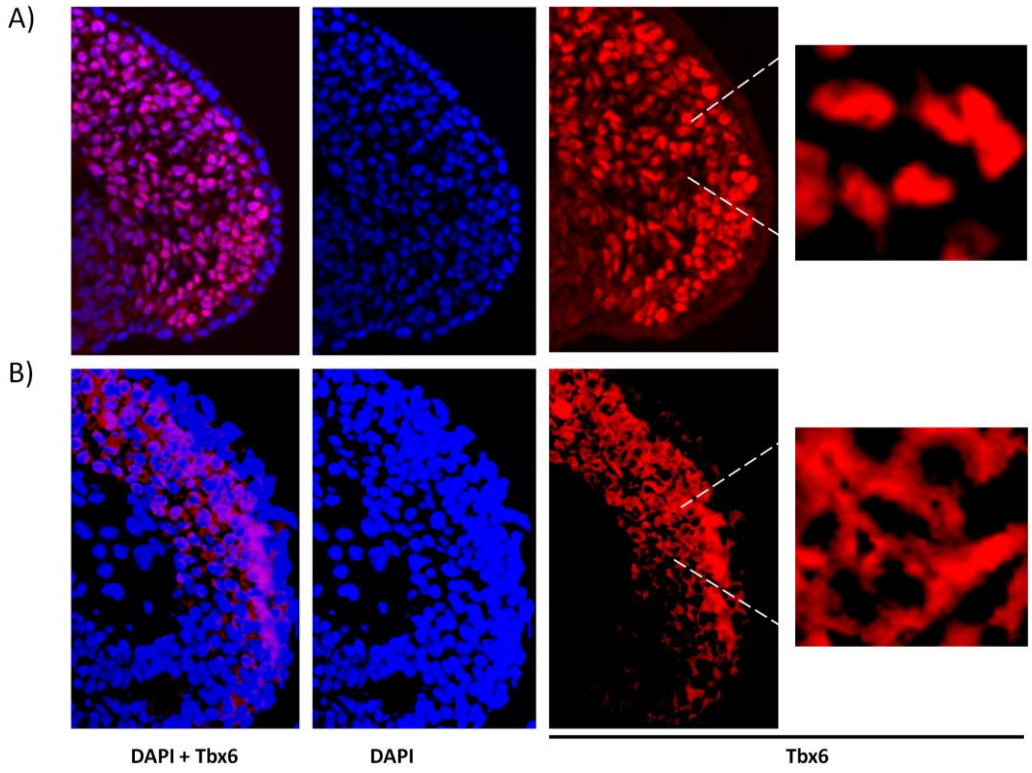

Figure S.4*: Immunofluorescence staining of embryonic caudal ends using a Tbx6-specific antibody. A) The transcription factor Tbx6 is usually found to have a nuclear localization in cells of embryonic caudal ends. B) In one* $Srf^{flex1/flex1}$; T_{streak} Cre *embryo Tbx6 was detected in the cytoplasm, while there is no staining in the nucleus.*

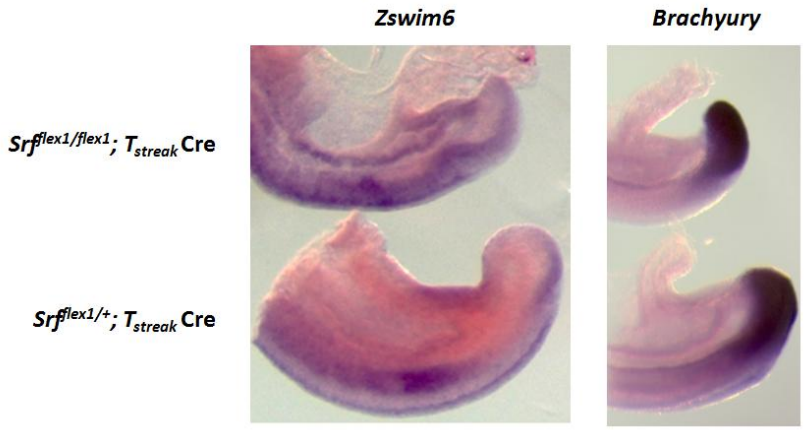

Figure S.5: In situ *hybridization on* $Srf^{flex1/flex1}$; T_{streak} Cre *embryos and heterozygous littermates using probes specific for* Zswim6 *or* Brachyury *revealed no observable changes in expression following loss of* Srf.

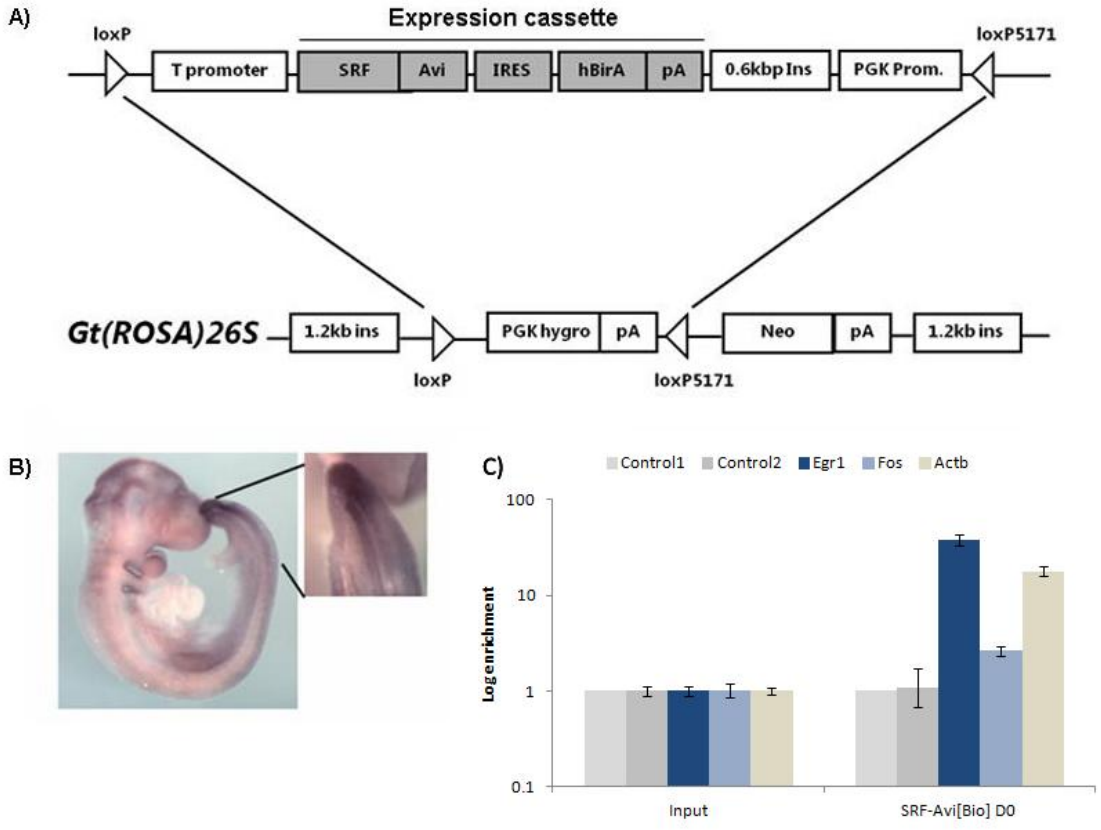

Figure S.6: *Improved specificity of transgene expression. A) Reconstructed donor and recipient constructs. The Tet responsive element was removed in the donor construct, the reverse tet silencer was removed in the recipient construct and a 1.2kb insulator was integrated into the 3'site of the Gt(ROSA)26S locus. Additionally the resistance genes were changed for a more efficient selection (compare with Fig.18). B) Whole mount* in situ *hybridization of transgenic embryos derived from* T_{streak}Srf-noTre-Avi[Bio]*#4 ES cells with a* hBirA*-specific probe reveals expression of the transgene that mainly resembles the* T_{streak} *promoter activity. Additional expression in the branchial arches and the brain has been previously noted for endogenous "leaky" transcriptional activity in the* Gt(ROSA)26S *locus. C) qPCR analysis of Srf-Avi[Bio] ChIP with chromatin from* T_{streak}Srf-noTre-Avi[Bio]*#4 ES cells differentiated into the mesodermal lineage. Enrichment of Srf target genes is comparably strong to ChIP experiments performed with the* in vitro *differentiated ES cell line* T_{streak}Srf-Avi[Bio]*#8.*

10.2. ChIP-Seq data

Table 10: *Srf peaks in P19 cells (identified by Srf ChIP-Seq); 183 genes (closest gene) were associated with one or several peaks when a high stringency threshold (fold enrichment ≥ 10; p-value ≤ 1x10^{-10}) was applied. The distance to the transcriptional start site (TSS) of the closest gene, the position of the peak and the enrichment over background are indicated.*

Chromosome	start	End	fold enrichment	distance to closest TSS (bp)	closest gene
chrX	84483520	84483929	46.74	435688	*BX088567.6*
chr13	108679807	108680924	36.71	135	*Zswim6*
chr19	8326567	8326948	36.14	10541	*D630002G06Rik*
chr18	40467558	40468108	35.54	49543	*Kctd16*
chrX	139917379	139917801	35.23	35755	*Pak3*
chr16	97191204	97191864	34.23	200679	*Dscam*
chr12	3109791	3110294	34.21	125720	*1700012B15Rik*
chr1	197067632	197068229	32.61	13160	*AC109145.11*
chr13	94073261	94073909	31.82	451	*Homer1*
chr19	61275425	61276123	31.6	15632	*EG547091*
chrX	73843921	73844639	31.43	4342	*4930468A15Rik*
chr19	53603109	53604150	31.2	529	*Dusp5*
chr6	103598905	103599356	30.6	6940	*Chl1*
chr3	5860464	5860997	30.59	20577	*AC137844.8*
chrX	109484294	109484865	29.85	24545	*Apool*
chr11	103128655	103129448	29.34	60	*Map3k14*
chr4	120288150	120289195	29	26801	*EG667063*
chr18	82390207	82390982	28.79	185962	*Galr1*
chr16	64884231	64884587	28.75	32320	*Cggbp1*
chr18	85866914	85867688	28.38	19005	*AC144861.3*
chr14	20110505	20110894	28.38	390	*AC165092.3*
chr17	13498667	13499282	28.28	16365	*AC166110.3*
chr11	118914451	118915270	27.49	12224	*Cbx8*
chr9	24346110	24347010	27.01	27856	*Dpy19l1*
chr15	79804202	79804587	26.87	41036	*Pdgfb*
chr8	87499881	87501049	26.71	2447	*Junb*
chr2	90235126	90235527	26.71	5693	*Olfr1274*
chr9	35112358	35113509	26.45	22938	*AC160116.2*
chr7	7230596	7231261	26.19	404	*Vmn2r29*
chr15	74916613	74917640	26.04	25484	*Ly6c2*
chr4	70038837	70039703	25.98	32564	*Cdk5rap2*
chr19	5488533	5489447	25.26	197	*Mus81*
chr13	77577777	77578485	25.14	270206	*1110033M05Rik*
chr11	69475597	69476374	25.12	476	*Mpdu1*
chr15	51697761	51698390	25	732	*Eif3h*
chr1	145549548	145550391	25	266	*AC175246.2*
chrX	138902399	138902829	25	17591	*SNORA17*
chrX	58372707	58373073	25	3146	*C230004F18Rik*
chrUn_random	5256151	5256558	25	421552	*Spt2*
chr14	52503000	52503464	25	1516	*Mett11d1*
chr9	110183312	110184308	24.42	24383	*2610002I17Rik*
chr16	42633780	42634333	24.36	242095	*AC161607.4*
chr2	45336212	45336746	24.29	151460	*AL773566.6*
chr4	124385926	124386867	24.06	6178	*Sf3a3*
chr17	36368031	36368929	23.95	5086	*CT030728.6-201*
chr2	23901395	23902065	23.68	2991	*Hnmt*

Chromosome	start	End	fold enrichment	distance to closest TSS (bp)	closest gene
chr18	61811543	61812721	23.53	2631	*mmu-mir-143*
chr3	102819209	102819876	23.49	5321	*Csde1*
chr7	19716565	19717360	23.48	11346	*Fbxo46*
chr7	28275521	28276501	23.3	3549	*Sertad1*
chr1	135989179	135990290	23.25	13447	*Btg2*
chr15	77735574	77736305	23.23	23824	*Txn2*
chr1	64167822	64168895	23.13	141	*Klf7*
chr14	21748008	21748858	23.03	646	*Vcl*
chr18	35019968	35021468	23.02	892	*Egr1*
chrX	163099284	163100078	22.77	75636	*Egfl6*
chr2	143740376	143741386	22.58	690	*Dstn*
chr5	114322457	114323492	22.5	36258	*Coro1c*
chr8	17653812	17654485	22.5	118427	*Csmd1*
chr11	108872420	108873898	22.41	3455	*AL732612.28*
chr5	147072398	147073099	22.41	12271	*AC113316.14*
chr11	110642613	110643598	22.22	186733	*Kcnj16*
chr12	86826504	86827576	21.8	11654	*Fos*
chr18	34268335	34269090	21.51	15236	*SNORA17*
chr8	48419184	48419987	21.43	18518	*Stox2*
chr10	17742052	17742533	21.43	948	*3110003A17Rik*
chr14	70474357	70475227	21.31	2894	*Egr3*
chr16	84045345	84046038	21.21	249435	*U6*
chr6	22367047	22367684	20.83	60917	*D6Wsu176e*
chr2	97250337	97250549	20.43	57493	*Lrrc4c*
chr17	23013238	23013904	20.43	9137	*A630033E08Rik*
chr16	22281845	22282579	20.43	13255	*AC154605.2*
chr6	85463316	85463947	20.09	220	*Egr4*
chr1_random	272315	273251	20.08	1908	*Aida*
chr19	5447048	5448143	19.75	499	*Fosl1*
chr8	87187463	87188561	19.71	712	*Ier2*
chr5	26333246	26334167	19.7	16210	*AC125090.3-202*
chr5	142074405	142075114	19.44	16160	*Sdk1*
chrX	148450732	148451452	19.35	106	*Ribc1*
chr2	45025081	45025795	19.32	52166	*Zeb2*
chr18	68851591	68852275	19.23	259591	*4930546C10Rik*
chr2	103401093	103401616	19.17	5373	*Abtb2*
chr7	77506372	77507199	19.14	893	*Nr2f2*
chr9	50505464	50506093	19.12	601	*Dixdc1*
chr9	31069828	31070546	19.09	12654	*EG638580*
chr13	100560710	100561512	18.99	98226	*SNORA17*
chr6	148588906	148589607	18.85	2271	*SNORA62*
chr3	68346527	68347255	18.83	5103	*U2*
chrUn_random	5475241	5476073	18.75	640642	*Spt2*
chr7	52252449	52253732	18.62	500	*Bcl2l12*
chr9	113635291	113635997	18.57	15299	*Clasp2*
chr1	174353570	174354191	18.52	46883	*Pigm*
chr15	78707012	78708008	18.46	607	*Gga1*
chr1	136720125	136720835	18.31	14273	*AC131591.13*
chr3	95462274	95463558	18.26	368	*Mcl1*
chr9	7238545	7239153	18.09	33983	*Mmp13*
chr11	88055017	88055955	18.08	37073	*Mrps23*
chr7	19895288	19896177	17.98	106	*Fosb*
chr16	91534273	91535905	17.91	6032	*AC159199.3*
chr1	95721550	95722543	17.86	266	*D2hgdh*
chr2	64876000	64876503	17.86	15169	*Grb14*
chr16	45743872	45744446	17.69	855	*Abhd10*

Chromosome	start	End	fold enrichment	distance to closest TSS (bp)	closest gene
chr16	4880274	4881069	17.5	456	*1500031H01Rik*
chr4	155628378	155629262	17.28	19433	*Klhl17*
chr1	47608107	47608805	17.28	184891	*AC102313.8*
chr7	144505768	144506677	17.24	360	*Ebf3*
chr10	41112496	41113497	17.16	4790	*U6*
chr8	25982441	25982899	17.05	47600	*Adam5*
chr11	3435215	3436529	17	2186	*Smtn*
chr5	143667957	143669036	16.94	446	*Actb*
chr11	104549660	104550730	16.89	2639	*AL645600.11*
chr11	53953075	53954337	16.81	11548	*AL596103.22*
chr11	87562475	87563585	16.67	7891	*mmu-mir-142*
chr5	38952443	38953463	16.67	391	*Wdr1*
chr19	22026529	22027038	16.67	14659	*AC129205.3*
chr18	13107328	13107897	16.46	7042	*Osbpl1a*
chr8	48760054	48761143	16.43	458	*Ing2*
chr5	143667124	143667913	16.4	237	*Actb*
chrUn_random	5248527	5250036	16.39	413928	*Spt2*
chr4	123444838	123445493	16.36	17250	*Akirin1*
chr11	67989735	67990477	16.27	80269	*AL663076.17*
chr18	3005159	3006508	16.08	118253	*LOC171266*
chr2	167361939	167364875	16.06	997	*AL589870.30*
chr2	181665247	181666150	16	33719	*AL928734.9*
chr6	124869652	124870547	15.96	1688	*Ptms*
chr14	21711115	21711729	15.94	35472	*AC154840.2*
chr9	65398606	65399663	15.93	1292	*Plekho2*
chr9	21904373	21904964	15.91	648	*Cnn1*
chr5	32437356	32438408	15.87	867	*Fosl2*
chr11	105937541	105938075	15.84	8381	*Map3k3*
chr4	145029478	145030291	15.74	7165	*AL606963.5*
chr7	126575441	126576220	15.71	27345	*Gp2*
chr18	37129090	37129894	15.71	1681	*Pcdha6*
chr4	145163783	145164386	15.71	648	*AL713973.7*
chr18	73692615	73693687	15.62	590	*5_8S_rRNA*
chr10	79450999	79451658	15.57	345	*Cnn2*
chr10	75257353	75258124	15.38	2961	*Gstt1*
chr3	153647315	153647832	15.32	10915	*Slc44a5*
chr12	81309440	81309866	15.28	17582	*Strm*
chr3	121304896	121305418	15.28	41507	*A730020M07Rik*
chr3	8245150	8246725	15.13	29275	*AC160969.3*
chr4	122563514	122564170	15.08	227	*Cap1*
chr13	3371960	3372980	15.07	165360	*Gdi2*
chr8	93649621	93650530	15.05	5429	*Aktip*
chr18	35046856	35047726	14.94	25995	*Egr1*
chr5	75213826	75214738	14.84	114890	*Lnx1*
chr10	21861979	21863166	14.81	6650	*E030030I06Rik*
chr9	91118827	91119603	14.71	126841	*AC152825.4*
chr3	145316216	145316716	14.67	3271	*Cyr61*
chr11	45900374	45901183	14.52	30787	*Adam19*
chr9	2999837	3001448	14.37	1233	*AC131780.5-209*
chr12	8895101	8895768	14.37	10836	*9930038B18Rik*
chr9	6726185	6726921	14.29	19739	*U6*
chr5	149630145	149630619	14.19	65129	*5S_rRNA*
chrUn_random	4320081	4320316	14.14	514518	*Spt2*
chr2	171634271	171634727	14.06	234598	*Cbln4*
chr6	5373129	5373928	14	32702	*Asb4*
chr10	67000013	67001991	13.89	603	*Egr2*

Chromosome	start	End	fold enrichment	distance to closest TSS (bp)	closest gene
chr11	101807639	101808232	13.87	6096	*4930417O22Rik*
chr7	89919670	89920912	13.86	16181	*5S_rRNA*
chr4	22284550	22285829	13.73	161	*Fbxl4*
chr12	106674166	106674859	13.57	127215	*Bdkrb2*
chr8	97126090	97127352	13.51	212	*Rspry1*
chr6	97557448	97558171	13.36	10128	*Frmd4b*
chr8	76142692	76143386	13.21	22712	*AC138317.11*
chr1	43220220	43221305	13.12	494	*Fhl2*
chr12	88228824	88229453	12.99	3063	*6430527G18Rik*
chr2	181651896	181652946	12.9	20368	*AL928734.9*
chr9	55889263	55890443	12.82	388	*Rcn2*
chr15	77095742	77096168	12.68	41293	*Rbm9*
chr17	46692619	46693705	12.5	492	*Srf*
chr17	34072219	34073334	12.5	5055	*H2-Ke2*
chr9_random	241853	243809	12.35	128810	*Nlrp4g*
chr9	118107622	118108112	12.26	21606	*AC159900.3*
chr10	80640565	80642241	11.9	1130	*SNORD37*
chr8	91094213	91094982	11.7	48944	*Nkd1*
chr6	30754328	30755379	11.67	63029	*mmu-mir-335*
chr3	3324358	3326677	11.61	183672	*Hnf4g*
chr12	56639290	56640067	11.61	26831	*Aldoart2*
chr14	58871857	58872920	11.54	177903	*Fgf9*
chrUn_random	2747645	2748905	11.54	11813	*Vmn2r123*
chr6	82662129	82663375	11.54	59261	*Pole4*
chr9	3035547	3036759	11.36	1564	*AC131780.5-202*
chr6	17631366	17632092	11.36	12581	*St7*
chr4	152405230	152406173	11.33	17537	*AL805898.11*
chr15	79905509	79906764	10.99	3512	*snoU83B*
chr2	17610957	17611802	10.91	41738	*Nebl*
chr5	135476177	135476928	10.91	1043	*Wbscr25*
chr11	94390201	94390918	10.89	7385	*Mycbpap*
chr6	29698175	29699293	10.55	2586	*AC069469.5*
chr14	70476296	70476907	10.43	955	*Egr3*
chr5	100779558	100780851	10.42	45289	*Sec31a*
chrUn_random	5495211	5497429	10.41	660612	*Spt2*
chrUn_random	4062457	4064448	10.34	772142	*Spt2*
chr11	75401098	75402307	10.22	500	*Pitpna*
chr11	25686567	25687416	10	15997	*5730522E02Rik*

Table 11: Srf peaks in P19 cells (identified by Srf-Avi[Bio] ChIP-Seq) *120 genes (closest gene) were associated with one or several peaks when a high stringency threshold (fold enrichment ≥ 10; p-value ≤ $1x10^{-10}$) was applied. The distance to the transcriptional start site (TSS) of the closest gene, the position of the peak and the enrichment over background are indicated.*

Chromosome	start	end	fold enrichment	distance to closest TSS (bp)	closest gene
chrX	84483520	84483929	46.74	435688	*BX088567.6*
chr13	108679807	108680924	36.71	135	*Zswim6*
chr19	8326567	8326948	36.14	10541	*D630002G06Rik*
chr18	40467558	40468108	35.54	49543	*Kctd16*
chrX	139917379	139917801	35.23	35755	*Pak3*
chr16	97191204	97191864	34.23	200679	*Dscam*
chr12	3109791	3110294	34.21	125720	*1700012B15Rik*
chr1	197067632	197068229	32.61	13160	*AC109145.11*
chr13	94073261	94073909	31.82	451	*Homer1*
chr19	61275425	61276123	31.6	15632	*EG547091*
chrX	73843921	73844639	31.43	4342	*4930468A15Rik*
chr19	53603109	53604150	31.2	529	*Dusp5*
chr6	103598905	103599356	30.6	6940	*Chl1*
chr3	5860464	5860997	30.59	20577	*AC137844.8*
chrX	109484294	109484865	29.85	24545	*Apool*
chr11	103128655	103129448	29.34	60	*Map3k14*
chr4	120288150	120289195	29	26801	*EG667063*
chr18	82390207	82390982	28.79	185962	*Galr1*
chr16	64884231	64884587	28.75	32320	*Cggbp1*
chr18	85866914	85867688	28.38	19005	*AC144861.3*
chr14	20110505	20110894	28.38	390	*AC165092.3*
chr17	13498667	13499282	28.28	16365	*AC166110.3*
chr11	118914451	118915270	27.49	12224	*Cbx8*
chr9	24346110	24347010	27.01	27856	*Dpy19l1*
chr15	79804202	79804587	26.87	41036	*Pdgfb*
chr8	87499881	87501049	26.71	2447	*Junb*
chr2	90235126	90235527	26.71	5693	*Olfr1274*
chr9	35112358	35113509	26.45	22938	*AC160116.2*
chr7	7230596	7231261	26.19	404	*Vmn2r29*
chr15	74916613	74917640	26.04	25484	*Ly6c2*
chr4	70038837	70039703	25.98	32564	*Cdk5rap2*
chr19	5488533	5489447	25.26	197	*Mus81*
chr13	77577777	77578485	25.14	270206	*1110033M05Rik*
chr11	69475597	69476374	25.12	476	*Mpdu1*
chr15	51697761	51698390	25	732	*Eif3h*
chr1	145549548	145550391	25	266	*AC175246.2*
chrX	138902399	138902829	25	17591	*SNORA17*
chrX	58372707	58373073	25	3146	*C230004F18Rik*
chrUn_random	5256151	5256558	25	421552	*Spt2*
chr14	52503000	52503464	25	1516	*Mett11d1*
chr9	110183312	110184308	24.42	24383	*2610002I17Rik*
chr16	42633780	42634333	24.36	242095	*AC161607.4*
chr2	45336212	45336746	24.29	151460	*AL773566.6*
chr4	124385926	124386867	24.06	6178	*Sf3a3*
chr17	36368031	36368929	23.95	5086	*CT030728.6-201*
chr2	23901395	23902065	23.68	2991	*Hnmt*
chr18	61811543	61812721	23.53	2631	*mmu-mir-143*
chr3	102819209	102819876	23.49	5321	*Csde1*
chr7	19716565	19717360	23.48	11346	*Fbxo46*
chr7	28275521	28276501	23.3	3549	*Sertad1*
chr1	135989179	135990290	23.25	13447	*Btg2*
chr15	77735574	77736305	23.23	23824	*Txn2*
chr1	64167822	64168895	23.13	141	*Klf7*
chr14	21748008	21748858	23.03	646	*Vcl*

Chromosome	start	end	fold enrichment	distance to closest TSS (bp)	closest gene
chr18	35019968	35021468	23.02	892	*Egr1*
chrX	163099284	163100078	22.77	75636	*Egfl6*
chr2	143740376	143741386	22.58	690	*Dstn*
chr5	114322457	114323492	22.5	36258	*Coro1c*
chr8	17653812	17654485	22.5	118427	*Csmd1*
chr11	108872420	108873898	22.41	3455	*AL732612.28*
chr5	147072398	147073099	22.41	12271	*AC113316.14*
chr11	110642613	110643598	22.22	186733	*Kcnj16*
chr12	86826504	86827576	21.8	11654	*Fos*
chr18	34268335	34269090	21.51	15236	*SNORA17*
chr8	48419184	48419987	21.43	18518	*Stox2*
chr10	17742052	17742533	21.43	948	*3110003A17Rik*
chr14	70474357	70475227	21.31	2894	*Egr3*
chr16	84045345	84046038	21.21	249435	*U6*
chr6	22367047	22367684	20.83	60917	*D6Wsu176e*
chr2	97250337	97250549	20.43	57493	*Lrrc4c*
chr17	23013238	23013904	20.43	9137	*A630033E08Rik*
chr16	22281845	22282579	20.43	13255	*AC154605.2*
chr6	85463316	85463947	20.09	220	*Egr4*
chr1_random	272315	273251	20.08	1908	*Aida*
chr19	5447048	5448143	19.75	499	*Fosl1*
chr8	87187463	87188561	19.71	712	*Ier2*
chr5	26333246	26334167	19.7	16210	*AC125090.3-202*
chr5	142074405	142075114	19.44	16160	*Sdk1*
chrX	148450732	148451452	19.35	106	*Ribc1*
chr2	45025081	45025795	19.32	52166	*Zeb2*
chr18	68851591	68852275	19.23	259591	*4930546C10Rik*
chr2	103401093	103401616	19.17	5373	*Abtb2*
chr7	77506372	77507199	19.14	893	*Nr2f2*
chr9	50505464	50506093	19.12	601	*Dixdc1*
chr9	31069828	31070546	19.09	12654	*EG638580*
chr13	100560710	100561512	18.99	98226	*SNORA17*
chr6	148588906	148589607	18.85	2271	*SNORA62*
chr3	68346527	68347255	18.83	5103	*U2*
chrUn_random	5475241	5476073	18.75	640642	*Spt2*
chr7	52252449	52253732	18.62	500	*Bcl2l12*
chr9	113635291	113635997	18.57	15299	*Clasp2*
chr1	174353570	174354191	18.52	46883	*Pigm*
chr15	78707012	78708008	18.46	607	*Gga1*
chr1	136720125	136720835	18.31	14273	*AC131591.13*
chr3	95462274	95463558	18.26	368	*Mcl1*
chr9	7238545	7239153	18.09	33983	*Mmp13*
chr11	88055017	88055955	18.08	37073	*Mrps23*
chr7	19895288	19896177	17.98	106	*Fosb*
chr16	91534273	91535905	17.91	6032	*AC159199.3*
chr1	95721550	95722543	17.86	266	*D2hgdh*
chr2	64876000	64876503	17.86	15169	*Grb14*
chr16	45743872	45744446	17.69	855	*Abhd10*
chr16	4880274	4881069	17.5	456	*1500031H01Rik*
chr4	155628378	155629262	17.28	19433	*Klhl17*
chr1	47608107	47608805	17.28	184891	*AC102313.8*
chr7	144505768	144506677	17.24	360	*Ebf3*
chr10	41112496	41113497	17.16	4790	*U6*
chr8	25982441	25982899	17.05	47600	*Adam5*
chr11	3435215	3436529	17	2186	*Smtn*
chr5	143667957	143669036	16.94	446	*Actb*
chr11	104549660	104550730	16.89	2639	*AL645600.11*
chr11	53953075	53954337	16.81	11548	*AL596103.22*
chr11	87562475	87563585	16.67	7891	*mmu-mir-142*
chr5	38952443	38953463	16.67	391	*Wdr1*

Chromosome	start	end	fold enrichment	distance to closest TSS (bp)	closest gene
chr19	22026529	22027038	16.67	14659	*AC129205.3*
chr18	13107328	13107897	16.46	7042	*Osbpl1a*
chr8	48760054	48761143	16.43	458	*Ing2*
chr5	143667124	143667913	16.4	237	*Actb*
chrUn_random	5248527	5250036	16.39	413928	*Spt2*
chr4	123444838	123445493	16.36	17250	*Akirin1*
chr11	67989735	67990477	16.27	80269	*AL663076.17*
chr18	3005159	3006508	16.08	118253	*LOC171266*
chr2	167361939	167364875	16.06	997	*AL589870.30*
chr2	181665247	181666150	16	33719	*AL928734.9*
chr6	124869652	124870547	15.96	1688	*Ptms*
chr14	21711115	21711729	15.94	35472	*AC154840.2*
chr9	65398606	65399663	15.93	1292	*Plekho2*
chr9	21904373	21904964	15.91	648	*Cnn1*
chr5	32437356	32438408	15.87	867	*Fosl2*
chr11	105937541	105938075	15.84	8381	*Map3k3*
chr4	145029478	145030291	15.74	7165	*AL606963.5*
chr7	126575441	126576220	15.71	27345	*Gp2*
chr18	37129090	37129894	15.71	1681	*Pcdha6*
chr4	145163783	145164386	15.71	648	*AL713973.7*
chr18	73692615	73693687	15.62	590	*5_8S_rRNA*
chr10	79450999	79451658	15.57	345	*Cnn2*
chr10	75257353	75258124	15.38	2961	*Gstt1*
chr3	153647315	153647832	15.32	10915	*Slc44a5*
chr12	81309440	81309866	15.28	17582	*Strm*
chr3	121304896	121305418	15.28	41507	*A730020M07Rik*
chr3	8245150	8246725	15.13	29275	*AC160969.3*
chr4	122563514	122564170	15.08	227	*Cap1*
chr13	3371960	3372980	15.07	165360	*Gdi2*
chr8	93649621	93650530	15.05	5429	*Aktip*
chr18	35046856	35047726	14.94	25995	*Egr1*
chr5	75213826	75214738	14.84	114890	*Lnx1*
chr10	21861979	21863166	14.81	6650	*E030030I06Rik*
chr9	91118827	91119603	14.71	126841	*AC152825.4*
chr3	145316216	145316716	14.67	3271	*Cyr61*
chr11	45900374	45901183	14.52	30787	*Adam19*
chr9	2999837	3001448	14.37	1233	*AC131780.5-209*
chr12	8895101	8895768	14.37	10836	*9930038B18Rik*
chr9	6726185	6726921	14.29	19739	*U6*
chr5	149630145	149630619	14.19	65129	*5S_rRNA*
chrUn_random	4320081	4320316	14.14	514518	*Spt2*
chr2	171634271	171634727	14.06	234598	*Cbln4*
chr6	5373129	5373928	14	32702	*Asb4*
chr10	67000013	67001991	13.89	603	*Egr2*
chr11	101807639	101808232	13.87	6096	*4930417O22Rik*
chr7	89919670	89920912	13.86	16181	*5S_rRNA*
chr4	22284550	22285829	13.73	161	*Fbxl4*
chr12	106674166	106674859	13.57	127215	*Bdkrb2*
chr8	97126090	97127352	13.51	212	*Rspry1*
chr6	97557448	97558171	13.36	10128	*Frmd4b*
chr8	76142692	76143386	13.21	22712	*AC138317.11*
chr1	43220220	43221305	13.12	494	*Fhl2*
chr12	88228824	88229453	12.99	3063	*6430527G18Rik*
chr2	181651896	181652946	12.9	20368	*AL928734.9*
chr9	55889263	55890443	12.82	388	*Rcn2*
chr15	77095742	77096168	12.68	41293	*Rbm9*
chr17	46692619	46693705	12.5	492	*Srf*
chr17	34072219	34073334	12.5	5055	*H2-Ke2*
chr9_random	241853	243809	12.35	128810	*Nlrp4g*
chr9	118107622	118108112	12.26	21606	*AC159900.3*

Chromosome	start	end	fold enrichment	distance to closest TSS (bp)	closest gene
chr10	80640565	80642241	11.9	1130	*SNORD37*
chr8	91094213	91094982	11.7	48944	*Nkd1*
chr6	30754328	30755379	11.67	63029	*mmu-mir-335*
chr3	3324358	3326677	11.61	183672	*Hnf4g*
chr12	56639290	56640067	11.61	26831	*Aldoart2*
chr14	58871857	58872920	11.54	177903	*Fgf9*
chrUn_random	2747645	2748905	11.54	11813	*Vmn2r123*
chr6	82662129	82663375	11.54	59261	*Pole4*
chr9	3035547	3036759	11.36	1564	*AC131780.5-202*
chr6	17631366	17632092	11.36	12581	*St7*
chr4	152405230	152406173	11.33	17537	*AL805898.11*
chr15	79905509	79906764	10.99	3512	*snoU83B*
chr2	17610957	17611802	10.91	41738	*Nebl*
chr5	135476177	135476928	10.91	1043	*Wbscr25*
chr11	94390201	94390918	10.89	7385	*Mycbpap*
chr6	29698175	29699293	10.55	2586	*AC069469.5*
chr14	70476296	70476907	10.43	955	*Egr3*
chr5	100779558	100780851	10.42	45289	*Sec31a*
chrUn_random	5495211	5497429	10.41	660612	*Spt2*
chrUn_random	4062457	4064448	10.34	772142	*Spt2*
chr11	75401098	75402307	10.22	500	*Pitpna*
chr11	25686567	25687416	10	15997	*5730522E02Rik*

10.3. Gene Expression data

Table 12: Dysregulated genes in $Srf^{flex1/flex1}$; T_{streak} Cre embryos; fold change ≥ $ln2^{0.4}$; p≤ 10^{-5}

	Gene symbol	Fold expr.	*p-value*		Gene symbol	Fold expr.	*p-value*
E8.5	*Downregulated genes*				*Tpm1*	*0.64*	*0.00030*
	Mocs1	*0.61*	*0.00002*		*Actb*	*0.68*	*0.00239*
	Acta2	*0.62*	*0.00493*		*Alas2*	*0.68*	*0.00408*
	Srf	*0.62*	*0.00001*		*Zfp101*	*0.68*	*0.00052*
	Wdr1	*0.68*	*0.00001*		*Gypa*	*0.71*	*0.02000*
	Pdlim7	*0.72*	*0.00041*		*Acta2*	*0.71*	*0.02146*
	Tpm1	*0.72*	*0.00255*		*Cited4*	*0.71*	*0.01162*
	Tpm4	*0.74*	*0.02100*		*Cnn2*	*0.72*	*0.00067*
E8.5	*Upregulated genes*				*Fhl2*	*0.72*	*0.00017*
	H2-Bl	*1.50*	*0.00526*		*Fgf17*	*0.72*	*0.03305*
	Mipep	*1.32*	*0.00662*		*Cdc20*	*0.72*	*0.03212*
	H2-T23	*1.32*	*0.00902*		*LOC100040573*	*0.73*	*0.00375*
E8.75	*Downregulated genes*				*Rbm13*	*0.73*	*0.03044*
	Ddx3y	*0.28*	*0.00336*		*D16Ertd472e*	*0.73*	*0.00049*
	Slc4a1	*0.50*	*0.03076*		*Rasa1*	*0.73*	*0.03479*
	Acta2	*0.53*	*0.01093*		*Tpm4*	*0.74*	*0.01052*
	Mocs1	*0.54*	*0.00002*		*Zyx*	*0.75*	*0.00839*
	Srf	*0.60*	*0.02061*		*BC003885*	*0.75*	*0.02075*
	Actb	*0.60*	*0.01115*		*Trp53*	*0.75*	*0.01480*
	Wdr1	*0.62*	*0.00010*		*Clptm1l*	*0.75*	*0.00937*
	Wdr1	*0.63*	*0.00759*	**E8.75**	*Upregulated genes*		
	Pdlim7	*0.64*	*0.00068*		*Apoa1*	*9.89*	*0.00070*

Gene symbol	Fold expr.	*p-value*	Gene symbol	Fold expr.	*p-value*
Ttr	*8.85*	*0.00031*	*LOC381629*	*1.46*	*0.00443*
Mt1	*5.37*	*0.00118*	*4933439C20Rik*	*1.46*	*0.00072*
Apoa2	*4.36*	*0.00533*	*Slc39a4*	*1.45*	*0.00123*
Spink3	*4.07*	*0.00106*	*2310016C08Rik*	*1.44*	*0.02668*
Amn	*4.01*	*0.00132*	*Slc23a3*	*1.44*	*0.00506*
Slc13a4	*3.73*	*0.00196*	*1500032D16Rik*	*1.44*	*0.03736*
Apom	*3.62*	*0.00133*	*0910001L09Rik*	*1.43*	*0.03116*
Spp2	*3.13*	*0.00183*	*Vamp8*	*1.43*	*0.01353*
Rhox5	*2.23*	*0.00151*	*Rabac1*	*1.43*	*0.00533*
Soat2	*2.21*	*0.00463*	*Atp5l*	*1.42*	*0.04905*
Dab2	*2.19*	*0.00618*	*Apoc1*	*1.42*	*0.00712*
Clic6	*2.17*	*0.00089*	*Ppp1r3c*	*1.42*	*0.00246*
Dab2	*2.14*	*0.00646*	*Rnase4*	*1.42*	*0.00348*
Ctsh	*2.08*	*0.00690*	*Neu1*	*1.41*	*0.00792*
Pdzk1	*2.07*	*0.01013*	*Plekha2*	*1.41*	*0.02702*
Apoe	*1.95*	*0.00174*	*Them2*	*1.40*	*0.01192*
Apoa4	*1.91*	*0.01210*	*Pigp*	*1.40*	*0.02337*
H2-Bl	*1.88*	*0.00481*	*Stx3*	*1.40*	*0.01634*
Lgals2	*1.88*	*0.00204*	*Tmem60*	*1.40*	*0.01370*
Apoc1	*1.86*	*0.00173*	*Gipc2*	*1.39*	*0.00908*
Ier3	*1.79*	*0.00004*	*Fga*	*1.39*	*0.00419*
Myo6	*1.73*	*0.00245*	*Fam162a*	*1.39*	*0.00683*
S100a1	*1.68*	*0.00118*	*Glo1*	*1.39*	*0.02501*
Trf	*1.65*	*0.01341*	*Ifi30*	*1.39*	*0.00120*
Ctsc	*1.64*	*0.02157*	*Fkbp11*	*1.39*	*0.04474*
Ddit4	*1.63*	*0.00476*	*Fga*	*1.38*	*0.00915*
Glrx	*1.63*	*0.00132*	*Bri3*	*1.38*	*0.04790*
Slc2a2	*1.61*	*0.00791*	*Lum*	*1.38*	*0.01931*
Cldn2	*1.60*	*0.00254*	*5133400G04Rik*	*1.38*	*0.04748*
Xist	*1.59*	*0.00441*	*Naca*	*1.38*	*0.03942*
Pla2g12b	*1.58*	*0.00477*	*Eif4ebp1*	*1.38*	*0.01478*
Dpp4	*1.57*	*0.00948*	*B4galnt2*	*1.38*	*0.00661*
Lgmn	*1.56*	*0.01811*	*Erdr1*	*1.37*	*0.02980*
Vkorc1	*1.56*	*0.00096*	*Plac1*	*1.37*	*0.01716*
Ang	*1.55*	*0.00105*	*1110001J03Rik*	*1.37*	*0.02583*
Fmr1nb	*1.54*	*0.00384*	*0610007C21Rik*	*1.37*	*0.00682*
Ifitm3	*1.54*	*0.00563*	*Tctex1d2*	*1.37*	*0.04822*
Rps3	*1.54*	*0.04906*	*Ccdc58*	*1.36*	*0.01162*
Apoc3	*1.54*	*0.00679*	*H2-T23*	*1.36*	*0.00564*
Lgals1	*1.53*	*0.01067*	*Tbl2*	*1.36*	*0.00543*
Xist	*1.53*	*0.00788*	*Sh3bgrl3*	*1.36*	*0.03578*
Cstb	*1.53*	*0.04299*	*2610029G23Rik*	*1.36*	*0.00526*
Egr1	*1.53*	*0.00000*	*Ndufc2*	*1.36*	*0.04602*
Car4	*1.52*	*0.01439*	*Slc7a9*	*1.36*	*0.01081*
Vkorc1	*1.52*	*0.00407*	*LOC675228*	*1.36*	*0.04985*
Rbm47	*1.51*	*0.00314*	*Trappc2l*	*1.36*	*0.04878*
Serpinf2	*1.51*	*0.00402*	*Ctsc*	*1.36*	*0.02459*
Ctsb	*1.49*	*0.00860*	*Lxn*	*1.35*	*0.00116*
Vkorc1	*1.49*	*0.00515*	*Gm129*	*1.35*	*0.04297*
Atp6v0e	*1.49*	*0.00618*	*Atp5o*	*1.35*	*0.01665*
Fxyd2	*1.48*	*0.00905*	*Ccdc53*	*1.35*	*0.04359*
Fgg	*1.48*	*0.01181*	*B2m*	*1.35*	*0.02562*
0610007C21Rik	*1.48*	*0.00560*	*Trappc1*	*1.34*	*0.04180*
Bex2	*1.47*	*0.00464*	*Vars2*	*1.34*	*0.02418*
4930583H14Rik	*1.46*	*0.00256*	*Ebpl*	*1.33*	*0.03896*
Reep6	*1.46*	*0.01782*	*Gpx4*	*1.33*	*0.02690*

	Gene symbol	Fold expr.	*p-value*	Gene symbol	Fold expr.	*p-value*
	Atpif1	*1.33*	*0.03831*	*Nptx2*	*0.63*	*0.00014*
	Phlda2	*1.33*	*0.00991*	*Calr*	*0.63*	*0.00001*
	BC028528	*1.33*	*0.03079*	*Ank1*	*0.63*	*0.00945*
	Dgat2	*1.33*	*0.00385*	*Smox*	*0.64*	*0.00035*
	Ap2s1	*1.33*	*0.04136*	*Gypc*	*0.64*	*0.00103*
	Ass1	*1.33*	*0.02696*	*Bcl9l*	*0.64*	*0.00013*
	Arhgef3	*1.32*	*0.00643*	*Nkd2*	*0.64*	*0.01956*
	Ndufs7	*1.32*	*0.04400*	*Mbd1*	*0.64*	*0.00001*
E9.0	*Downregulated genes*			*Plagl2*	*0.64*	*0.00001*
	Hba-x	*0.07*	*0.00106*	*Asna1*	*0.64*	*0.00003*
	Hba-a1	*0.07*	*0.00203*	*Olfml3*	*0.64*	*0.00001*
	Hbb-bh1	*0.09*	*0.00274*	*Endod1*	*0.65*	*0.00020*
	Hbb-y	*0.10*	*0.00306*	*Ssr3*	*0.65*	*0.00222*
	Hbb-y	*0.10*	*0.00325*	*Tspan33*	*0.65*	*0.00906*
	Slc4a1	*0.12*	*0.00099*	*Nedd9*	*0.65*	*0.00000*
	Eraf	*0.16*	*0.00261*	*Eif2s3x*	*0.65*	*0.01468*
	Alas2	*0.25*	*0.00118*	*Hoxc10*	*0.65*	*0.01298*
	Blvrb	*0.28*	*0.00169*	*Klf1*	*0.65*	*0.01250*
	Cited4	*0.33*	*0.00080*	*Mcm6*	*0.65*	*0.00190*
	1190007F08Rik	*0.42*	*0.00661*	*Actb*	*0.65*	*0.00001*
	Mgst3	*0.43*	*0.00798*	*Mest*	*0.66*	*0.00013*
	Mocs1	*0.44*	*0.00000*	*H19*	*0.66*	*0.00396*
	Emcn	*0.45*	*0.00000*	*LOC100046690*	*0.66*	*0.00709*
	Tpm1	*0.45*	*0.00002*	*Gorasp2*	*0.66*	*0.00072*
	Wdr1	*0.46*	*0.00000*	*Tmsb4x*	*0.66*	*0.01005*
	Gypa	*0.47*	*0.00294*	*Uros*	*0.66*	*0.01014*
	Car2	*0.47*	*0.00054*	*Tgm2*	*0.67*	*0.02567*
	Tmem14c	*0.48*	*0.00113*	*LOC100048313*	*0.67*	*0.00185*
	Trim10	*0.50*	*0.00479*	*Dlst*	*0.67*	*0.00010*
	Slc25a37	*0.50*	*0.00519*	*Sfrs1*	*0.67*	*0.00124*
	Actb	*0.51*	*0.00003*	*Smox*	*0.67*	*0.00064*
	Mgst3	*0.53*	*0.01005*	*Pvrl2*	*0.67*	*0.00070*
	Eif4ebp2	*0.53*	*0.00096*	*Cbfb*	*0.67*	*0.00034*
	Nfe2	*0.54*	*0.00604*	*Hsp90b1*	*0.68*	*0.00295*
	Kel	*0.54*	*0.00526*	*LOC100040573*	*0.68*	*0.00010*
	1200009O22Rik	*0.55*	*0.00173*	*Kif23*	*0.68*	*0.00006*
	Wdr1	*0.55*	*0.00000*	*Hspd1*	*0.68*	*0.00011*
	Psme3	*0.56*	*0.00023*	*Alad*	*0.68*	*0.03493*
	Tpm4	*0.57*	*0.00004*	*Acss1*	*0.68*	*0.00947*
	Furin	*0.57*	*0.00001*	*4930403O06Rik*	*0.68*	*0.00064*
	Pnpo	*0.59*	*0.00478*	*Rrp12*	*0.68*	*0.00038*
	Mbnl1	*0.59*	*0.00559*	*Cmpk*	*0.68*	*0.00018*
	Wnt5a	*0.59*	*0.01433*	*Rnf10*	*0.68*	*0.00010*
	Dlk1	*0.59*	*0.01171*	*Tmsb10*	*0.68*	*0.00018*
	Tfrc	*0.60*	*0.00170*	*Ssbp3*	*0.68*	*0.00009*
	Tmem14c	*0.60*	*0.00279*	*Ube1l*	*0.68*	*0.03251*
	2810003C17Rik	*0.61*	*0.00000*	*Tmem98*	*0.68*	*0.00010*
	Srprb	*0.61*	*0.00172*	*Hnrpl*	*0.69*	*0.00202*
	Gmpr	*0.61*	*0.00300*	*Reln*	*0.69*	*0.00968*
	Fen1	*0.61*	*0.00000*	*Ywhah*	*0.69*	*0.00005*
	Glrx5	*0.61*	*0.00609*	*1110008P14Rik*	*0.69*	*0.00024*
	Slc39a8	*0.62*	*0.00259*	*Atg5*	*0.69*	*0.00011*
	Nol6	*0.62*	*0.00001*	*Psmd7*	*0.69*	*0.00010*
	Ccnd2	*0.63*	*0.00253*	*Srm*	*0.69*	*0.00245*
	Atic	*0.63*	*0.00002*	*D3Ucla1*	*0.69*	*0.00006*
	Tfrc	*0.63*	*0.00614*	*Atp2a2*	*0.69*	*0.00015*

Gene symbol	Fold expr.	*p-value*	Gene symbol	Fold expr.	*p-value*
Hspa8	*0.69*	*0.00023*	*Rrm1*	*0.73*	*0.00016*
Tomm6	*0.69*	*0.00003*	*Nol5*	*0.73*	*0.00297*
Ddx51	*0.70*	*0.00003*	*Dstn*	*0.73*	*0.00206*
Gtf3c2	*0.70*	*0.00177*	*Trim10*	*0.73*	*0.00525*
Crk	*0.70*	*0.00083*	*Rbm12*	*0.73*	*0.00141*
Immt	*0.70*	*0.00042*	*Phlda1*	*0.73*	*0.04721*
1110007M04Rik	*0.70*	*0.00006*	*Rad17*	*0.73*	*0.00006*
Prkaca	*0.70*	*0.00058*	*Vcl*	*0.73*	*0.00012*
Cdr2	*0.70*	*0.00002*	*Serpinh1*	*0.73*	*0.00087*
Tmem14c	*0.70*	*0.02525*	*Ssx2ip*	*0.73*	*0.00292*
Mfap4	*0.70*	*0.00075*	*Mrps23*	*0.73*	*0.00002*
Reep3	*0.70*	*0.00006*	*Eif3a*	*0.73*	*0.00127*
Hoxd13	*0.70*	*0.00039*	*Slc30a5*	*0.73*	*0.00237*
Ccnd3	*0.70*	*0.00058*	*Sepw1*	*0.73*	*0.01143*
Tpm3	*0.70*	*0.00002*	*Scamp4*	*0.73*	*0.00006*
Vegfc	*0.71*	*0.00001*	*Hebp1*	*0.74*	*0.04629*
Zfp131	*0.71*	*0.00006*	*A830059I20Rik*	*0.74*	*0.03447*
Mllt3	*0.71*	*0.00742*	*D16H22S680E*	*0.74*	*0.00598*
BC022224	*0.71*	*0.00358*	*LOC100048721*	*0.74*	*0.01212*
Anxa3	*0.71*	*0.00061*	*Fblim1*	*0.74*	*0.00047*
AU022870	*0.71*	*0.00188*	*Tmsb10*	*0.74*	*0.00135*
Arrdc4	*0.71*	*0.00002*	*Pfn1*	*0.74*	*0.00139*
Clptm1l	*0.71*	*0.00000*	*LOC100048413*	*0.74*	*0.00037*
Zmiz1	*0.71*	*0.00102*	*Rad23b*	*0.74*	*0.00034*
2610028A01Rik	*0.71*	*0.00015*	*Clptm1*	*0.74*	*0.00003*
Marcksl1	*0.71*	*0.00006*	*Olfm1*	*0.74*	*0.00012*
Gclm	*0.71*	*0.00312*	*LOC100048046*	*0.74*	*0.00031*
6720458F09Rik	*0.71*	*0.00329*	*Caprin1*	*0.74*	*0.00033*
Pigq	*0.71*	*0.00002*	*Anxa3*	*0.74*	*0.00069*
Nudt4	*0.71*	*0.00038*	*Caprin1*	*0.74*	*0.00322*
Fscn1	*0.72*	*0.00222*	*Hoxc10*	*0.74*	*0.01516*
Eef1d	*0.72*	*0.00282*	*Twist1*	*0.74*	*0.00532*
LOC100047261	*0.72*	*0.00089*	*Fth1*	*0.74*	*0.00393*
Myl1	*0.72*	*0.00068*	*Sfrs6*	*0.74*	*0.00030*
Srf	*0.72*	*0.00003*	*Slc30a5*	*0.74*	*0.00561*
Uros	*0.72*	*0.00459*	*Cnot4*	*0.74*	*0.00029*
Smarce1	*0.72*	*0.00032*	*Atic*	*0.74*	*0.00036*
Wdr6	*0.72*	*0.00126*	*Crcp*	*0.74*	*0.00074*
Hnrpll	*0.72*	*0.00205*	*Narg1*	*0.74*	*0.00033*
Rbm38	*0.72*	*0.00045*	*Cited2*	*0.74*	*0.00002*
Stip1	*0.72*	*0.00009*	*Noc3l*	*0.74*	*0.00002*
Ccdc86	*0.72*	*0.00032*	*Slc35a4*	*0.74*	*0.00036*
Tmem98	*0.72*	*0.00059*	*Tpcn1*	*0.75*	*0.01136*
Hprt1	*0.72*	*0.01402*	*Serbp1*	*0.75*	*0.00573*
Mrap	*0.72*	*0.01158*	*2810008M24Rik*	*0.75*	*0.00164*
Rbm8a	*0.72*	*0.00614*	*Abcf2*	*0.75*	*0.00000*
Josd2	*0.72*	*0.00914*	*Pop5*	*0.75*	*0.01865*
Sec13	*0.72*	*0.00000*	*Dph3*	*0.75*	*0.00304*
Hsd11b2	*0.72*	*0.00001*	*Eif2ak1*	*0.75*	*0.00020*
Tgs1	*0.72*	*0.00111*	*H2afy*	*0.75*	*0.02048*
Srm	*0.73*	*0.00235*	*Klhl9*	*0.75*	*0.00010*
Nedd9	*0.73*	*0.00000*	*Nup85*	*0.75*	*0.00000*
Tprgl	*0.73*	*0.00000*	*AU021838*	*0.75*	*0.00003*
Gpx1	*0.73*	*0.00247*	*Hnrnph1*	*0.75*	*0.00635*
Mllt3	*0.73*	*0.01077*	*Tal1*	*0.75*	*0.00652*
Ide	*0.73*	*0.00281*	*Rhox5*	*0.75*	*0.00078*

	Gene symbol	Fold expr.	*p-value*	Gene symbol	Fold expr.	*p-value*
	Ift20	*0.75*	*0.00036*	*1500032D16Rik*	*1.66*	*0.00011*
	Mlp	*0.75*	*0.00005*	*Acat2*	*1.65*	*0.00066*
	Slc12a2	*0.75*	*0.00017*	*Cyp51*	*1.65*	*0.00106*
	Cpsf6	*0.75*	*0.00004*	*Grb7*	*1.64*	*0.00410*
	Ecm1	*0.75*	*0.00001*	*Dbi*	*1.61*	*0.00020*
	Hemgn	*0.75*	*0.00397*	*Fam162a*	*1.60*	*0.00005*
	Cd40	*0.75*	*0.02642*	*Mrpl2*	*1.60*	*0.00080*
	Ahdc1	*0.75*	*0.00031*	*Haghl*	*1.60*	*0.00003*
	Ppa1	*0.75*	*0.00042*	*Ptprs*	*1.59*	*0.00004*
	LOC100044322	*0.75*	*0.00015*	*Bhlhb2*	*1.59*	*0.00007*
	Pno1	*0.75*	*0.00074*	*Upp1*	*1.58*	*0.01284*
	Mat2a	*0.75*	*0.00152*	*Eif4ebp1*	*1.58*	*0.00054*
	Prpf40a	*0.75*	*0.00187*	*Hoxa7*	*1.57*	*0.00001*
	Stk4	*0.75*	*0.00157*	*LOC100047934*	*1.57*	*0.00000*
	Flrt3	*0.75*	*0.01432*	*H2-Ab1*	*1.57*	*0.00001*
	Cdv3	*0.75*	*0.00191*	*Rbp1*	*1.56*	*0.00152*
	Mrps7	*0.76*	*0.00004*	*Nrn1*	*1.56*	*0.00006*
	Creld2	*0.76*	*0.00010*	*LOC100045403*	*1.56*	*0.00507*
	Mapk14	*0.76*	*0.00166*	*Pcyt2*	*1.55*	*0.00051*
	Ubxd8	*0.76*	*0.00152*	*Cdt1*	*1.55*	*0.00027*
	Pon2	*0.76*	*0.00008*	*Sqle*	*1.55*	*0.00123*
E9.0	*Upregulated genes*			*Csnk1d*	*1.55*	*0.00009*
	Ddit4	*3.56*	*0.00000*	*Pold1*	*1.55*	*0.00109*
	Scd1	*2.12*	*0.00001*	*5133400G04Rik*	*1.55*	*0.00001*
	Stard4	*2.08*	*0.00001*	*Fhl1*	*1.55*	*0.00003*
	Aldoc	*2.04*	*0.00013*	*Hoxd4*	*1.54*	*0.00034*
	Plekha2	*2.03*	*0.00003*	*Rdm1*	*1.54*	*0.00002*
	Anxa2	*1.97*	*0.00066*	*Vegfa*	*1.54*	*0.00001*
	Aacs	*1.94*	*0.00000*	*Suv420h2*	*1.54*	*0.00024*
	Mvd	*1.94*	*0.00034*	*Asah1*	*1.54*	*0.00058*
	Ier3	*1.91*	*0.00012*	*Maf1*	*1.53*	*0.00004*
	Chac1	*1.90*	*0.02319*	*Vldlr*	*1.53*	*0.00002*
	LOC100040592	*1.89*	*0.00004*	*Insig1*	*1.53*	*0.00021*
	Stc2	*1.88*	*0.00000*	*5133400G04Rik*	*1.53*	*0.00001*
	Dbp	*1.86*	*0.00000*	*Tpi1*	*1.53*	*0.00020*
	Srebf2	*1.84*	*0.00000*	*Akap8l*	*1.53*	*0.00023*
	Actb	*1.83*	*0.00137*	*Csnk1d*	*1.52*	*0.00003*
	Pgm2	*1.83*	*0.00012*	*6430706D22Rik*	*1.52*	*0.00002*
	Fdps	*1.82*	*0.00080*	*LOC100048445*	*1.52*	*0.00004*
	Vldlr	*1.81*	*0.00000*	*Gm129*	*1.52*	*0.00007*
	Galk1	*1.80*	*0.00001*	*Pdk1*	*1.52*	*0.00018*
	Pfkp	*1.78*	*0.00000*	*0610007C21Rik*	*1.51*	*0.00002*
	Rps3	*1.77*	*0.00006*	*Triobp*	*1.51*	*0.00005*
	Actb	*1.77*	*0.00093*	*Them2*	*1.51*	*0.00009*
	4930583H14Rik	*1.74*	*0.00001*	*Chst1*	*1.51*	*0.00006*
	Renbp	*1.73*	*0.00000*	*Fam125a*	*1.51*	*0.00028*
	Prnp	*1.72*	*0.00008*	*Dym*	*1.51*	*0.00138*
	1500032D16Rik	*1.71*	*0.00031*	*Klk8*	*1.50*	*0.00000*
	1810027O10Rik	*1.70*	*0.00020*	*Grina*	*1.50*	*0.00277*
	Tpi1	*1.69*	*0.00002*	*Eef2*	*1.50*	*0.00122*
	5133400G04Rik	*1.69*	*0.00001*	*Irx3*	*1.50*	*0.01827*
	Sc4mol	*1.68*	*0.00475*	*Neurl*	*1.49*	*0.00002*
	Nipsnap1	*1.68*	*0.00001*	*Ccdc58*	*1.49*	*0.00001*
	2310022B05Rik	*1.67*	*0.00001*	*Nr6a1*	*1.49*	*0.02161*
	Aldoa	*1.67*	*0.00010*	*Ift172*	*1.48*	*0.00000*
	Ppp1r3c	*1.66*	*0.00000*	*Pla2g12a*	*1.48*	*0.00031*

Gene symbol	Fold expr.	*p-value*	Gene symbol	Fold expr.	*p-value*
Espn	*1.47*	*0.00005*	*Rnf19a*	*1.41*	*0.00002*
Vldlr	*1.47*	*0.00016*	*Bri3*	*1.40*	*0.00007*
Gdi1	*1.47*	*0.00033*	*Ppp2r5c*	*1.40*	*0.00001*
Rras	*1.47*	*0.00013*	*Nfatc4*	*1.40*	*0.00398*
Rps5	*1.46*	*0.00001*	*Scamp1*	*1.40*	*0.00034*
LOC100043671	*1.46*	*0.00000*	*Hoxc6*	*1.40*	*0.00123*
Zfyve21	*1.46*	*0.00003*	*Jmjd1a*	*1.39*	*0.00006*
Sqle	*1.46*	*0.00038*	*Crabp1*	*1.39*	*0.00001*
Egr1	*1.46*	*0.00000*	*Osgep*	*1.39*	*0.00001*
Grhpr	*1.46*	*0.00004*	*Mrpl17*	*1.39*	*0.00315*
Dag1	*1.46*	*0.00022*	*Rxrb*	*1.39*	*0.00137*
Rps6ka1	*1.45*	*0.00206*	*Uaca*	*1.39*	*0.01728*
Neu1	*1.45*	*0.00010*	*Vars2*	*1.39*	*0.00018*
Use1	*1.45*	*0.00021*	*Itgb5*	*1.39*	*0.00001*
Nsdhl	*1.45*	*0.00131*	*Slc25a11*	*1.39*	*0.00005*
Cox7a2l	*1.45*	*0.00004*	*Agrn*	*1.38*	*0.00257*
Tmem205	*1.45*	*0.00000*	*5830434P21Rik*	*1.38*	*0.00043*
LOC100044692	*1.45*	*0.00068*	*Tpd52*	*1.38*	*0.00168*
Nsdhl	*1.45*	*0.00070*	*Cope*	*1.38*	*0.00177*
H2-Ke6	*1.44*	*0.00007*	*Ikbkap*	*1.38*	*0.00000*
Sap30l	*1.44*	*0.00014*	*Phc2*	*1.38*	*0.00455*
Fam173a	*1.44*	*0.00007*	*LOC641240*	*1.38*	*0.00001*
Tubb2b	*1.44*	*0.00336*	*Dap3*	*1.38*	*0.00008*
Rpl22	*1.44*	*0.00004*	*Ndufa6*	*1.38*	*0.00008*
Pygl	*1.44*	*0.00089*	*Litaf*	*1.38*	*0.00000*
LOC100044298	*1.44*	*0.00001*	*Esm1*	*1.38*	*0.00007*
Wnt6	*1.44*	*0.00000*	*LOC100048589*	*1.38*	*0.00468*
Wtip	*1.44*	*0.00093*	*Mic2l1*	*1.37*	*0.00039*
Triobp	*1.43*	*0.00002*	*1600016N20Rik*	*1.37*	*0.00001*
Trappc6a	*1.43*	*0.00000*	*Araf*	*1.37*	*0.00001*
Eif3k	*1.43*	*0.00143*	*LOC654467*	*1.37*	*0.00218*
Lcmt1	*1.43*	*0.00010*	*Paox*	*1.37*	*0.00012*
Insig1	*1.43*	*0.00050*	*Hsd17b7*	*1.37*	*0.00436*
Dusp4	*1.43*	*0.00302*	*Hint2*	*1.37*	*0.00036*
Nup210	*1.43*	*0.00003*	*2900010M23Rik*	*1.37*	*0.00584*
Pcsk9	*1.43*	*0.01109*	*D15Ertd682e*	*1.37*	*0.01029*
Flywch2	*1.43*	*0.00023*	*Abhd14b*	*1.37*	*0.00209*
Ppme1	*1.43*	*0.00001*	*Smc1a*	*1.37*	*0.00043*
Bbc3	*1.42*	*0.00053*	*Gstz1*	*1.37*	*0.00326*
Fhl1	*1.42*	*0.00004*	*Nsdhl*	*1.37*	*0.00163*
1500010J02Rik	*1.42*	*0.00012*	*Cmc1*	*1.37*	*0.00003*
Eef2	*1.42*	*0.00374*	*Supt3h*	*1.37*	*0.00066*
Gabarapl1	*1.42*	*0.00005*	*Spag5*	*1.37*	*0.00008*
Dap3	*1.42*	*0.00002*	*Cdkn1c*	*1.36*	*0.00094*
Klhl17	*1.42*	*0.00000*	*Vps28*	*1.36*	*0.00026*
Hbp1	*1.42*	*0.00011*	*Smtnl2*	*1.36*	*0.00660*
LOC381629	*1.42*	*0.00016*	*0610007C21Rik*	*1.36*	*0.00003*
Ypel3	*1.42*	*0.00277*	*Glb1*	*1.36*	*0.00004*
Egln3	*1.42*	*0.00004*	*Rpl13a*	*1.36*	*0.00071*
D10Ertd610e	*1.41*	*0.00084*	*Polr1c*	*1.36*	*0.00009*
Mrps9	*1.41*	*0.00004*	*Tmem118*	*1.36*	*0.00004*
Bnip3l	*1.41*	*0.00002*	*LOC100045343*	*1.36*	*0.00049*
Arrdc3	*1.41*	*0.00156*	*Mpnd*	*1.36*	*0.00002*
Msx1	*1.41*	*0.04662*	*Hsd17b10*	*1.36*	*0.00001*
Lss	*1.41*	*0.00426*	*Fxyd6*	*1.36*	*0.00033*
Xpnpep1	*1.41*	*0.00002*	*Ttc3*	*1.36*	*0.00007*

Gene symbol	Fold expr.	*p-value*	Gene symbol	Fold expr.	*p-value*
Cln3	*1.36*	*0.00003*	*Idh2*	*1.34*	*0.00243*
Cercam	*1.36*	*0.00284*	*Tbl2*	*1.34*	*0.00000*
Pgls	*1.36*	*0.00138*	*Acp6*	*1.34*	*0.00032*
Tubb2b	*1.36*	*0.00846*	*Pfdn5*	*1.34*	*0.00000*
Edem2	*1.35*	*0.00523*	*Parp2*	*1.34*	*0.00064*
Psmg2	*1.35*	*0.00007*	*Aplp2*	*1.34*	*0.00039*
Nlgn2	*1.35*	*0.00000*	*Slc7a3*	*1.34*	*0.04398*
Mnat1	*1.35*	*0.00001*	*Aasdh*	*1.33*	*0.00001*
Slc25a1	*1.35*	*0.00130*	*Eef2*	*1.33*	*0.00324*
Tctex1d2	*1.35*	*0.00739*	*Kctd2*	*1.33*	*0.00136*
Rpl31	*1.35*	*0.00001*	*Arhgef4*	*1.33*	*0.00024*
Cd81	*1.35*	*0.00115*	*Klhl26*	*1.33*	*0.00005*
Mnat1	*1.35*	*0.00010*	*Tkt*	*1.33*	*0.00716*
Pcdh21	*1.35*	*0.00975*	*Zfp277*	*1.33*	*0.00216*
Nusap1	*1.35*	*0.00211*	*2310047M10Rik*	*1.32*	*0.00011*
Ankzf1	*1.35*	*0.00000*	*Rdh11*	*1.32*	*0.00076*
Csnk1e	*1.35*	*0.00540*	*Alg8*	*1.32*	*0.00177*
Pfkl	*1.34*	*0.00135*	*Ankzf1*	*1.32*	*0.00002*
Smarca2	*1.34*	*0.00016*	*2810046M22Rik*	*1.32*	*0.00007*
Prr14	*1.34*	*0.00391*	*Tubb2b*	*1.32*	*0.00036*
Copz2	*1.34*	*0.00013*	*4833426J09Rik*	*1.32*	*0.00011*
Dag1	*1.34*	*0.00178*	*2410001C21Rik*	*1.32*	*0.00086*
Usp52	*1.34*	*0.00016*			

Table 13: Deregulated genes in caudal end halves of wt embryos after 4 hour treatment with the chemical FGF receptor inhibitor SU5402; fold change ≥ $ln2^{0.4}$; p≤ 10^{-5}

Gene symbol	Fold Expr.	p-value
Downregulated		
Dusp6	*0,18*	*0,00000*
Egr1	*0,26*	*0,00514*
Tnfrsf12a	*0,29*	*0,00010*
Sgk1	*0,43*	*0,00010*
Ctgf	*0,44*	*0,00524*
Atf3	*0,45*	*0,03517*
Fzd5	*0,45*	*0,00046*
Cdkn1a	*0,46*	*0,00665*
Cdkn1a	*0,47*	*0,00613*
Slco4a1	*0,48*	*0,01395*
Lor	*0,48*	*0,00074*
Coq10b	*0,48*	*0,00005*
Fos	*0,48*	*0,00457*
Coq10b	*0,48*	*0,00028*
Dusp1	*0,49*	*0,00005*
Herpud1	*0,49*	*0,00714*
Junb	*0,50*	*0,00053*
Gch1	*0,50*	*0,00001*
Herpud1	*0,50*	*0,00737*
Klf5	*0,50*	*0,00146*
Sertad1	*0,50*	*0,00063*
Epha2	*0,52*	*0,00334*
Etv5	*0,52*	*0,00163*
Cdkn1a	*0,52*	*0,01979*
Gadd45a	*0,53*	*0,00141*
Ccng1	*0,53*	*0,02327*
Plk2	*0,53*	*0,00206*
Vasn	*0,53*	*0,00007*
Hoxb7	*0,56*	*0,00982*
Fzd7	*0,56*	*0,01630*
Errfi1	*0,56*	*0,00004*
Dusp2	*0,57*	*0,00019*
Hexim1	*0,57*	*0,03744*
Atp1b1	*0,58*	*0,00367*
Gcnt2	*0,58*	*0,00252*
Ier3	*0,58*	*0,00087*
Mfsd2	*0,58*	*0,00225*
Snora65	*0,58*	*0,00023*
Nrip3	*0,58*	*0,00000*
Axud1	*0,59*	*0,00085*
Trp53inp1	*0,59*	*0,01380*
Hes1	*0,60*	*0,00190*
Zfp36	*0,60*	*0,00010*
Enc1	*0,60*	*0,00201*
Ddit4l	*0,60*	*0,02824*
Tuft1	*0,61*	*0,00156*
Gata2	*0,61*	*0,00260*
Gadd45a	*0,61*	*0,00600*
LOC100046120	*0,61*	*0,00074*
Klf6	*0,61*	*0,01293*
Klhl21	*0,61*	*0,00157*
Klf9	*0,62*	*0,00000*
Tubb6	*0,63*	*0,00000*
Ccrn4l	*0,63*	*0,00004*
Gfpt2	*0,63*	*0,00060*
Hspa2	*0,63*	*0,01843*
Tcfap2c	*0,64*	*0,00257*
Hoxb6	*0,64*	*0,00635*
Dusp6	*0,64*	*0,00013*
Etsrp71	*0,64*	*0,03744*
Phlda3	*0,65*	*0,00414*
Zswim4	*0,65*	*0,00916*
Cdr2	*0,65*	*0,00002*
Tuba6	*0,65*	*0,00658*
Dusp4	*0,65*	*0,00484*
Dusp16	*0,65*	*0,01360*
Pdlim7	*0,65*	*0,00228*
Vangl2	*0,66*	*0,00023*
Rbm38	*0,66*	*0,00058*
Mat2a	*0,66*	*0,00226*
4732471D19Rik	*0,66*	*0,00310*
Maff	*0,66*	*0,00113*
1190002H23Rik	*0,66*	*0,02075*
Dido1	*0,66*	*0,03777*
Dusp11	*0,66*	*0,02327*
Klf4	*0,66*	*0,01019*
Id1	*0,66*	*0,00004*
Dusp8	*0,67*	*0,03114*
Tob1	*0,67*	*0,01835*
Nedd9	*0,67*	*0,00256*
Hbegf	*0,67*	*0,00799*
Nedd9	*0,67*	*0,00433*
Vegfa	*0,67*	*0,03018*
Tacstd2	*0,67*	*0,02074*
Gtf2f1	*0,67*	*0,01033*
Gpha2	*0,67*	*0,00232*
Flt1	*0,67*	*0,04322*
Ehd4	*0,67*	*0,00003*
Mcam	*0,68*	*0,00023*
Pik3r3	*0,68*	*0,00008*
Klhl6	*0,68*	*0,00782*
Ppp1r13b	*0,68*	*0,00071*
5730593F17Rik	*0,68*	*0,01401*
Phlda3	*0,68*	*0,01488*
LOC100047261	*0,68*	*0,00432*
Hap1	*0,68*	*0,00743*
Gsto2	*0,68*	*0,04195*
Mrps6	*0,68*	*0,00007*
Flrt3	*0,69*	*0,00286*
Gdf9	*0,69*	*0,04983*
Cebpb	*0,69*	*0,00419*
Frat2	*0,69*	*0,00124*
Gjb3	*0,69*	*0,00001*

Gene symbol	Fold Expr.	p-value
Sirt1	*0,69*	*0,00150*
Arc	*0,69*	*0,02416*
LOC100048721	*0,70*	*0,00390*
Sp6	*0,70*	*0,00048*
Rbm38	*0,70*	*0,00623*
Uchl1	*0,70*	*0,00113*
Csnrp2	*0,70*	*0,00000*
Klhl6	*0,70*	*0,00657*
Hspa1a	*0,70*	*0,00931*
Wsb1	*0,70*	*0,00058*
2310061F22Rik	*0,70*	*0,00142*
Tpbg	*0,71*	*0,00807*
Dusp16	*0,71*	*0,03632*
Ckb	*0,71*	*0,01711*
Rhob	*0,71*	*0,00001*
Rnd3	*0,71*	*0,00018*
Nptx2	*0,71*	*0,02521*
Nuak2	*0,71*	*0,00878*
Oat	*0,71*	*0,00100*
F2rl1	*0,71*	*0,00024*
Slc19a2	*0,71*	*0,00186*
Pou3f1	*0,71*	*0,02293*
Per2	*0,71*	*0,00000*
Nfkbiz	*0,71*	*0,00021*
Vegfa	*0,71*	*0,01282*
Creb3	*0,71*	*0,03425*
Midn	*0,72*	*0,01365*
Irs2	*0,72*	*0,00821*
Sp6	*0,72*	*0,00016*
Fjx1	*0,72*	*0,00041*
Pdgfa	*0,72*	*0,00469*
Ddx26b	*0,72*	*0,00035*
Bbc3	*0,72*	*0,00474*
Polr1e	*0,72*	*0,02877*
Traf4	*0,72*	*0,01870*
LOC100046232	*0,73*	*0,00005*
Rln1	*0,73*	*0,00789*
Alas1	*0,73*	*0,00283*
Tcfap2c	*0,73*	*0,00450*
Stat3	*0,73*	*0,00001*
LOC100045567	*0,73*	*0,00161*
Pim3	*0,73*	*0,00105*
Mknk2	*0,73*	*0,00023*
2310008H09Rik	*0,73*	*0,00098*
Arf6	*0,73*	*0,00016*
Irx2	*0,73*	*0,01540*
Hspb1	*0,73*	*0,00241*
Jmjd1a	*0,73*	*0,00006*
Ngfr	*0,73*	*0,00713*
Pkp2	*0,73*	*0,00483*
Ier5l	*0,73*	*0,01454*
Per2	*0,73*	*0,00000*
Midn	*0,73*	*0,02117*
Fosb	*0,74*	*0,00724*
Ndrl	*0,74*	*0,00300*
Irf1	*0,74*	*0,00883*

Gene symbol	Fold Expr.	p-value
Fpgs	*0,74*	*0,00518*
Slc6a8	*0,74*	*0,00067*
8430408G22Rik	*0,74*	*0,04136*
Tcf21	*0,74*	*0,03860*
Pcdh17	*0,74*	*0,00015*
Pkp2	*0,74*	*0,01242*
Dlx3	*0,74*	*0,00918*
Bag3	*0,74*	*0,00378*
Timp3	*0,74*	*0,00049*
Fbxo33	*0,74*	*0,00356*
Plec1	*0,74*	*0,00257*
Dnajb9	*0,74*	*0,00349*
Tes	*0,75*	*0,00077*
Elf1	*0,75*	*0,00449*
Zfp36l1	*0,75*	*0,00044*
Zfp36l1	*0,75*	*0,00231*
Esco2	*0,75*	*0,00001*
Tax1bp3	*0,75*	*0,00295*
Hhex	*0,75*	*0,00441*
Traf4	*0,75*	*0,02786*
Xbp1	*0,75*	*0,00005*
LOC216443	*0,75*	*0,02828*
Ankrd50	*0,75*	*0,00309*
Pnpla2	*0,75*	*0,00281*
Cklf	*0,75*	*0,01620*
Elk3	*0,75*	*0,00685*
Hspa2	*0,75*	*0,02591*
Gtse1	*0,75*	*0,01610*
Neu1	*0,75*	*0,00137*
Mcl1	*0,75*	*0,00645*
Pank2	*0,76*	*0,00001*
2310003H01Rik	*0,76*	*0,00384*
Pop5	*0,76*	*0,03794*
Upregulated genes		
Mid1	*2,68*	*0,00010*
Laptm5	*2,40*	*0,00262*
Adra2a	*1,98*	*0,00008*
Pdgfra	*1,79*	*0,00594*
Avil	*1,78*	*0,00014*
Ampd1	*1,73*	*0,00172*
Myo6	*1,70*	*0,00004*
Zfp292	*1,65*	*0,04656*
Ampd1	*1,63*	*0,00018*
Celsr3	*1,63*	*0,00034*
Tnrc6a	*1,60*	*0,00701*
Smoc1	*1,60*	*0,02246*
Cdk5r1	*1,60*	*0,04120*
Zfp326	*1,60*	*0,00370*
Psg23	*1,59*	*0,00150*
Creb1	*1,59*	*0,00000*
Rasl11b	*1,58*	*0,01175*
Tnrc6a	*1,58*	*0,00926*
Rasl11b	*1,58*	*0,01017*
Sdf4	*1,58*	*0,00014*
B230342M21Rik	*1,56*	*0,00070*
Rps15a	*1,55*	*0,00036*

Gene symbol	Fold Expr.	p-value
Lrrc40	*1,55*	*0,00007*
Hvcn1	*1,54*	*0,01223*
Tecta	*1,54*	*0,00017*
H13	*1,52*	*0,00047*
Ankfy1	*1,51*	*0,00135*
Traf3	*1,51*	*0,00016*
Pknox2	*1,50*	*0,00036*
Traf2	*1,50*	*0,00009*
Fzd2	*1,49*	*0,00146*
Zfp251	*1,49*	*0,01698*
4930403C10Rik	*1,49*	*0,00308*
Cldn15	*1,48*	*0,00675*
Traf3	*1,48*	*0,00055*
Gemin4	*1,48*	*0,00089*
F730047E07Rik	*1,47*	*0,00285*
Zfp326	*1,47*	*0,00643*
Zfp41	*1,47*	*0,03253*
Fut10	*1,47*	*0,00040*
D130043K22Rik	*1,47*	*0,00345*
Gad1	*1,45*	*0,03738*
Zfp326	*1,45*	*0,00012*
Plch2	*1,45*	*0,00023*
Hes5	*1,44*	*0,02796*
4930519N16Rik	*1,44*	*0,00004*
Sephs1	*1,44*	*0,00106*
2510009E07Rik	*1,44*	*0,02446*
Dlg3	*1,43*	*0,01123*
C87436	*1,43*	*0,00213*
4632411B12Rik	*1,43*	*0,02697*
Setd1b	*1,43*	*0,04254*
Clspn	*1,43*	*0,01439*
Xpnpep3	*1,42*	*0,00007*
Zer1	*1,42*	*0,00916*
Dusp11	*1,42*	*0,00482*
Bahcc1	*1,41*	*0,01204*
Zfp292	*1,41*	*0,03234*
Gemin4	*1,41*	*0,00020*
Hsd3b7	*1,41*	*0,04314*
Rab22a	*1,41*	*0,00086*
Ubr1	*1,41*	*0,00907*
LOC100044324	*1,41*	*0,03019*
Mkrn3	*1,41*	*0,00936*
Zfp41	*1,41*	*0,03904*
Cachd1	*1,41*	*0,00202*
Ank1	*1,40*	*0,04428*

Gene symbol	Fold Expr.	p-value
Clspn	*1,40*	*0,01639*
Zfhx3	*1,40*	*0,04241*
2810410P22Rik	*1,39*	*0,04393*
Cbx5	*1,39*	*0,00295*
Rnmt	*1,39*	*0,00039*
Glis1	*1,39*	*0,00008*
Rab22a	*1,38*	*0,00410*
Crbn	*1,38*	*0,00082*
Fbxo9	*1,38*	*0,02798*
LOC100047173	*1,38*	*0,00040*
Dok3	*1,38*	*0,01152*
2610207I05Rik	*1,38*	*0,01791*
Sdc3	*1,38*	*0,00468*
Lman2	*1,38*	*0,00301*
Mfap3	*1,37*	*0,00090*
Rhof	*1,37*	*0,00014*
Usp1	*1,37*	*0,00051*
F730047E07Rik	*1,36*	*0,00091*
Prkacb	*1,36*	*0,03411*
2210417D09Rik	*1,36*	*0,00339*
Clspn	*1,36*	*0,00601*
Amigo3	*1,36*	*0,00009*
Chd5	*1,36*	*0,00175*
Ctns	*1,36*	*0,01571*
Mtmr12	*1,35*	*0,00008*
1110018G07Rik	*1,35*	*0,00030*
Ddx6	*1,35*	*0,00296*
Sfxn1	*1,35*	*0,00676*
LOC100046690	*1,34*	*0,04083*
Ncoa6	*1,34*	*0,01356*
Yeats2	*1,34*	*0,00050*
Papd1	*1,34*	*0,00053*
A930010I20Rik	*1,34*	*0,03920*
Ppp2r5e	*1,34*	*0,00118*
D0H4S114	*1,34*	*0,03572*
1700029G01Rik	*1,34*	*0,00068*
Igsf9	*1,34*	*0,02407*
Mysm1	*1,34*	*0,04675*
Fign	*1,33*	*0,01066*
Cxxc4	*1,33*	*0,00045*
Rbm4	*1,33*	*0,00022*
Mfap3	*1,33*	*0,00082*
Snx14	*1,33*	*0,00112*
1700102P08Rik	*1,33*	*0,00036*
Slc17a8	*1,33*	*0,02437*

Gene symbol	Fold Expr.	p-value
Tshz3	*1,33*	*0,02987*
Rbak	*1,33*	*0,02575*
Setd5	*1,33*	*0,02691*
Sox12	*1,33*	*0,00215*
AW549877	*1,33*	*0,01198*
Zmiz1	*1,33*	*0,02466*
Zbtb24	*1,33*	*0,01538*
LOC381302	*1,32*	*0,00081*

2310047C04Rik	*1,32*	*0,00683*
Usp1	*1,32*	*0,00256*
Stx8	*1,32*	*0,02046*
Rsrc2	*1,32*	*0,02698*
Trpc2	1,32	0,02133