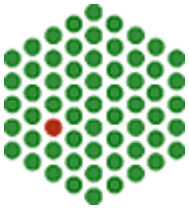
Genetic Analysis of the *Sall* Transcription Factor Family in Murine Development

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.) DER NATURWISSENSCHAFTLICHEN FAKULTÄT III – BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG



vorgelegt von

Ulrich Elling aus Freising

2005

Promotionsgesuch eingereicht am:	21. November 2005
Die Arbeit wurde angeleitet von:	Dr. M. Treier
Doktorvater:	Prof. Dr. Schneuwly
Thesis advisory committee, EMBL:	Dr. S. Cohen
	Dr. C. Nerlov

Prüfungsausschuß:

Vorsitzender:	Prof. Dr. Kunzelmann
1. Gutachter	Prof. Dr. Schneuwly
2. Gutachter	Prof. Dr. Witzgall
3. Prüfer:	Prof. Dr. Seufert
Ersatzperson:	Prof. Dr. Warth

Zusammenfassung

Spalt Proteine bilden eine Familie von Transkriptionsfaktoren, die von *Caenorhabdits elegans* über *Drosophila melanogaster* bis hin zu Vertebraten konserviert sind. In Säugetiergenomen findet man jeweils vier Gene, die als *Spalt-like* (*SALL1 bis SALL4*) bezeichnet werden. Mutationen in *SALL1* und *SALL4* verursachen im Menschen Townes-Brocks-Syndrom (TBS) beziehungsweise Okihiro-/Duane Radial Ray-Syndrom (DRRS). Beide Syndrome variieren stark in ihrem klinischen Befund, beschrieben sind unter anderem Verdopplung oder Fehlen von Fingern und Zehen, Schäden im zentralen Nervensystem sowie Missbildungen im Urogenitalsystem. Wie es zu diesen multiplen Fehlbildungen kommt ist nicht verstanden. Aufbauend auf eine systematische, vergleichende Untersuchung der Expressionsmuster wurden in der vorliegenden Arbeit die Auswirkungen von Mutationen in *Sall1* bis *Sall4* auf die Embryonalentwicklung von Mäusen einzeln oder in Kombination untersucht. Bisher wurde nur gezeigt, dass Verlust von *Sall1* zur Nierenagenese führt. Das Ziel war deshalb, durch Untersuchung der verschiedenen *Sall* Mutationen in unterschiedlichen Organsystemen zugrunde liegende Gemeinsamkeiten herauszuarbeiten, die eine generelle Rolle von *Sall* Genen während der Embryonalentwicklung aufzeigen.

Sall4 ist der erste *Sall* Transkriptionsfaktor, der während der Embryonalentwicklung exprimiert wird. Die innere Zellmasse (IZM) der Blastula, aus der die embryonalen Stammzellen und das Extraembryonale Endoderm hervorgehen, exprimiert *Sall4* besonders stark. Im Trophoblast hingegen wird die Transkription von *Sall4* ab dem Tag e5.5 vollständig reprimiert. Die Deletion von *Sall4* aus dem Mausgenom führt bei homozygoten Tieren zum Fehlen der beiden IZM Zellinien und damit zum Absterben des Embryos direkt nach der Einnistung in die Gebärmutter. Das Studium dieses Prozesses in einem Kultursystem bestätigte den sukzessiven Verlust der IZM und führte zu der Aussage, dass Sall4 für den Erhalt der embryonalen Stammzellen und des primitiven Endoderms notwendig ist. In Abwesenheit von Sall4 kommt es jedoch nicht zum Zelltod sondern zu reduzierter Zellteilung dieser Zellinien und zu deren Differenzierung. Auch embryonale Stammzellen verlieren in Kultur die Fähigkeit zur Zellteilung, wenn *Sall4* deletiert wird.

Um zu testen, ob *Sall4* die Identität von Zellen zellautonom steuert, wurden chimäre Mäuse aus Zellen mit und ohne *Sall4* hergestellt. Dabei wurde festgestellt, dass Zellen, deren *Sall4* Gen entfernt ist, auch im Verband mit Wildtyp-Zellen nicht zum Embryo beitragen können. Dies lässt die Schlussfolgerung zu, dass die Funktion von Sall4 zellautonom ist.

Im Gegensatz zum frühen Embryo werden in Vorläuferzellen des zentralen Nervensystems alle vier *Sall* Gene abgelesen. Es konnte gezeigt werden, dass die Mutation einzelner Gene nur milde Effekte auf die Entwicklung hat wohingegen das Entfernen mehrerer Sall Faktoren zu Verlust, Verkleinerung oder mangelnder Organisation verschiedener Bereiche des Nervensystems führt. Dies wurde auf die Beobachtung zurückgeführt, dass die Zellteilungsrate verringert ist. Beispiele hierfür sind die Grosshirnrinde und der Bulbus olfactorius. In *Sall1/Sall2* Doppelmutanten kommt es so zum kompletten Verlust des Geruchssinns. Auch die wachsende Extremitätenknospe aktiviert mehrere *Sall* Gene, in späteren Stadien vor allem *Sall1* und *Sall3*. Hier konnte gezeigt werden, dass die vollständige Ausbildung aller Elemente des Skeletts von der Dosis der Sall Aktivität abhängt und totale Abwesenheit von *Sall1* und *Sall3* Oligodaktylie zur Folge hat.

Mutation mehrerer Gene führt in Mäusen demnach zu Entwicklungsschäden vergleichbar denen, wie sie in Menschen, für TBS und DRRS beschrieben sind. Funktionen von Sall Faktoren überlappen stark. In allen untersuchten Entwicklungsprozessen kommt es in Abwesenheit aller hier exprimierten Sall Transkriptionsfaktoren zum Fehlen von differenzierten Zellen in den betroffenen Geweben, das sich nicht durch Zelltod sondern durch den Verlust der Vorläuferpopulation erklärt. Da Sall Proteine immer in Vorläufer- und Stammzellen aktiv sind lässt sich also ein allgemeines Modell postulieren, nachdem Sall Proteine für den Erhalt dieser nicht differenzierten Zellen nötig sind. Dieser Blickwinkel eröffnet neue Interpretationsmöglichkeiten für das Zustandekommen der Vielzahl an Missbildungen in TBS und DRRS.

Table of Contents

Introduction

In development, a single, totipotent cell undergoes multiple rounds of division accompanied by a determination and differentiation process. Differentiation occurs at the expense of a loss of developmental potential in almost all daughter cells of a progenitor cell. While several cell types remain competent to differentiate into multiple others as well as to self-renew, the only truly totipotent cells constitute the germline of an adult animal. Loss of totipotency is tightly linked to differentiation and controlled by multiple cellular signalling pathways. As a consequence to the signals received, much of the developmental state of a cell is defined by the genes expressed, facilitating the cell to carry out the designated functions. To translate the proper set of proteins, cells evolved a complex circuitry of transcriptional regulators. Control of transcription will reflect on the level of messenger RNA and ultimately contribute to the regulation of protein abundance. Many transcription factors themselves are thus expressed in a tightly controlled fashion in space and time. One such transcriptional regulator family that plays a crucial role in multiple developmental processes is founded by the *Drosophila* gene *spalt*. The mutation was initially described in Drosophila melanogaster (Tsonis and Goetinck, 1988). Subsequently, cloning revealed that the observed phenotypes are caused by mutations of a Zn-finger transcription factor (Kuhnlein et al., 1994).

1. The Spalt-like Transcription Factor Family

1.1. Orthologs and Similarities

Spalt (*sal*) like proteins are transcription factors that are described for nematodes, shrimp, insects and multiple chordata analyzed till date. 175 members of the Spalt family of proteins have been annotated in different sequenced organisms (ensembl). The simplest organism described to contain a Spalt homolog is *C. elegans*. Only one *spalt* like gene is found in *C. elegans*, the genome of *Drosophila melanogaster* harbours two paralogs and in vertebrates typically four genes are present. These paralogs seem to have arisen by several duplication events of the *spalt* locus and are called *Sall* for *spalt-like*. Figure 1 shows the relative homology of a selection of 23 *spalt-like* genes based on protein conservation. Judged by sequence homology, the vertebrate *spalt* genes evolved from one ancestor more closely related to the Drosophila *spalt major* (*salm*) ortholog than to *spalt-related* (*salr*). *Sem-4* of *C. elegans* and *spalt-related* form a distant subgroup (figure 1) with the two *Drosophila spalt*

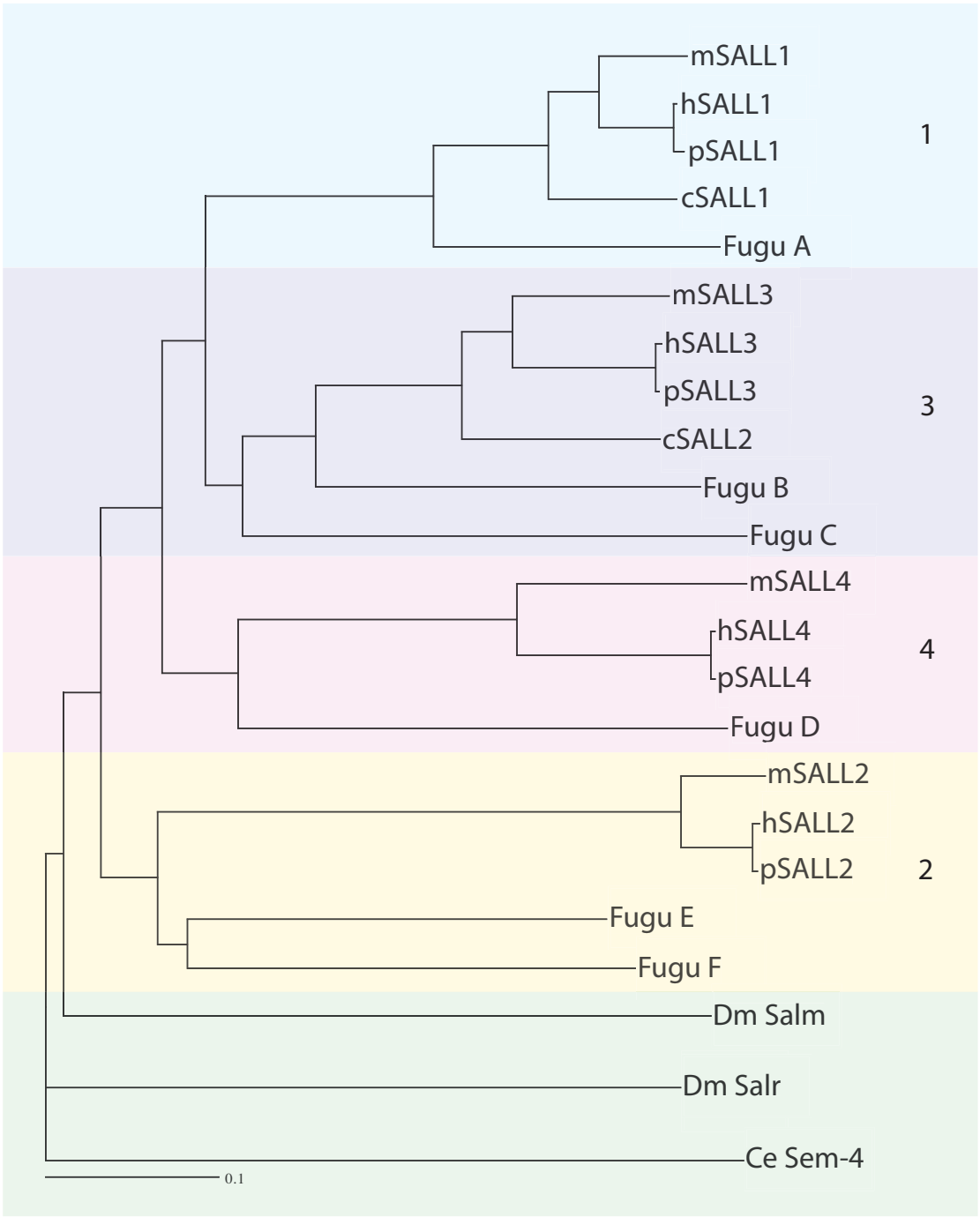


Figure 01: Relative sequence similarity of Spalt proteins

Sequence comparison of Spalt proteins in human (hSALL1-4), chimpanzee (pSALL1-4), mouse (mSALL1-4), chicken (cSALL1-2), Fugu, Drosophila (Dm Salm, Dm Salr) and C. elegans (Sem-4). Sequence was compared in ClustalX, similarity is illustrated by the length of arms separating genes from junctions or junctions from one another.

While completed genomes for human and mouse allow precise alignment, other vertebrate genomes still miss information. For some fugu orthologs, only fragments are known till date and therefore final alignment will have to await completion of the sequencing projects.

Underlaying colores emphasize the four groups of vertebrate spalt genes as well as a distant group of Spalt homologuos genes. In accordance with phylogenetic distances known for human, chimpanzee, mouse and chicken, SALL proteins relate in the subgroup for each paralog. It can therefore be assumed that the four SALL proteins found in mammals were present already before separation into different lineages and that homolog genes can directly be assigned for mammalian genomes.

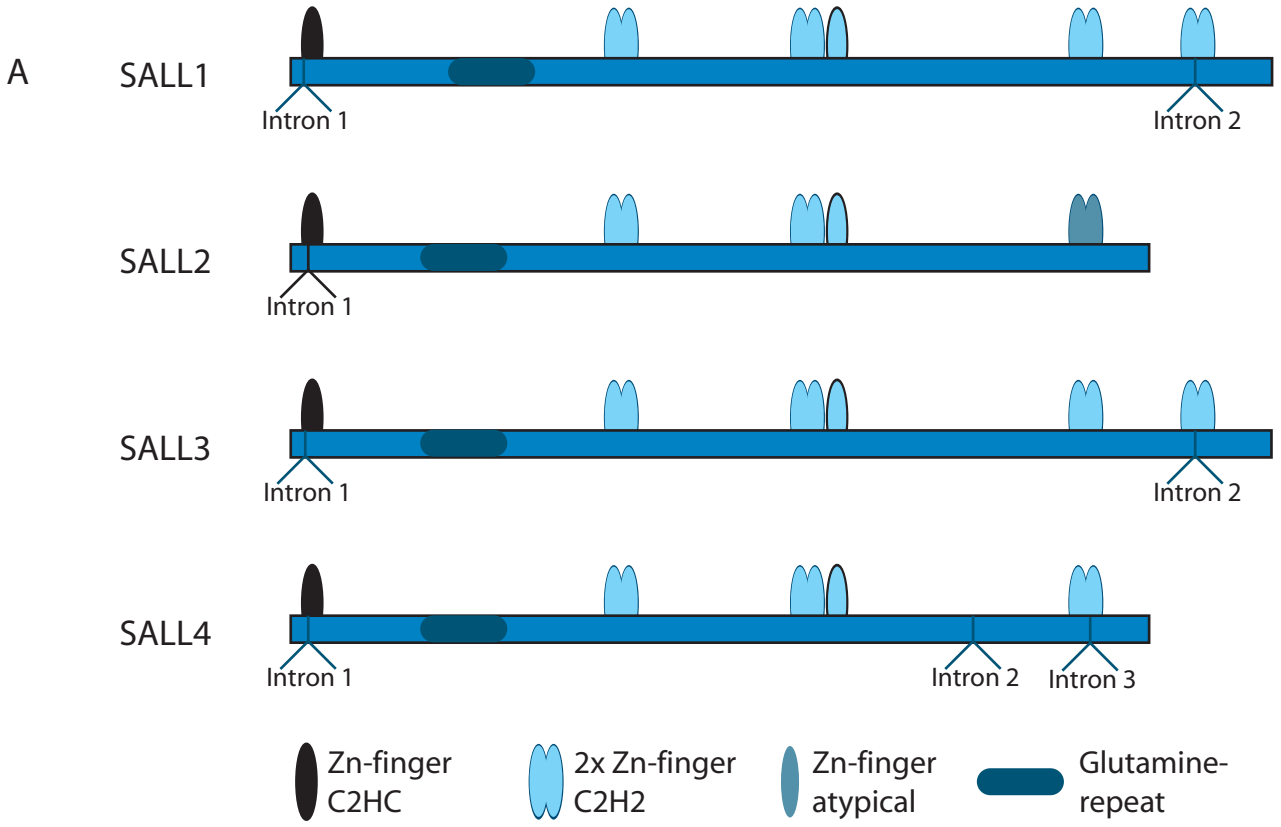
paralogs being less similar to one another than any of the vertebrate alleles are to each other. The presence of four paralog genes was established very early in vertebrate development judging from the presence of closely related orthologs for most *Sall* genes in fish. The presence of two Fugu paralogs in subgroup two and four represents recent duplications in fish genomes. In Danio rerio, Medaka and Fugu the number of paralogs will also exceed the four found in higher vertebrates, once the genome sequencing projects are completed. Already six *spalt* like genes are annotated in the Fugu genome.

Based on sequence homology, *Sall1* and *Sall3* derive from one common ancestor while *Sall4* is more distant. *Sall2* shares the least homology with the other three orthologs. Nevertheless, all four SALL paralogs are almost identical in protein sequence level within the functional domains.

In support of the sequence comparison, close relationship between all four *SALL/Sall* orthologs is also seen in the organization of the mRNA in humans and mice. All four genes are encoded by a small first exon harbouring the translational start followed by a big intron and a second large exon with most of the coding sequence followed by small 3´exons (figure2).

1.2. Protein Structure and Subcellular Localization

Spalt proteins are large proteins of 1000 to 1300 amino acids. Typical for Spalt proteins is the presence of a double zinc finger domain of the C2H2 type, in which two zinc fingers are located directly adjacent to one another. Zinc finger motifs form a distinct three dimensional structure stabilized by the complexation of a zinc ion between cysteine and histidine residues. They are known to bind DNA with the alpha-helical part of the loop between the first two and second two zinc binding sites (Kuhnlein et al., 1994; Pavletich and Pabo, 1991). The double zinc finger motif is repeated three to four times within the protein structure of all Spalt proteins with high similarity even on sequence level. Next to the second double Zn-finger domain, a single Zn-finger is found in all Spalt-like proteins. Moreover, all vertebrate SALL proteins contain an additional N-terminal zinc finger of the C2HC class (Kohlhase et al., 2002b). The N-terminus including the C2HC zinc finger motif had been shown to facilitate repression of target genes by binding and recruiting histone deacetylase (HDAC) (Kiefer et al., 2002), suggesting that Spalt proteins act as transcriptional repressors. However, fusion experiments of XsalF and VP16 in a Xenopus oocyte system indicate that Spalt proteins can also act as transcriptional activators in some contexts (Onai et al., 2004) .



B Glutamine rich domain

```
mSall1  173 KVAVAQF 210SQEARCCGASGG-----KLLISTLMEQLLALQQQQIHQLQLIEQIRHQI LLLASQSADLP
hSALL1  173 KVAVAQF 210SQEARCCGASGG-----KLAVPALMEQLLALQQQQIHQLQLIEQIRHQI LLLASQNADLP
mSall3  140 KVAVAQF 157SQGARAGGTTGAGGSVGAVAIPMILEQLVALQQQQIHQLQLIEQIRSQV ALMSRQ----P
hSALL3  135 KVAVAQF 181SQGARAAGGSGAGGGVAAAAVPLILEQLMALQQQQIHQLQLIEQIRSQV ALMQRP----P
mSall4  129 KVAVNQR 191GAEAPMAPMPAAQ------GIPWVLEQILCLQQQQLQQIQLTEQIRVQV NMWAAH-----
hSALL4  130 KVAVNQR 196SADALPAPVPGAN------SIPWVLEQILCLQQQQLQQIQLTEQIRIQV NMWASH-----
mSall2  134 STPAPPP 162PPPPPPPPPGVGSG----HLNIPLILEELRVLQQRQIHQMQMTVQICRQV LLLGSLG----
hSALL2  134 STPAPPP 162PPPPPPPPPGVGSG----HLNIPLILEELRVLQQRQIHQMQMTEQICRQV LLLGSLG----
```

First double Zn-finger domain

```
mSall1 438 AKSTSDEAFFKHKCRFCAKVFGSDSALQIHLRS HTGERPFKCNICGNRFSTKGNLKVHFQRHKEKYPHIQMNPYPVPEHLDNVPTSIGI
hSALL1 437 AKSTSDEAFFKHKCRFCAKVFGSDSALQIHLRS HTGERPFKCNICGNRFSTKGNLKVHFQRHKEKYPHIQMNPYPVPEHLDNIPTSTGI
mSall3 386 PKASAEDPFFKHKCRFCAKVFGSDSALQIHLRS HTGERPFKCNICGNRFSTKGNLKVHFQRHKEKYPHIQMNPYPVPEYLDNVPTCSGI
hSALL3 408 PKASAEDPFFKHKCRFCAKVFGSDSALQIHLRS HTGERPFKCNICGNRFSTKGNLKVHFQRHKEKYPHIQMNPYPVPEYLDNVPTCSGI
mSall4 375 VKAKDEVVLGKHKCRYCPKVFGTDSSLQIHLRS HTGERPYVCPICGHRFTTKGNLKVHLQRH----PEVKANPQLLAPFQDKGAVSAAS
hSALL4 370 VKPKDEAALYKHKCKYCSKVFGTDSSLQIHLRS HTGERPFVCSVCGHRFTTKGNLKVHFHRH----PQVKANPQLFAPFQDKVAAGNGI
mSall2 360 VISSLEKPGGRHKCRFCAKVFGSDSALQIHLRS HTGERPYKCNVCGNRFTTRGNLKVHFHRHREKYPHVQMNPHPVPEHLDYVITSSGL
hSALL2 361 VMGPLEKPGGRHKCRFCAKVFGSDSALQIHLRS HTGERPYKCNVCGNRFTTRGNLKVHFHRHREKYPHVQMNPHPVPEHLDYVITSSGL
                     * *         *   *     * *         *   *
```

Figure 02: Protein structure of vertebrate SALL genes

A) SALL proteins are transcription factors of 1000-1300 aminoacid length. The most prominent sequence structure found in all spalt genes is a double zinc-finger domain which is repeated several times within the coding sequence with the second repeat being accompained by an additional Zn-finger. The N-terminal Zn-finger scetched in black has been shown to bind histone deacetylase and mediate transcriptional repression. SALL proteins are believed to act as dimers with the glutamine rich stretch serving as dimerzation domain. Human (shown here) and murine Sall genes even show conservation in mRNA structure, as seen by conserved positions of intronic sequence.

B) multiple sequence alignment with ClustalW shows a close to complete conservation of primary sequence between mouse and human homologues within the glutamine rich as well as the Zn-finger domains. These domains are also highly conserved between paralog genes. An asterisk indicates the postions of cystine and histidine chelating the Zn-ion to form Zn-fingers.

In addition to zinc fingers, the only other domain recognizable on primary structure level is a glutamine rich region prior to the first double zinc finger motif. In chicken it has been shown that this domain is necessary for dimerization of Spalt proteins. It is therefore proposed that Spalt proteins function as homo- or heterodimers (Sweetman et al., 2003). Figure 2B illustrates the high level of sequence identity in functional domains of *Sall* genes exemplified by the glutamate rich domain as well as the first double zinc finger motif for human and mouse *spalt* homologous genes.

Zinc finger domains show a very high degree of homology and in *Drosophila*, *salm* and *salr* show strong genetic interaction arguing for redundancy of protein function (Mollereau et al., 2001). Taken together with relative protein homologies, functional redundancy of *Sall* genes in mouse might also be expected.

Studies in cell culture revealed the subcellular localization of SALL proteins. By tagging full length Csal1 and Csal3 it could be shown that Csal1 is found localized to the cell nucleus, while Csal3 remains in the cytoplasm (Sweetman et al., 2003). In a coexpression of both orthologs, Csal1 and Csal3 colocalized and Csal3 tethered Csal1 to the cytoplasm. It remains to be determined though whether cytosolic Csal3 localization was due to modification of the protein by fusion of a flag tag.

Similarly, a study in transiently transfected NIH-3T3 cells, where human *SALL1* was tagged with GFP, revealed localization to pericentromeric heterochromatin (Netzer et al., 2001).

2. Mutations in Spalt-like Genes Cause Inherited Diseases in Humans

Mutations in genes of the *spalt* like transcription factor family have been shown to manifest in phenotypic aberrations in several model organisms. Moreover, humans carrying mutations in *spalt* like genes suffer from various developmental defects.

Two autosomal inherited diseases have been linked to mutations in genes of the Spalt like transcription factor family in humans. While mutations in *SALL2* (*spalt-like*, ortholog 2) and *SALL3* have not been reported to result in pathologic variations, both, mutations in *SALL1* and *SALL4*, result in multiple developmental aberrations inherited in an autosomal dominant fashion.

2.1. Townes-Brocks Syndrome (TBS)

Patients suffering from Townes-Brocks Syndrome display a combination of anal, renal, limb and ear abnormalities (Townes 1972).The clinical presentation is highly variable even within families. Patients typically display outer ear deformities, an imperforate anus and sensineural hearing loss (Townes 1972) combined with several other conditions. The most common limb defects in TBS are preaxial polydactyly, syndactyly and club feet. Reported renal malformations can lead to renal failure (Rossmiller and Pasic 1994). Typically one or both kidneys are hypoplastic (Kohlhase et al 1999). Further rare conditions are mild mental retardation, a hypoplasia of the corpus callosum, coloboma as well as heart defects. Various genital malformations like vaginal aplasia, bifid uterus and scrotum as well as chryptorchidism have also been reported. The frequency of TBS could not be determined to date since penetrance seems incomplete but due to its dominant nature it is a rare syndrome.

TBS is inherited in an autosomal dominant fashion. Therefore haploinsufficiency of *SALL1* has been discussed as a likely explanation for TBS (Kohlhase et al., 1998) but formally a dominant negative or gain of function effect can not be excluded.

35 independent mutations in TBS patients have been mapped and all but two cluster prior or in the first double zinc finger motif and found to result in truncated SALL1 protein. Other mutations include a variety of deletions and insertions (Botzenhart et al., 2005; Kohlhase et al., 1999). Taken together with the finding that SALL proteins form homo- and heterodimers through the glutamine rich domain which lies N-terminal to the to the cluster of truncation sites in TBS patients, a dominant negative effect was proposed (Sweetman et al., 2003). Until date, no correlation between the molecular basis and site of the mutation with the clinical presentation could be drawn (Botzenhart et al., 2005) and high phenotypic variability is observed even within families.

2.2. Okihiro Syndrome/Radial Ray Syndrome

Mutations in the most recently identified member of the Spalt class of transcription factors, *SALL4*, have recently been described to result in a range of clinically overlapping phenotypes. The two most commonly observed effects are eye movement disorders and limb malformations. Patients diagnosed of Okihiro Syndrome, Radial Ray Syndrome (DRRS), Holt-Oram Syndrome, Acro-Renal Ocular Syndrome and Thalidomide Embryopathy have been found to carry mutations in SALL4 (Al-Baradie et al., 2002; Borozdin et al., 2004;

Kohlhase et al., 2005; Kohlhase et al., 2002b; Kohlhase et al., 2003). These patients typically display limitation of eye abduction associated with retraction of the eye globe as well as absence or dysplasias of the radius and fingers, most often the thumb. Close examination also reveals slight deformities of the ears, hearing loss in some cases, and kidney hypoplasia. Overall, a clear overlap to clinical manifestation of Townes Brocks Syndrome is apparent. The high variability of the expressivity is also a common feature of the syndromes described above.

Initial reports (Kohlhase et al., 2002b) found a clustering of mutations in *SALL4* that will result in premature truncation prior to the first double zinc finger motif similar to mutations seen in *SALL1*. Later reports also see large deletions of the whole *SALL4* locus in DRRS patients (Borozdin et al., 2004), arguing that in contrast to *SALL1,* mutations in SALL4 do indeed manifest in a haploinsufficient fashion. Further studies with higher numbers of patients analyzed will clarify this observation.

In summary, clinical manifestations of *SALL* mutations are highly variable and affect sensory organs as well as CNS, kidney and limb development. Phenotypes in TBS and DRRS vary and a clear genotype-phenotype correlation is not evident (Kohlhase et al., 1999).

The molecular effect of mapped mutations is not understood, one would for example expect, that oligodactyly versus polydactyly is caused by different types of mutations. Nevertheless, a correlation between the molecular effect of a mutation and the phenotypic manifestation is not clear. Due to the complicated genetics, no underlying developmental principal was identified till date that could account for TBS and DRRS.

3. Mutations in Spalt and Sem-4

Spalt proteins were first described in *Drosophila melanogaster* by virtue of its capacity to promote terminal structures (Tsonis and Goetinck, 1988). *Spalt major* (*Salm*), the founding member of this protein family, has been shown to play important roles in multiple steps of development (Dong et al., 2003) such as sensory organ development, the anterior-posterior boundary formation in the wing disc as well as determination of trunk identities. Mutations in *spalt* cause head and tail segments to develop trunk structures (Kuhnlein et al., 1994). This publication also is the first one to report the protein sequence and the one to identify the Zn-finger domains since the transcriptional unit had previously been misidentified.

Further, mutation of *salm* and *salr* in the eye indicates a requirement of *spalt* genes for development of photoreceptors R3/R4 (Domingos et al., 2004b) and R7/R8 (Domingos et al.,

2004a). Loss of *spalt* genes in the developing tracheal system as well as their overexpression result in defects in tracheal branching and morphogenesis (Ribeiro et al., 2004).

A null allele of the C. elegans homolog *sem-4* (*sex myoblast*) causes a failure in sex myoblast differentiation and neuronal determination (Basson and Horvitz, 1996). In addition, several alleles exhibit vulval cell fate determination defects (Grant et al., 2000). *Sem-4* function is required for expression of *lin-39*, the *Drosophila sex combs reduced* homolog in *C. elegans*. Together with a genetic interaction of *spalt* and *antennapedia* as well as other studies from *Drosophila* and mammals, it has been proposed that *spalt* genes may be part of a bigger complex required for general transcriptional silencing (*polycomb* complex) and activation (Franch-Marro and Casanova, 2002; Toker et al., 2003). Localization of Spalt to heterochromatin as discussed earlier might fit this idea.

4. Spalt-like Transcription Factors and Cell Signalling

Till date, induction of *spalt* transcription has been linked to a number of different signalling pathways. While *Salm* is repressed by *antennapedia* in the leg imaginal disc (Wagner-Bernholz et al., 1991), expression of *spalt-major* and *spalt-relates* (*salr*) is induced in response to *decapentaplegic* (*dpp*), a *BMP4* homolog, in the wing disc (Barrio R, 1996; de Celis and Barrio, 2000) and required for the positioning of veins in the wing disc. *Spalt* expression in the tracheal system is induced by *wingless* (Ribeiro et al., 2004). *Wnt* genes in Xenopus have also been reported to activate homologs to *spalt* that in turn will antagonize canonical *Wnt* signalling (Onai et al., 2004). Contrary to that, expression of *Sall1* in cell culture potentiated Wnt signalling (Sato et al., 2004). In the fish Medaka, *spalt* is discussed as a target of hedgehog signalling (Koster et al., 1997). It appears from these studies that Spalt function can not be assigned into a single cell signalling pathway but rather is involved in a more general fashion to facilitate a series of developmental processes. One model generated in Xenopus suggests that SALL proteins facilitate general cellular responsiveness to cell signalling inputs (Onai et al., 2004). How responsiveness might be achieved is not understood.

5. Aim of this Study

Mutations in genes of the *spalt-like* family manifest in multiple developmental phenotypes. Genetic interaction of Spalt proteins with known cell signalling pathways can not explain the observed general cellular responses. Function of *Sall* genes could neither be attributed to specific cell types nor specific cellular pathways. Especially, observations in mammals are too few and – in case of humans - difficult to interpret. Mutations of human *SALL* genes lead to a variety of different malformations arguing that SALL transcription factors are involved in multiple developmental processes in mammals. A deeper understanding of the role of Sall genes in mammalian development will therefore gain insights into general development as well as elucidate the underlying principles of Townes-Brocks Syndrome, Duane-Radial Ray Syndrome and related diseases.

Mouse is the closest system to humans that is accessible to reverse genetics. Therefore, I decided to study the effects of loss of *Sall* orthologous genes in *Mus musculus* by means of genetic manipulation. The fully sequenced genome and a good correlation of *spalt* orthologs between mice and men facilitate such an approach.

Mice homozygous mutant for *Sall1* do not show the phenotypes observed in patients suffering from TBS. Rather, *Sall1* knockout mice are born without kidneys or with severe renal hypoplasia. Defective kidneys are supposed to lead to the perinetal lethality in these animals (Nishinakamura et al., 2001). No other phenotypes due to the loss of *Sall1* have been described and heterozygous mice show no defects. The allele used for this study is believed to be a genetic null while mapped mutations in humans typically lead to truncations of the SALL1 or SALL4 proteins. This argues that humans are either more sensitive to the level of SALL protein or the nature of these mutations is acting in a dominant negative or gain of function manner. Dimerization of Spalt proteins has been reported previously (Sweetman et al., 2003) and could explain such effects.

Further support that expression of a truncated *Sall1* protein rather than haploinsufficiency causes the defects seen in TBS patients comes from the observation that expression of a truncated version of *Sall1* in mouse results in stronger phenotypes, more similar to the malformations seen in humans (Kiefer et al., 2002). Like in humans, the expressivity of the phenotype is low, variable and difficult to study.

If the pathology observed in humans is indeed due to expression of truncated Spalt protein that forms homo- or heterodimers with full length Spalt protein, genetic analysis has to be evaluated carefully. It is not clear if and to what extend dimerization occurs (Sweetman

et al., 2003), if all *Sall* genes heterodimerize and if or how truncated protein interferes with the function of full length protein. In addition, expression of Spalt proteins has never been studied carefully or comparatively in mouse to ask the question, which *Sall* genes may interact.

Clean analysis of the effects of a loss of SALL activity requires a well controlled genetic system as well as good phenotypic penetrance. Based on the high degree of protein identity and limited phenotypes observed in a mouse knockout system I reasoned for possible redundancy in function amongst the different Sall proteins in mouse. To assess redundancy, expression needs to be taken in account and so I intended to study it in a precise and comparative manner.

The design of this study further included the genetic disruption of each known *Sall* paralog in the mouse genome. Based on the expressional analysis, I intended to study phenotypic effects in murine development caused by loss of all Sall activity in the tissue under observation. By comparison of several independent developmental systems I intended to link Sall function to a general cellular attribute that would be able to interpret several of the observed phenotypes caused by mutations in *Sall* genes.

This approach generated phenotypes with full penetrance that enables the investigation of cellular responses to SALL proteins. As a result, I have identified a novel function for *Sall* genes in maintenance of pluripotency within progenitor populations.

Results

1. Comparison of Sall Proteins

1.1. Comparative Expression Analysis of Sall mRNA Transcripts in Mouse

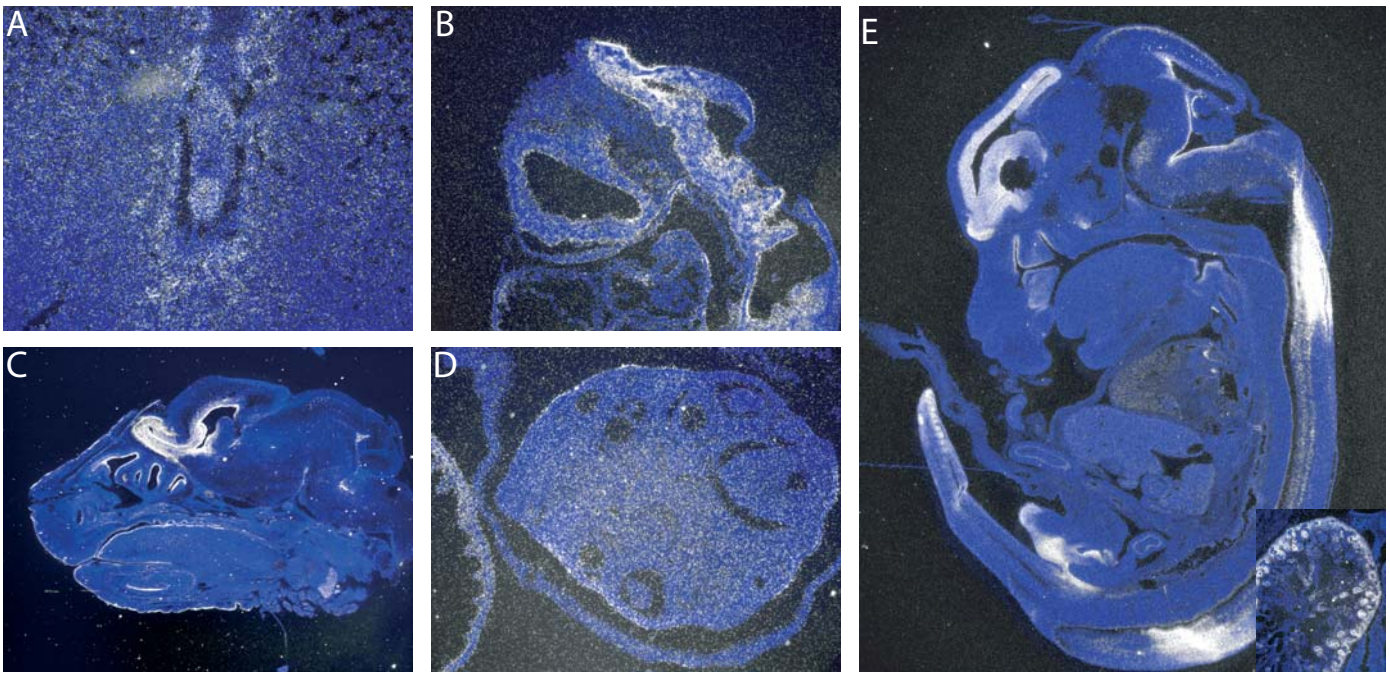
In order to get a better overview of all four Spalt like proteins in mouse the decision was taken to compare their expression patterns. Due to the high degree of conservation of the protein sequence amongst the different orthologs of the *spalt* family, a genetic loss of function analysis of their role in development can only be investigated when taking precise expression patterns of all four orthologs into account. Spatially and temporally overlapping expression of two or more orthologs could mask the developmental effects of a loss of one *Sall* gene.

A comprehensive expression analysis of the mouse *spalt-like* mRNA transcripts was performed by *in situ* hybridization at various developmental stages. To reach maximal spatial resolution and optimal comparability, adjacent histological sections were processed side by side for each paralog. Stages analyzed included e6.5 pre-gastrulation implantation sites, e9.5 post turning embryos, e11.5, e12.5, e13.5, e15.5 and P0 as well as various adult tissues.

Sall1 transcripts are barely detectable at e6.5 but start to be strongly transcribed shortly after. At e9.5, strong expression can be seen in the neuroectoderm (figure 3B). With further growth and differentiation of the embryo (e13.5, figure 3E), the transcript is found mainly in the central nervous system around the ventricles as well as in pituitary gland, thyroid gland and kidney (see insert e15,5 in E). Strongest expression can be seen in the developing telencephalon. It is interesting to note that expression within here restricts more and more to the ventricular and subventricular regions (figure 3C). Generally it can be said that expression is seen in tissues where progenitor populations are expected to reside. Adult tissues do not express *Sall1* at high levels with exception of the olfactory bulbs.

In contrast to *Sall1*, *Sall2* is expressed more broadly in space and time. Where both genes are expressed, levels of *Sall2* transcript are generally lower than those seen for *Sall1*. First expression can be seen in the embryo proper region of an e6.5 pre-gastrulation embryo. It will then remain expressed all through development within the CNS, but in other tissues *Sall2* transcript detection fades away between e13.5 and birth. At e15.5 (J, insert), expression

Sall1



Sall2

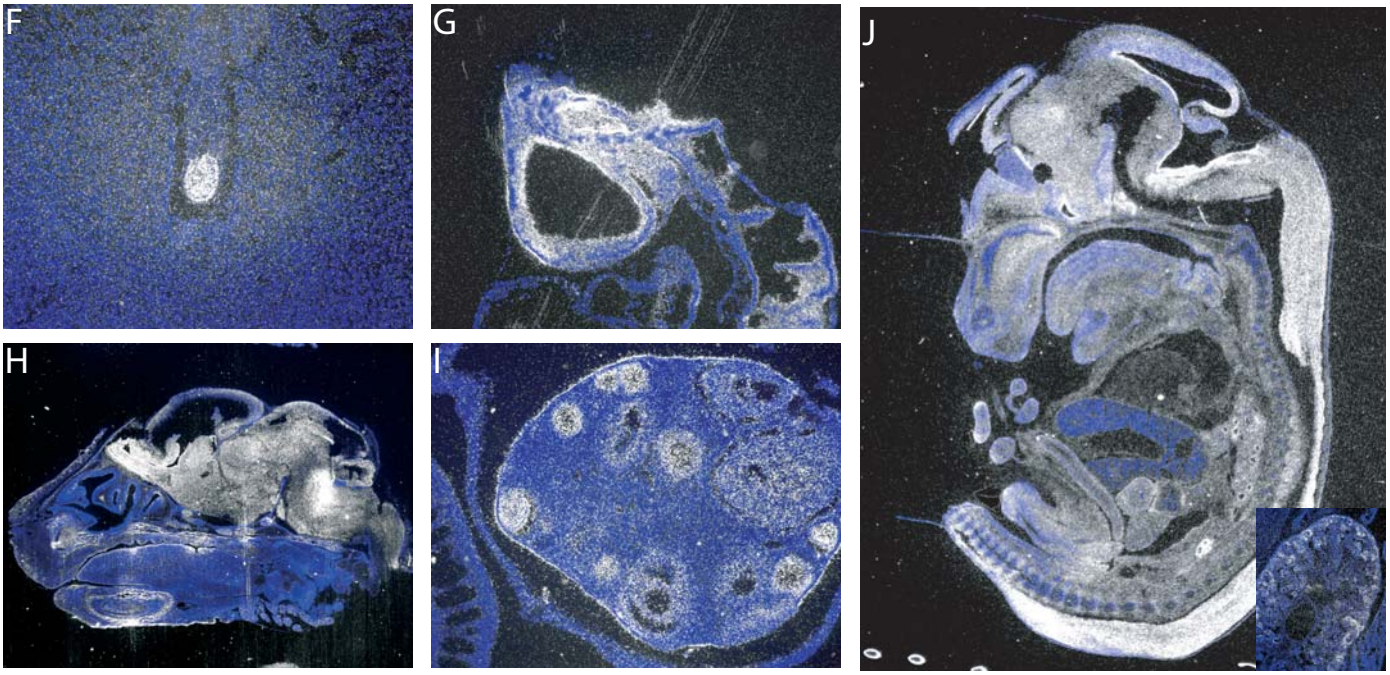
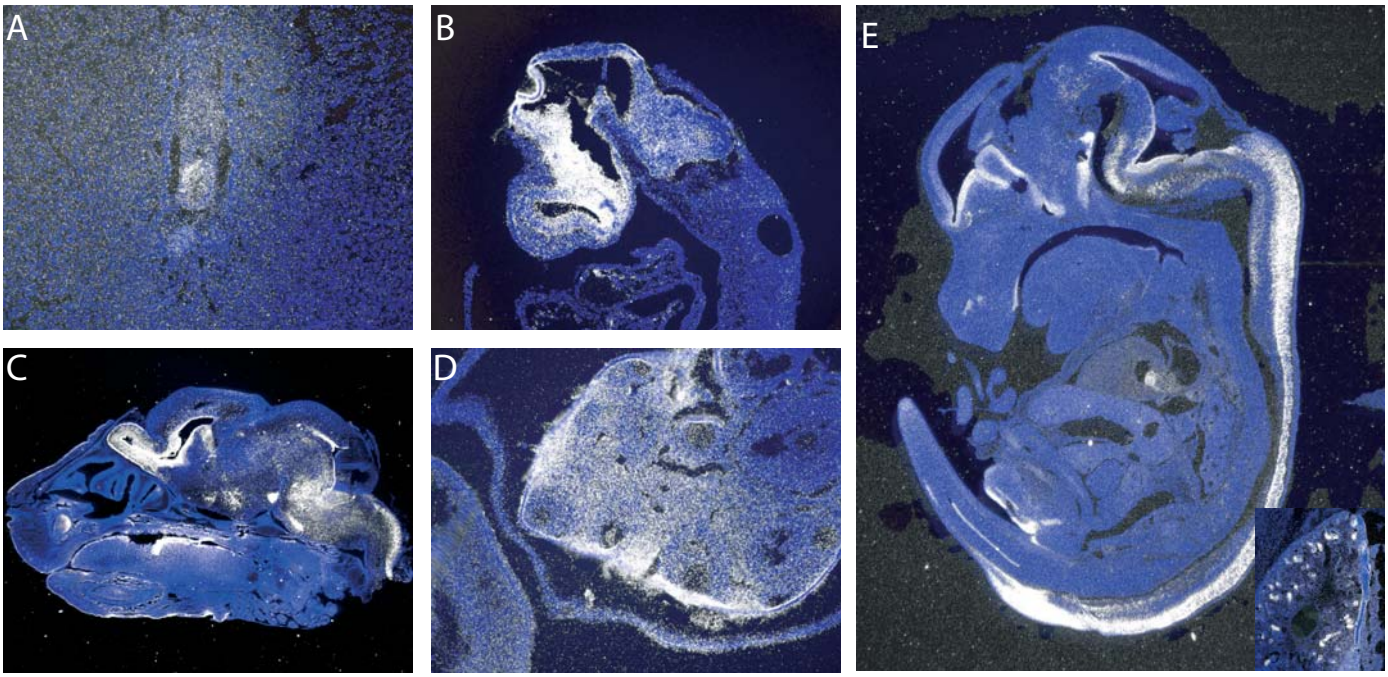


Figure 03: Comparative expression analysis of *Sall1* and *Sall2*

In situ hybridisation of *Sall1* and *Sall2* on cryostat sections. Tissues shown are e6.5 embryos including extraembryonic and maternal tissues (A, F), as well as sagittal sections of embryos at stages e9.5 (head region: B, G), e13.5 (E,J), e19.5 (C, H) as well as adult ovary (A, F). Autoradiographic staining indicating expression appears white in dark field microscopy. *Sall1* expression was observed very weakly in the embryo proper at e6.5 as well as in the central nervous system in later stages restricting successively to the ventricular zones. Renal expression was observed on lateral sections (inserts in E, J are e15.5). *Sall2* is expressed in the whole embryo proper from e6.5 onwards with exeption of the liver.

Sall3



Sall4

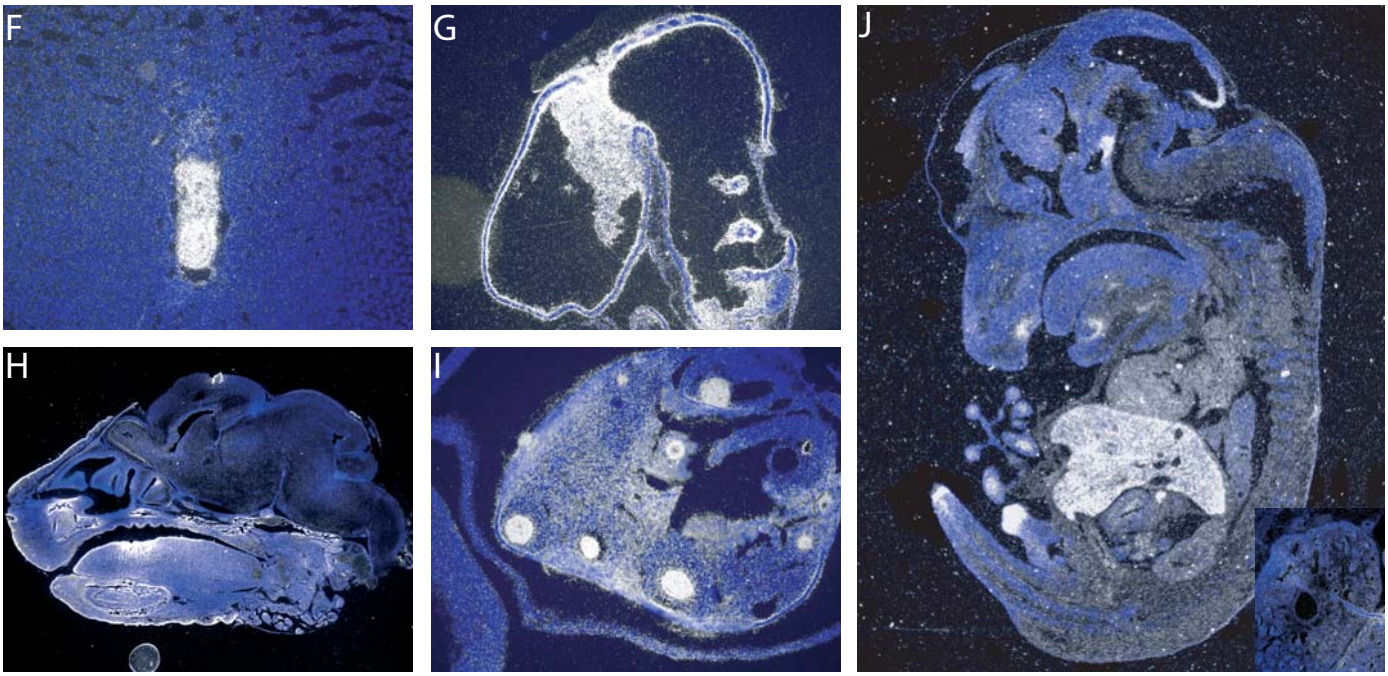


Figure 04: Comparative expression analysis of *Sall3* and *Sall4*.
In situ hybridisation of *Sall3* and *Sall4* on cryostat sections. Tissues shown are e6.5 embryos including extraembryonic and maternal tissues (A, F), as well as sagittal sections of embryos at stages e9.5 (head region: B, G), e13.5 (E,J), e19.5 (C, H) as well as adult ovary (A, F). Autoradiographic staining indicating expression appears white in dark field microscopy. *Sall3* transcript is not present in the adult ovary (D) and only weakly expressed at e 6.5 (A). At later stages, expression was detected in brain and spinal cord as well as heart. Kidney with high level of transcript can not be seen on the midline section but in the insert taken from e15.5.
Sall4, the only spalt ortholog expressed in the maturing oocyte (I) and can be detected in all inner cell mass derivatives at e6.5 (F). At e9.5 (G) the transcript is still present at high levels in every tissue. from then onwards, expression declines and restricts mainly to the liver (J).

is still visible in the kidney. The only tissue, that *Sall2* transcript is excluded from is the liver. In adult mice, *Sall2* can still be detected in olfactory bulb, hippocampal neurons and granulosa cells in the ovary (figure 3I).

Sall3 expression is reminiscent of *Sall1* expression (figure 4A-E). First strong expression was visualized e9.5. Transcription will then restrict mainly to the CNS again (C, E). Notably, only weak expression can be seen in the telencephalon at e13.5 while diencephalic and paleocortic precursor cells display active transcription. After completion of neurogenesis, *Sall3* expression restricts to the ventricular zone of the neocortex at birth (figure 4C). Like *Sall1* and *Sall2* the third family member too is expressed in the olfactory bulbs at high levels and remains to be also in adult animals. Expression outside the central nervous system localizes to heart and mesoderm derived structures in the kidney but not its cortex.

Sall4 differs considerably in temporal expression. For the three other genes, transcription peaks around mid-gestation and generally fades out until birth. *Sall4* is expressed much earlier. Strong expression can already be seen in post implantation embryos. Cells staining positive for *Sall4* localize to embryo proper, extraembryonic mesoderm and extraembryonic endoderm. Post implantation, *Sall4* transcription soon starts to be down regulated. E9.5 embryos still express high levels of *Sall4* but already two days later at e11.5 transcription disappears in most tissues. In e13.5 embryos expression is weak and remains active mainly in the liver (figure 4J). *Sall4* thus expresses at high level in tissues that show no expression of other *spalt* genes like extraembryonic mesoderm and extraembryonic endoderm. At later stages of development, expression is neither detectable in kidney nor in CNS. In adult mice, strong *Sall4* transcription was detected in developing oocytes (figure 4I)

Whether conservation of phenotypes caused by loss of a gene product can be expected in different species depends on their expression patterns *in vivo*. Promoter regions change at a higher rate than open reading frames making predictions across species difficult. I therefore decided to compare expression patterns as an approximation based on unigene clusters. The unigene library contains a big number of expressed sequence tags (ESTs), mapped to genes and sorted by the tissue that they were cloned from. Information is publicly available. I compared the origins of expressed sequence tags mapping to *Sall4* normalized to the total number of clones. Data were obtained from unigene webpage based on ESTs included until 20th of September 2005.

Table 1 shows a list of tissues with the relative expression level for each ortholog in mouse and human based on EST hits in the respective tissue. A comparison of presented *in situ* data with the abundance of mouse unigene clones indicates that a tendency of expression is also represented in unigene cluster quantification. For example mouse *Sall1*-3 transcripts are found in kidney while mouse *Sall4* is the only transcript present in liver. Therefore, a tendency for expression can be retrieved from this comparison. An overall similarity in expression levels between human and mouse can be seen, e.g. *Sall4* shows especial high abundance in embryonic stem cells in both organisms while all other members are found in the brain. *Sall1* in humans is not comparable due to a very high number of clones obtained from the liver.

	Human		Mouse	
Sall1	non cancerous liver tissue:	97.67	Undifferentiated ES Cell:	27.81
	soft tissue:	1.04	uterus:	21.49
	kidney:	0.32	unclassified:	16.40
	pluripotent cell line derived from blastocyst inner cell mass:	0.26	Embryonic Stem Cell:	14.54
	cervix:	0.18	brain:	6.18
	liver:	0.09	pituitary gland:	3.33
	placenta:	0.08	other:	3.23
	other:	0.08	kidney:	2.72
	muscle:	0.06	spleen:	2.61
	brain:	0.05	whole body:	1.70
Sall2	embryonic stem:	32.51	Pituitary gland:	17.08
	Embryonic Stem cells:	10.05	eye:	14.84
	human embryonic stem cells:	7.34	brain:	13.93
	peripheral nervous system:	7.23	pancreatic islet:	7.83
	pluripotent cell line derived from blastocyst inner cell mass:	6.90	other:	6.61
	bladder:	6.34	ovary:	5.29
	thyroid:	5.76	mixed:	4.99
	small intestine:	3.96	Embryonic Stem Cell:	4.15
	pancreatic islet:	3.83	lung:	4.12
	brain:	3.80	lymph node:	3.93
Sall3	pluripotent cell line derived from blastocyst inner cell mass:	42.42	ES Cell:	30.66
	brain:	15.89	brain:	26.07
	eye:	9.47	kidney:	22.90
	kidney:	7.38	other:	10.41
	blood:	6.80	whole body:	5.74
	other:	6.62	eye:	4.22
	prostate:	6.42		
	testis:	2.86		
	skin:	2.14		
Sall4	Embryonic Stem cells:	67.58	embryo, late gastrula:	25.73
	human embryonic stem cells diff. to an early endod. cell type:	18.09	Undifferentiated ES Cell:	20.90
	small intestine:	5.33	Embryonic Stem Cell:	20.48
	pluripotent cell line derived from blastocyst inner cell mass:	3.09	unclassified:	10.12
	other:	1.77	Trophoblast stem cell:	9.14
	testis:	1.25	Whole body:	4.02
	ovary:	0.97	pancreatic islet:	2.58
	muscle:	0.73	liver:	1.47
	mixed:	0.53	other:	1.32
	skin:	0.47	heart:	1.10

Table 1: Abundance of spalt transcripts among expressed sequence tags (ESTs) in human and mouse: Values are normalized to total number of sequenced ESTs from each tissue. Data were retrieved from unigene (September 20th, 2005).

On a northern blot, expression of human *SALL1* has been reported for the fetal brain but not the fetal liver. Strong expression could also be observed in adult endocrine organs like pituitary gland, adrenal gland and the gonads (Ma et al., 2002). These data fit well with my observations in mouse but they contradict the high number of clones unigene retrieved from the liver.

1.2. Expression Domains of Spalt-like Proteins Overlap.

The expression analysis shows, that *spalt-like* genes strongly overlap in their expression profile. At e9.5 all four *spalt* transcripts are very abundant and widely expressed. Especially in the CNS, overlapping expression is seen until birth. At e13.5 to P0 overlapping prominent expression is seen in ventricular zones of the central nervous system including the spinal cord. Strong expression in the olfactory bulb is observed only in *Sall1, Sall2* and *Sall3*, as opposed to a very low abundance of *Sall4* transcript in mitral cells (figure 4H) *Sall1-3* will still be expressed in adulthood. Notably, the roof of the lateral ventricles, giving rise to the neocortex, shows expression of *Sall1* and *Sall2* while *Sall3* expression is weak. A strong overlap of expression of *Sall1*-3 can also be observed in the developing kidney.

Sall4 is expressed earlier than the other three *Sall* members and is detected already in female gametes. At e6.5 and e9.5, the expression level exceeds that of the paralogous transcripts.

To understand the developmental processes, in which *spalt* genes are involved *in vivo*, organ systems of prominent *spalt* expression like the brain were studied. This approach was complicated by the fact that high levels of expression coincided with a high degree of coexpression with other members of this transcription factor family. It has been suggested before, that Spalt proteins have redundant functions due to the high degree of sequence similarity. To overcome a potential masking of the loss of function phenotype by the presence of other members of the *spalt* family it was thus necessary to have null alleles for each *spalt* gene available and to delete all Spalt activity in the tissues examined.

2. Generation of Mouse Lines

Previous to the presented study, a mouse line genetically null for *Sall1* had already been generated by our lab (Mathias Treier, unpublished). The second exon of *Sall1* had been excised and the sequence was replaced by enhanced green fluorescent protein (EGFP) fused in frame to the remaining amino acids of *Sall1*. 3´ to that in the transcript is an internal ribosomal entry site (IRES) and lacZ encoding for β-galactosidase. Such a setup should enable me to visualize cells expressing this bicistronic construct either through GFP or an enzymatic colour reaction of β-galactosidase in a cellular resolution. I will herein refer to this allele as *Sall1-lacZ.*

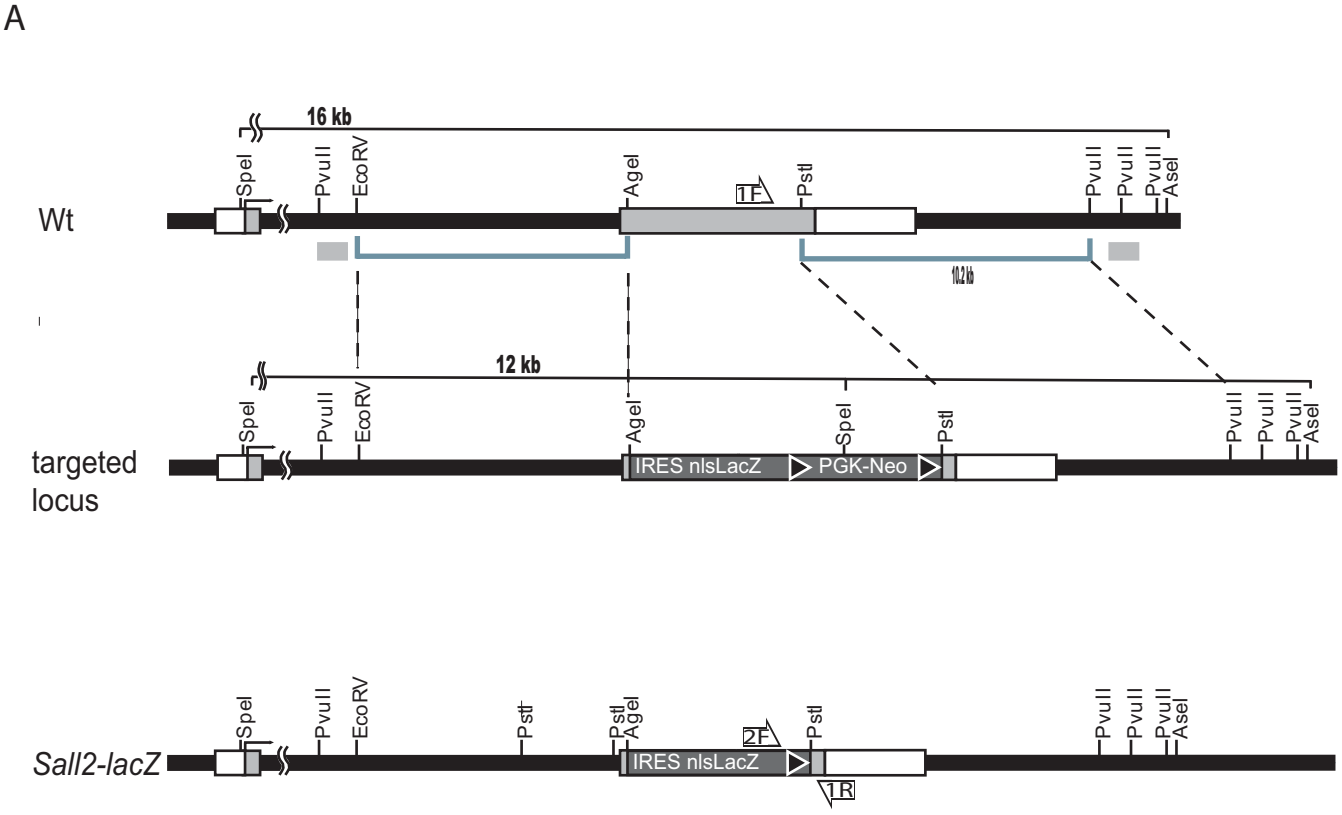
Mice heterozygous for *Sall3* (Parrish et al., 2004) were provided by Prof. Dr. Schütz, DKFZ Heidelberg, as frozen blastocysts and rederived by the EMBL transgenic service. The *lacZ* fused to the *Sall3* coding region only results in very weak expression of β-galactosidase. I will refer to this allele as *Sall3-lacZ*

2.1. Knockout Vector Design for Sall2

In order to obtain a *Sall2* null allele, the genomic region of *Sall2* was cloned. I screened a chromosomal mouse BAC (bacterial artificial chromosome) library with a *Sall2* exon two probe and identified three BAC clones that contained the *Sall2* genomic region, the identification numbers are 15P22, 137K18 and 177B13. A seven kilobase PstI fragment as well as a twelve kilobase XbaI fragment was subcloned and used for further assembly of the *Sall2* knockout vector.

The coding sequence for *Sall2* is distributed on two exons with 24 amino acids being encoded on the first exon. It is therefore reasonable to assume that a deletion of the second exon will result in a genetic null allele. Facilitating an AgeI site at the 5´end of exon two, it was possible to generate a vector that terminates the *Sall2* open reading frame (ORF) after 61 amino acids (figure 5). The C2HC amino-terminal zinc finger is just encoded within the first exon and the few amino acids remaining from exon two. The glutamine rich domain and all double zinc-finger motifs are encoded on the second exon and will be deleted. The knockout vector contained a positive selection cassette conferring neomycin resistance flanked by loxP sites, a negative selection marker encoding diphtheria toxin A to select against random integration events, a 5820 base pair 5’ arm and a 4153 base pair 3’ arm, to enable homologous

Targeting Strategy *Sall2-lacZ*



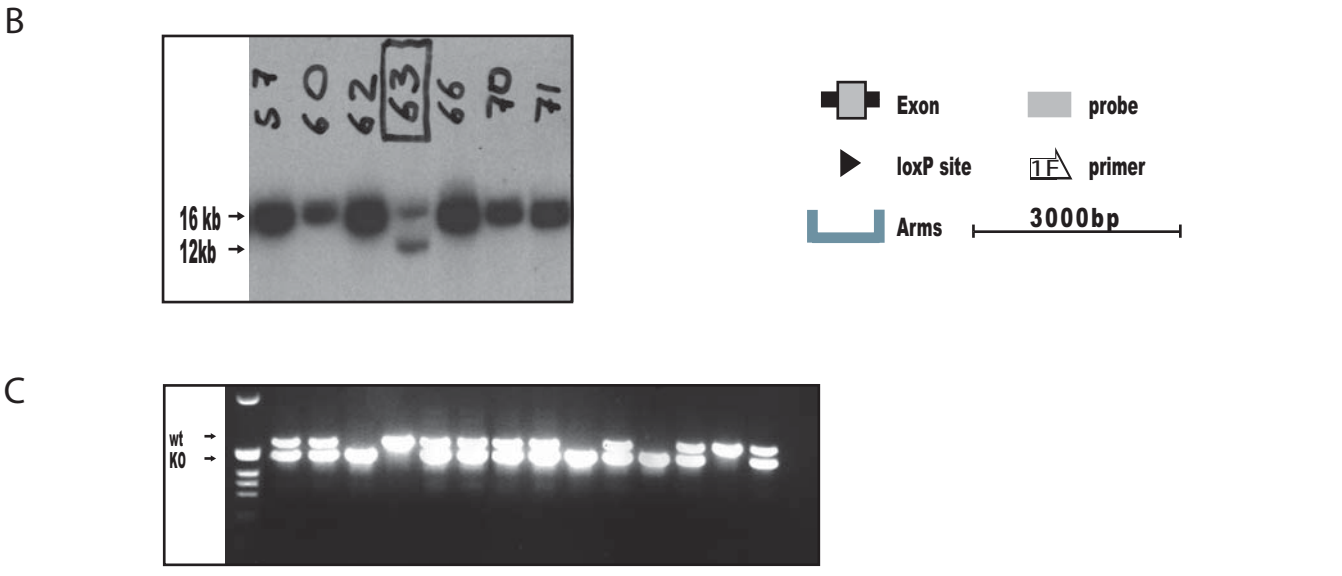


Figure 05: Targeting strategy for *Sall2-lacZ*.

A) Exon two carries the coding sequence for the glutamine repeat as well as all double zinc fingers. A replacement vector carried an IRES-lacZ as well as a positive selection cassette (PGK-Neo) between the two arms. Homologous recombination resulted in genetic removal of exon two and introduces a lacZ allele. Cre-mediated excision of PGK-Neo generated the allele termed Sall2-lacZ.

B) Screening for homologous recombination was done by digestion of genomic DNA with SpeI and subsequent southern blotting and probing with a probe 5´external to the 5´arm (grey box).

C) Interbreeding mice heterozygous for *Sall2-lacZ* revealed that *Sall2* loss does not lead to any obvious phenotype in an experimental environment. Genotyping done with the primers indicated in A.

recombination. The restriction sites used to clone the 5´and 3´arm were EcoRV-AgeI and PstI-PvuII respectively. Inserting an IRES lacZ cassette into the remaining transcript enables expression analysis of the *Sall2* locus.

2.2. Establishment of a Mouse Line Carrying a Sall2 Null Allele

The knockout vector was electroporated into IB10 embryonic stem cells (ES cells). G418 resistant clones, which had incorporated the knockout vector, were picked after 10 days of selection and cultured individually in a 96 well plate. To screen the ES-cell clones for homologous recombination at the *Sall2* locus, clones were split in two replicas, expanded and DNA was purified from one replicate. A Southern blot strategy was devised based on a SpeI-AseI digest followed by hybridization with a 3' external probe as shown in figure 5. A total of 240 ES cell colonies were picked. One clone had undergone positive homologous recombination at the *Sall2* locus resulting in a shift of the band containing the *Sall2* locus from 16 kilobases to 12 kilobases.

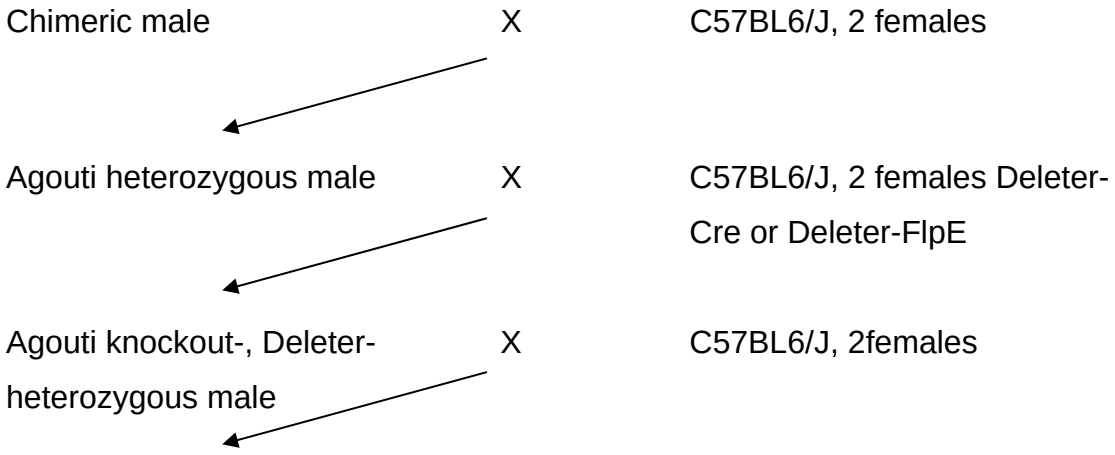


Table 2: Breeding scheme illustrating the sequence of crossings to remove selection cassettes used for positive selection in cell culture. To enable genetic excision, selection markers are located between loxP or FRT sites, recognized by recombinases Cre-recombinase and flippase respectively. For the *Sall2-lacZ* and *Sall4-GFP* allele, cassettes were removed with Cre; FlpE recombination was used for *Sall4-loxP.*

The clone was expanded, trypsinized and always twelve ES cells were injected into one of approximately 50 e3.5 blastocysts by the EMBL transgenic facility and all living blastocysts were subsequently transferred into pseudopregnant females to obtain chimeric animals. Male chimeras that had high contribution of the injected ES cells, which is reflected

by a high agouti to black coat colour ratio, were bred to black Bl-6 females. Agouti offspring indicates contribution of the modified ES cells to the germ line and those pups were screened for the presence of the engineered *Sall2* allele. Heterozygous animals were then bred into a *Deleter-Cre* line (Schwenk et al., 1995), a mouse strain expressing bacterial Cre recombinase in the germ line.

The positive selection cassette used for targeting *Sall2* is flanked by loxP sites, which are 34 base pair recognition sites for Cre recombinase. In mice carrying a *Sall2* allele and the *Deleter-Cr*e transgene, recombinase will excise any sequence between tandemly repeated loxP sites in the germ line. Mice heterozygous for both engineered alleles were subsequently bred back to wild type to obtain the desired final genotype. Thus, a replacement of the second exon of *Sall2* by a traceable marker with minimal further modifications in the genome was achieved. Successful recombination was analyzed by PCR; the allele will be called *Sall2-lacZ* from here on.

2.3. Mice Mutant for Sall2 are Viable and Fertile

Intercrossing mice heterozygous for *Sall2* resulted in 25% homozygous offspring (figure 5C). These mice were viable and showed no obvious phenotypes even at close examination, neither in Bl-6 nor in Bl-swiss background. The life expectancy is not altered. Furthermore, breeding pairs of individuals lacking the *Sall2* gene had normal litter sizes. It thus had to be concluded that *Sall2* function is not necessary for mice in the experimental environment, possibly due to expression of other *Sall* members with redundant function.

2.4. EGFP Knockin of Sall4

The exon-intron structure of *Sall4* also suggested that deletion of almost the entire second exon would result in a null allele. Similar to the strategy used for *Sall2*, a BAC library was screened, to identify a clone containing the genomic region of *Sall2*. The 5´ and 3´ arms used for homologous recombination were subcloned with SalI-AgeI and SpeI-KpnI respectively and assembled to a knockout vector. Positive and negative selection was facilitated by the same neomycin and diphtheria toxin A cassettes. The arms used for homologous recombination had a length of 2646 and 4261 base pairs respectively and would result in a remaining open reading frame encoding for 58 amino acids thereby deleting all zinc

Targeting Strategy *Sall4-GFP*

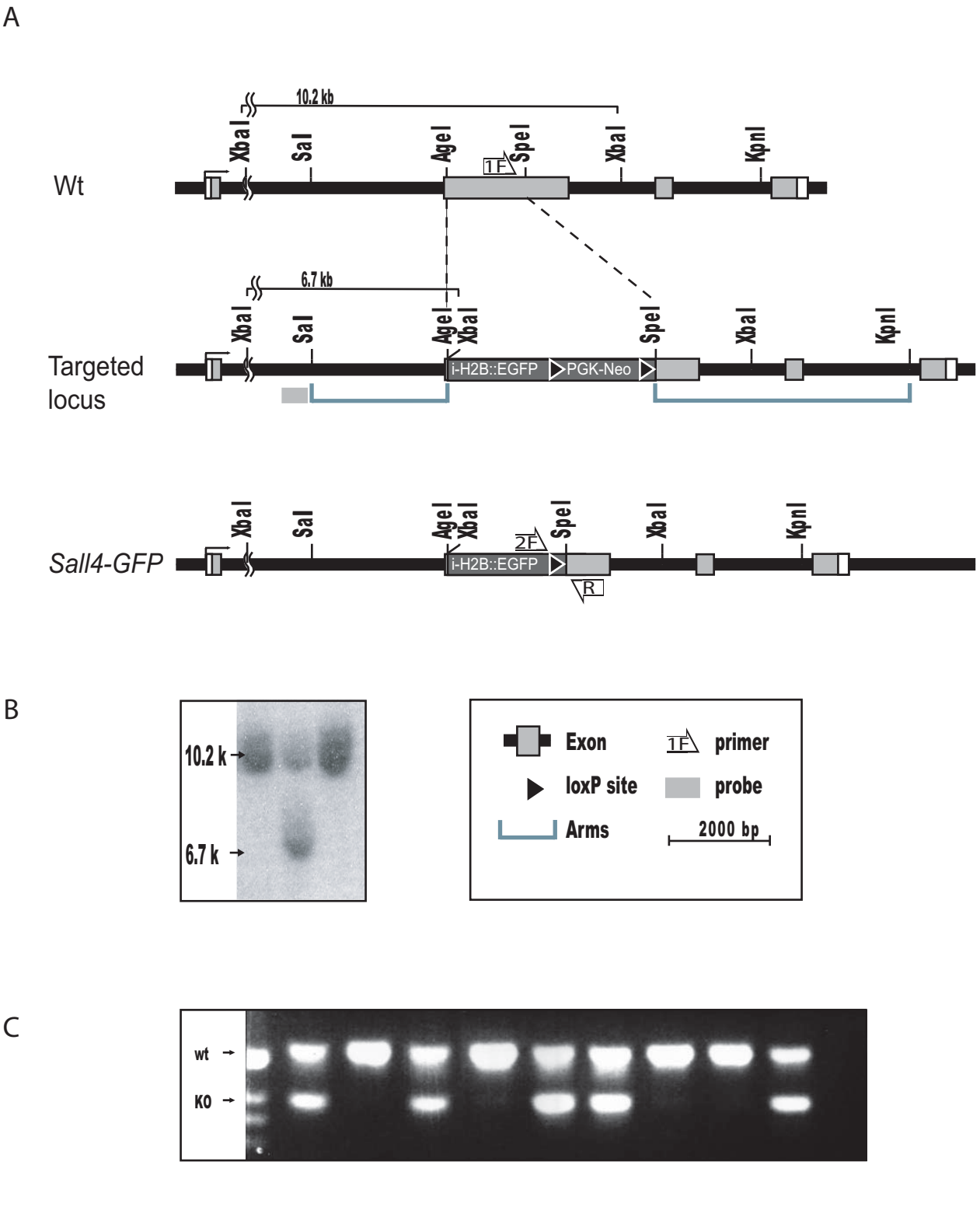


Figure 06: Targeting strategy for *Sall4-GFP*.

A) Exon two carries the coding sequence for the glutamine repeat as well as the first two double zinc finger motifs. A replacement vector carried IRES-GFP as well as a positive selection cassette (PGK-Neo) between the two arms indicated in A. Homologous recombination resulted in genetic removal of exon two and introduces a GFP allele as well as a transcriptional stop codon preventing transcription of downstream exons. Cre-mediated excision of PGK-Neo generated the allele termed *Sall4-GFP*.

B) Screening for homologous recombination was possible by digestion of genomic DNA with XbaI and subsequent southern blotting and probing with a probe 5´external to the 5´arm (grey box).

C) Successful deletion of the Neoomycin resistance cassette as well as further genotyping utilized the primers indicated in A. The 520 basepair band is indicative of the wildtype allele, a 400 basepair band is amplified from Sall4-GFP.

fingers including the one of the H2CH type at the N-terminus. A transcriptional stop signal should prevent transcription of exons three and four, thereby excluding alternative splicing.

Knowing that *Sall4* was expressed very early in development a high resolving nuclear labelling of *Sall4* to label expressing cells seemed important. I therefore designed an IRES-Histone2b::EGFP cassette and inserted it into the *Sall4* knockout vector after testing it in cell culture. This resulted in enhanced green fluorescent protein incorporated in the chromatin of all cells expressing *Sall4.*

Electroporation of embryonic stem cells resulted in 110 colonies that were picked, cultured and analyzed for recombination using an outside probe southern blot strategy shown in figure 6. DNA was digested with XbaI, separated by gel electrophoresis, blotted and the *Sall4* locus was screened with a 5´outside probe. Seven clones that had undergone homologous recombination were detected and Histone2b::EGFP expression assayed. Cells of two independent clones were transferred into wild type blastocysts. High percentile chimeras were taken through the same breeding scheme as *Sall2-lacZ* to remove the positive selection cassette. PCR analysis revealed successful deletion and the engineered allele was called *Sall4-GFP*.

2.5. Conditional Knockout of Sall4

As discussed before, *Sall4* is expressed very early and the expression is largely not overlapping with other members of the *spalt-like* family. This raised the possibility of a very early lethal phenotype in mice homozygous null for *Sall4.* Early phenotypes are often difficult to analyze. It appeared indicated to account for this by the generation of a conditional allele to be able to delete *Sall4* in a time- and tissue-specific fashion (figure 7). In this approach, the second exon of *Sall4* was flanked by loxP sites (*floxed*) to allow its controlled genomic deletion upon expression of Cre recombinase. This would leave only 39 amino acids of coding sequence prior to exon two behind. Possible splicing of exon one into exon three upon deletion of exon two would generate a protein lacking the N-terminal zinc finger required for histone deacetylase binding as well as the glutamine repeat and two of the three double zinc finger motifs suggesting that such a protein would be non-functional. The positive selection marker was surrounded by Flippase Recombination Targets, FRT sites, to allow deletion by flippase (FlpE).

Subsequent to the establishment of 159 stable transfected R1 ESC clones through G418 selection, a southern blot identified six clones that had inserted the neomycin resistance

Targeting Strategy *Sall4-loxP*

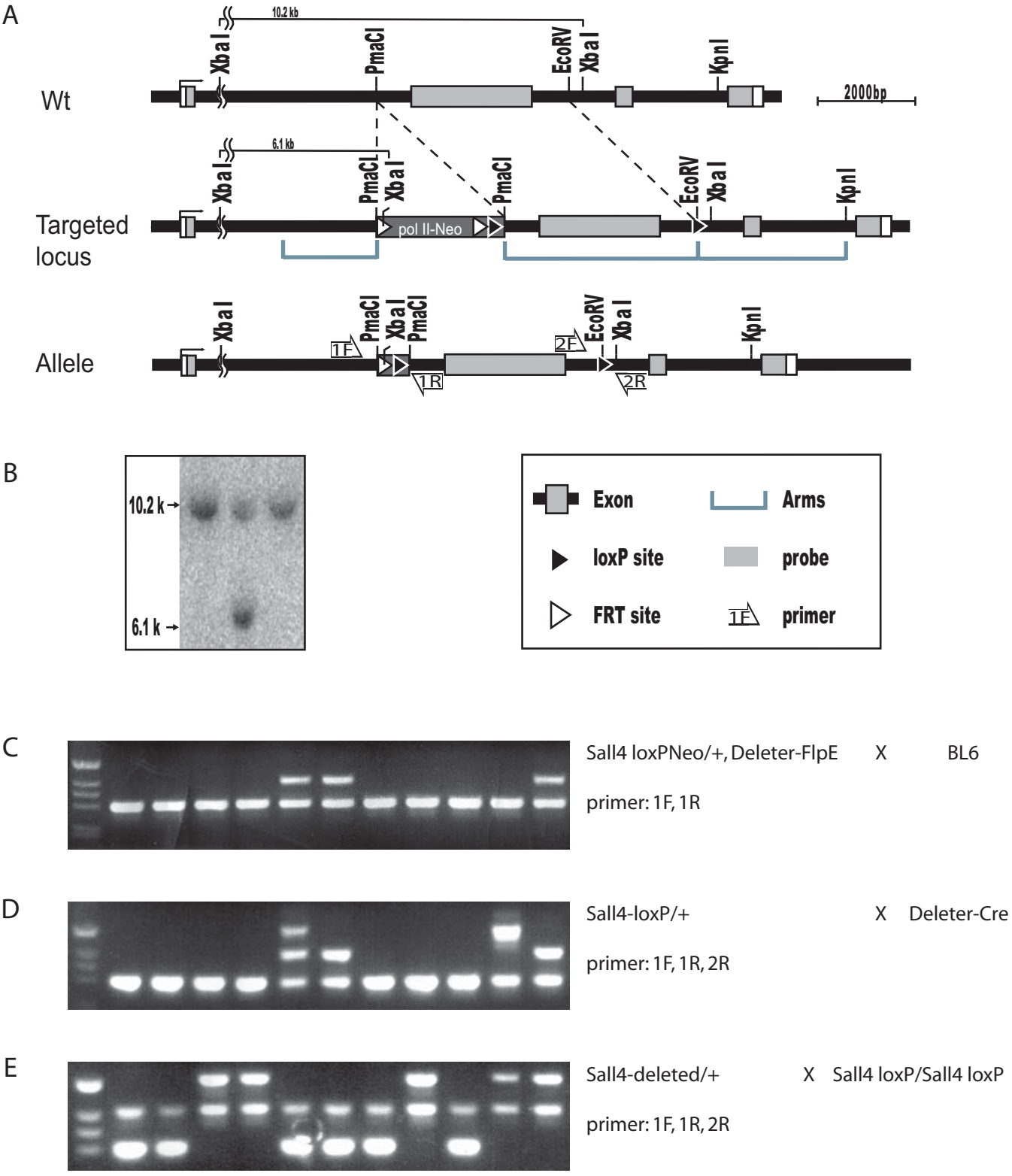


Figure 07: Targeting strategy for *Sall4loxP.*

A) Exon two carries the coding sequence for the glutamine repeat as well as the first two double zinc finger motifs. Floxing the second exon was chosen as a strategy to interfere with Sall4 function . Homologous recombination distal to either loxP site resulted in the insertion of a loxP site plus positive selection marker 5´and a loxP site 3´to the exon. FlpE-mediated excision of PGK-Neo generated the allele termed *Sall4-loxP*.

B) Screening for homologous recombination was possible by digestion of genomic DNA with XbaI and subsequent southern blotting and probing with a probe 5´external to the 5´arm (grey box).

C) Successful deletion of the Neomycin resistance cassette as well as further genotyping utilized the primers 1F and 1R.

D) Deletion of *Sall4* exon 2 can be assayed by primers 1F, 1R and 2R, the highest band at 500 basepairs is indicative of successful deletion.

E) Mice carrying a floxed as well as a deleted allele for *Sall4* are viable showing that *Sall4-loxP* functions as a wildtype allele.

cassette into the *Sall4* locus by means of homologous recombination (see figure 7B). In such a conditional knockout approach recombination 3´ to the neomycin selection cassette could occur either proximal (within exon two) or distal (in the 3´ arm) to the second loxP site. Only a recombination event distal to the loxP site would integrate it into the chromosome, a prerequisite for subsequent Cre mediated recombination. A PCR strategy was used to test the identified clones for the presence of the 3´loxP site. The primers S4loxP2-F and S4loxP2-R anneal to either side of the second loxP site and generated an amplicon of 230 base pairs on wild type DNA or one increased to 270 base pairs by the insertion of additional sequence. The amplicons were analyzed by gel electrophoresis. Of the six clones tested only one (#A12) had recombined distally and thus integrated both loxP sites.

Cells of this clone were injected in blastocysts and taken through the breeding scheme indicated in table 2 using the mouse line hACTP-FlpE expressing FlpE in the germline (Rodriguez et al., 2000). Deletion was assayed by PCR. The presence of a second band in addition to the wildtype band is indicative of excision of the selection marker (figure 7C).

2.6. The Floxed Sall4 Allele is Functional.

In order to be able to use the floxed *Sall4* mouse line, I tested if the inserted loxP sites interfere with the normal function of *Sall4* as well as if the second exon of *Sall4* could be efficiently excised by Cre recombinase. Two breedings provided proof for the genetic functionality of the floxed *Sall4* allele.

First, heterozygous floxed mice were mated to the Deleter-Cre line as it had been done for the other alleles (see table 2). This efficiently deleted exon 2 of *Sall4* from the genome (figure 7D) of mice heterozygous for the generic *Sall4* allele and *Deleter-Cre*. Excision was shown by the presence of a 500 base pair PCR amplicon with primers S4loxP1-F and S4loxP1-R on genomic DNA. It shows that the loxP sites were recognized by Cre recombinase and sequence between was deleted efficiently. This breeding also generated a second null allele similar to *Sall4-GFP*. The new allele was termed *Sall4-deleted*.

Secondly, to address whether the floxed *Sall4* locus was still able to produce *Sall4* protein, mice heterozygous for the floxed allele were intercrossed and genotyped. Offspring homozygous, heterozygous and wild type for *Sall4* was seen in the expected ratio. This argues either that *Sall4* loss of function does not affect viability or that the introduction of loxP sites into intron one and two does not affect transcription and processing of the mRNA. The *Sall4 loxP* mutation was subsequently maintained in a homozygous strain. To validate that *Sall4-*

loxP does not interfere with viability, *Sall4-loxP* homozygous mice were bred to a mouse heterozygous for *Sall4-deleted*. Genotyping pups at weaning age (figure 7E) revealed the expected 50% of animals that carried a null allele over a floxed allele. This serves as a proof that *Sall4-loxP* does not cause a major phenotype and also no other lethal mutations are present in the *Sall4-loxP* strain.

2.7. Breeding and Strain Maintenance

Sall2-lacZ as well as *Sall4-loxP* alleles are homozygous viable and are therefore kept homozygous. *Sall1-lacZ*, *Sall3-lacZ* and *Sall4-GFP* are maintained heterozygous, bred to wildtype mice each generation and offspring is genotyped.

In order to be able to analyze compound knockouts of *Sall1* or *Sall3* together with *Sall2*, these mutations were bred on a *Sall2* null background through two generations. *Sall2* null mice were used as controls for further analysis of the *Sall1* and *Sall3* alleles since loss of *Sall2* shows no phenotype in the genetic and environmental background tested. Crossing of *Sall1-lacZ* animals with *Sall3-lacZ* lines also generated mice double heterozygous for *Sall1* and *Sall3*. These lines were intercrossed both on a *Sall2* wildtype as well as on a *Sall2-lacZ* background.

3. Sall4 is Required for Periimplantation Development

In contrast to the overlapping expression of *Sall* transcription factors in the brain, early expression in the inner cell mass derivatives of extraembryonic endo- and mesoderm is unique to *Sall4*. Thus the effect of a deletion of *Sall4* can not be masked by redundancy with other *Sall* genes. If Sall activity is required for early stages of development, an early phenotype would be expected.

3.1. No Homozygous Embryos were Found at e10.5

To analyze the effects of a loss of *Sall4* on murine development, heterozygous *Sall4-GFP* mice were interbred. Genotyping the offspring revealed that *Sall4* knockout mice were not born alive. Subsequently, the onset of pregnancy was determined by vaginal plug check. To find out whether *Sall4* deficient embryos successfully undergo implantation, gastrulation and turning, 55 embryos were collected 10.5 days post coitum. At this age, genotyping of a normally developed embryo can be achieved by PCR analysis. Embryonic tissue was carefully separated from maternal tissue and DNA was purified. Wildtype and heterozygous embryos were found at the expected ratio but no embryos homozygous for *Sall4-GFP* could be detected. Consistent with this, a high number of small, empty implantation sites were observed. Genotyping of absorbed embryos or empty implantation sites is not possible by PCR due to the mixing of maternal, heterozygous tissue with the embryonic tissue of interest. Thus, amplicons indicative for both alleles would always be detected. It was therefore speculated that these empty implantation sites were due to embryos homozygous for *Sall4-GFP*. To support this assumption, embryos were collected at e3.5, prior to implantation into the uterus, and genotyped by PCR. Roughly one quarter of the blastocysts collected from an interbreeding of *Sall4-GFP* heterozygous animals was genotyped to be *Sall4* deficient. It was conducted that termination of development is occurring later than e3.5 and significantly before e10.5.

3.2. Generation and Testing of an Antibody Raised Against Murine Sall4

The presence of empty implantation sites suggests that *Sall4* deficiency does indeed lead to early embryonic lethality as it was hypothesized from comparative expression studies.

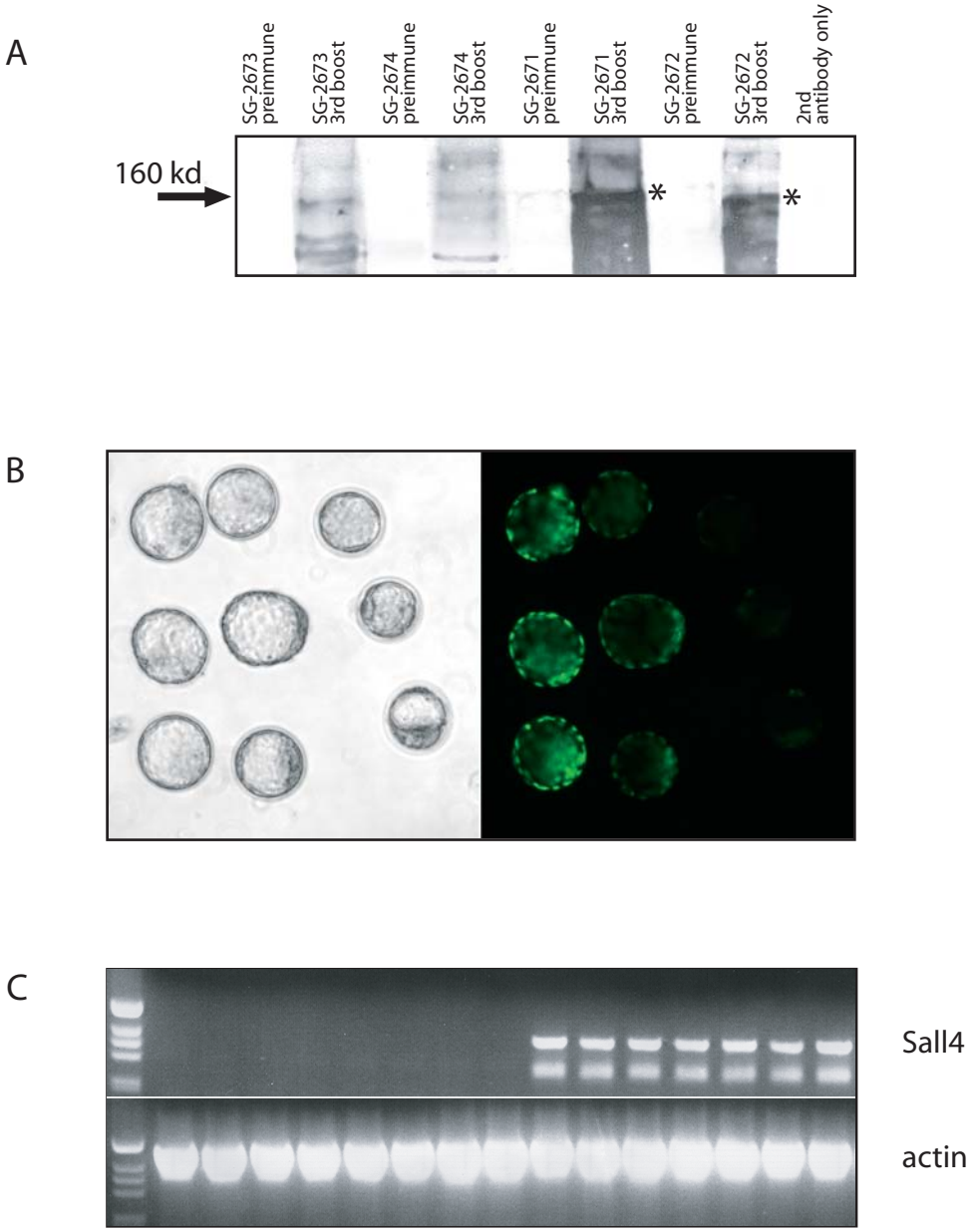


Figure 08: Testing of peptide antibodies generated against Sall4; Quantitanive differences in GFP intensity of e3.5 blastocysts enable separation of genotypes

A) Testing of sera obtained from SG-2671 till SG-2674. Equal amounts of protein extract from ES cells was loaded on an acrylamide gel, run, blotted and each lane was incubated independently in preimmune serum and third boost serum (1:1000). Animals SG-2673 and SG-2674 were injected with the C-terminal peptide and do not show immunoreactivity. SG-2671 and SG-2672, immunized against the N-terminal peptide show robust immunoreactivity against Sall4.

B) Nine blastocysts in nomarski optics (left) and GFP bioluminescence (right). Blastocysts are sorted according to GFP intensity into strong (left column), medium (center column) and no GFP signal (right column).

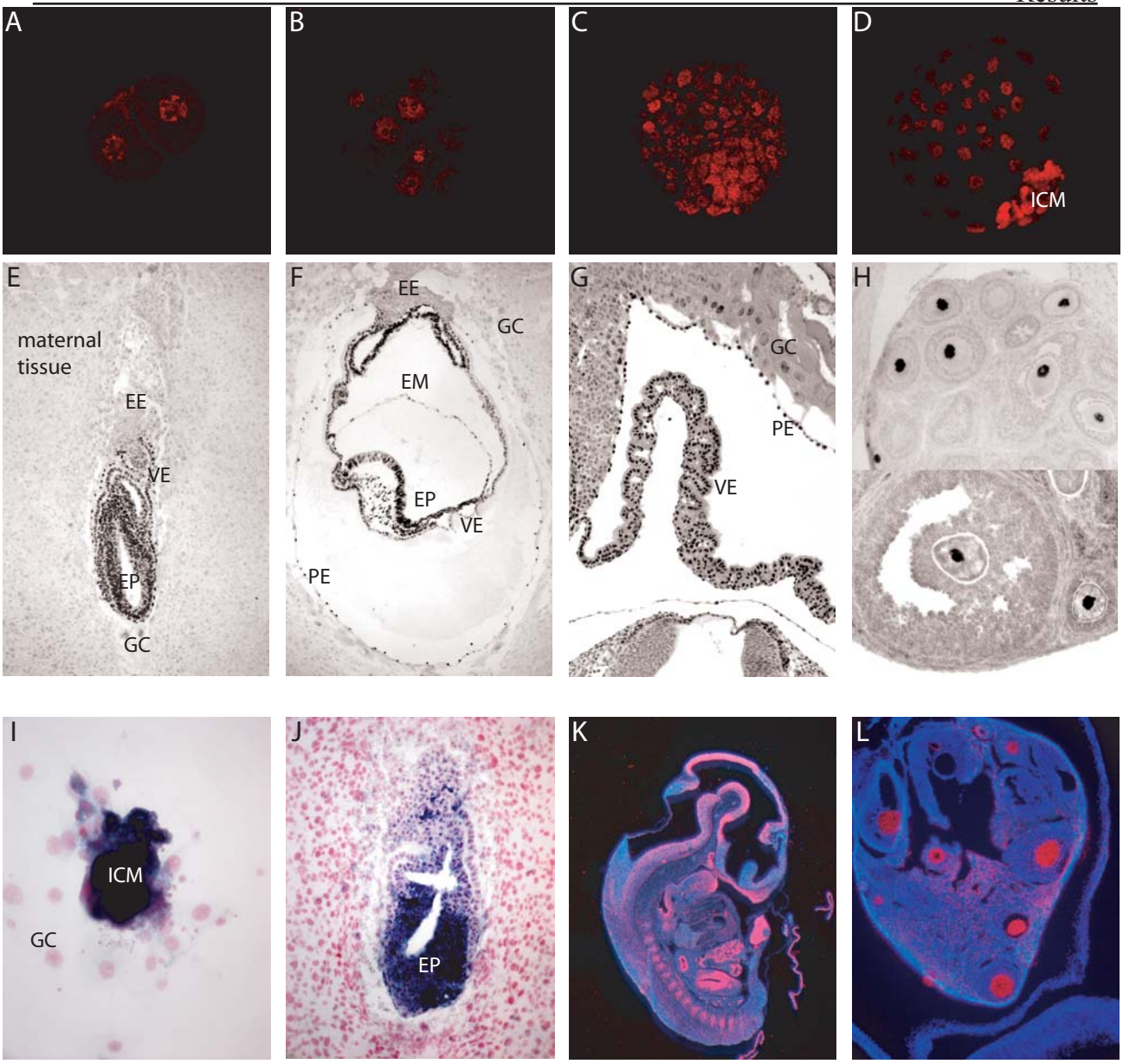
C) Subsequent to sorting, RNA was purified from 15 individual blastocysts and analyzed for presence of Sall4 transcript. Equal amounts of RNA were successfully purified, reverse transcribed and can be used for PCR as seen by the strong band obtained in all samples with primers specific for actin RNA. The eight blastocysts quantified to show strong GFP expression did indeed not contain Sall4 mRNA;

More detailed analysis of *Sall4* protein distribution was required to interpret the observed defects. GFP expression in the knockin of H2b::EGFP into the *Sall4* locus gives a good indication of *Sall4* expression in heterozygous animals but protein stability and translational control of the mRNA transcript are not addressed. High levels of *Sall4* expression had been observed by *in situ* hybridization in developing oocytes. This opened possibility that maternal contribution of protein could affect development in a *Sall4* knockout blastocyst. To address Sall4 protein presence, an antibody recognizing Sall4 was required.

The decision to generate antibodies against mouse *Sall4* was taken. Peptides identical to the N- and C-terminus of *Sall4* were synthesized commercially and used to immunize two rabbits each. Affinity of antibodies present in the serum was analyzed by western blots after the third bleed. While the C-terminal peptide did not result in an antibody recognizing murine *Sall4*, both rabbits immunized with the N-terminal peptide show robust immunoreactivity against full length endogenous protein (figure 8A). Antibodies were affinity purified after the 5^{th} boost and tested in immunohistochemistry. Optimal dilution was determined to be 1:1000 in immunofluorescence and immunohistochemistry as well as 1:5000 for western blotting.

3.3. Precise Expression Analysis for Sall4

It had already been established by *in situ* hybridization that *Sall4* transcript is present in mature oocytes and e6.5 implantation sites until mid gestation. Protein expression studies thus spanned all embryonic stages until e10.5 (figure 9). Protein could be visualized from cleavage stage embryos continuously until e8.5 embryos where levels start to decrease. In oocytes as well as in all other cell types examined, *Sall4* is localized in the nucleus. High magnification of fluorescent immunostainings shows localization in distinct spots within the nucleus. Expression can still be detected at two cell stage and is likely due to maternal contribution of protein and mRNA (figure 9A). In cleavage stage embryos, expression of *Sall4* is again seen at 16 cell stage and can subsequently be visualized in all blastomeres. Expression of a paternal *Sall4-GFP* allele also starts at morula stage, e2.5, in all blastomeres, arguing for an onset of zygotic *Sall4* expression at the 16 cell stage. Blastocysts at e3.5 start to show higher level of expression in the inner cell mass while the trophoblast still expresses considerable amounts of *Sall4*. After expanding blastocysts in culture for 24 hours, *Sall4* is expressed stronger in the inner cell mass stem cells than in the surrounding trophectoderm but expression is still visible in all cells.



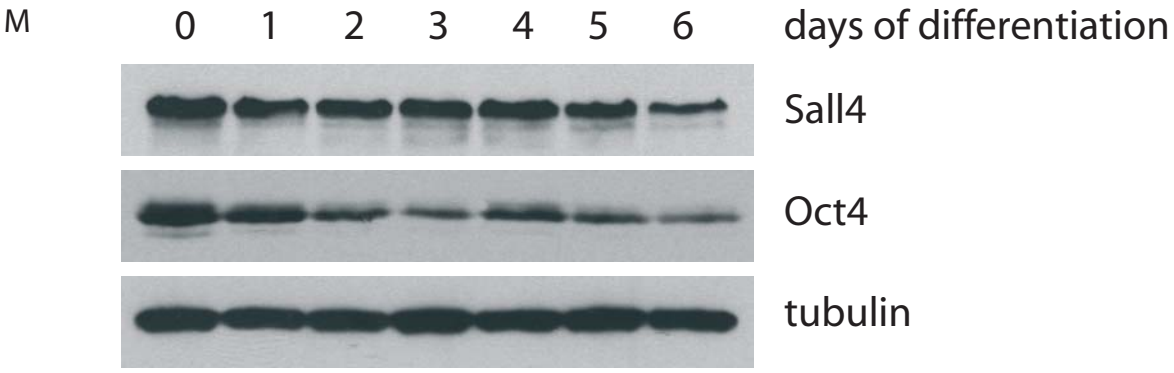


Figure 09: Sall4 expression at periimplantation stages

Analysis of Sall4 expression by immunofluorescence (A-D), immunohistochemistry (E-H) and in situ hybridization (I-L). Maternal Sall4 protein is still detectable at the two cell stage (A) prior to zygotic transcription. The first zygotic transcription is visible between 8 and 16 cell stage (B), will be ubiquitously expressed in early blastocysts (C) and localize mainly to the inner cell mass in expanded blastocysts (D). At early postimplantation stages e6.5 (E), e7.5 (F), specific expression in the inner cell mass (ICM) including embryo proper (EP), extraembryonic mesoderm (EM) and exrtaembryonic encoderm (visceral: VE, parietal:PE) reaches completion (GC: giant cells of the trophoblast; EE: extraembryonic ectoderm). Data obtained by antibody staining were confirmed using in situ hybridization. Sall4 transcript is exclusively found in the inner cell mass of a blastocyst ouwtgrowth in culture (I), an embryo e6.5 (J) and e11.5 (K). H and L show expression in maturing oocytes at P3 (H, top) and eight weeks of age (H, bottom and L);

M) differentiation of wildtype ESCs by withdrawl of LIF and addition of retinoic acid results in a decrease of Sall4 and Oct4 present. Tubulin was used as a loading control.

Post implantation at e6.5, expression is seen in embryo proper and primitive endoderm exclusively. All these tissues are ICM derivates. The primitive or extraembryonic endoderm surrounds the epiblast cells and forms an epithelium all around the blastocoel separating apolar epiblast cells from the lumen of the blastula. Sall4 protein abundance is similar in all ICM cell types. At e7.5, all three embryonic germlayers can easily be recognized to express *Sall4*. Extraembryonically, Sall4 immunoreactivity is visible in allantois and chorion of the extraembryonic mesoderm as well as in visceral and parietal primitive endoderm.

Cells of the trophectoderm lineage comprise two cell types, polar (extraembryonic ectoderm, EE) and mural trophectoderm (giant cells, GC). In the presence of FGF4 (fibroblast growth factor 4) secreted from the ICM, trophoblast cells remain proliferative as polar trophectoderm. Without the stimulus of this growth factor trophoblast cells will undergo endoreplication and form postmitotic, polyploid giant cells (Tanaka et al., 1998). Both cell types do not express *Sall4* protein at detectable levels. The expression in the embryo is regulated down upon onset of differentiation at e8.5 but expression is still detected in embryo proper, extraembryonic mesoderm as well as in visceral and parietal primitive endoderm. At this stage, all membranes surrounding the embryo are fully developed and show strong immunoreactivity while the developing placenta still shows only very faint expression. In figure 9 I-L *in situ* hybridization is visualizing transcripts enzymatically (I, J) or radiographically (K, L). Expression of transcript and protein fully overlaps and confirms the data obtained from immunostaining. In figure 9I, a blastocyst outgrowth (BO) is shown expressing *Sall4* in the ICM exclusively. A BO is a blastocyst attached and grown on a surface for five days (see also below).

Sall paralogous proteins show strong correlations of expression of in stem cell pools of ventricular zones of the brain and in early embryos. I next wanted to know if *Sall4* expression is linked to an undifferentiated state of pluripotency. If such a connection can be made, expression levels of *Sall4* should drop quantitatively upon induction of differentiation in ES cell culture. The transcription factors nanog and Oct4 had previously been shown to be required for stem cell identity. Protein abundance of these stem cell factors is known to drop in differentiation assays of embryonic stem cells (Boyer et al., 2005).

ES cells were plated at very low density on feeder free 10cm culture dishes in ESC medium and kept in culture for seven days. Medium was exchanged every day. Each day, colonies of one more plate were deprived of ESC medium and medium was replaced by differentiation medium containing all-trans retinoic acid at 100 nMol/liter. Removal of LIF from ES cell medium together with the addition of retinoic acid induces differentiation of

ESCs (Gribnau et al., 2005). Protein levels of Sall4 and Oct4 were analyzed by western blot. Equal protein loading was assured by reprobing the membrane against tubulin.

Figure 9M shows strong immunoreactivity against Sall4 and Oct4 protein in an ES cell lysate. Both protein levels decrease upon LIF withdrawal to a similar extend as well as with comparable kinetics. Reduction of *Sall4* expression could thus be linked to differentiation in embryonic stem cells. Whether loss of *Sall4* is sufficient to inhibit ESC self renewal was then analyzed *in vivo* as well as ex vivo utilizing the previously generated *Sall4-GFP* allele.

3.4. In Vivo Analysis of Embryos Deleted of Sall4

Sorting Embryos at Blastocyst Stage

Since no *Sall4-GFP* homozygous embryos could be found at e10.5 and initial analysis revealed presence of blastocysts homozygous for *Sall4-GFP,* I wanted to establish their developmental potential prior to e10.5. Heterozygous *Sall4-GFP* animals were interbred and cleavage stage embryos, morulas and blastulas of superovulated females were collected at e 0.5 to e3.5 by flushing oviduct and uteri. Presence of *Sall4* knockout embryos was established by different methods.

During the handling of embryos it was observed that three distinct classes of GFP intensity are present in blastulas. While 50% of the observed blastulas repeatedly showed weak signal, 25% showed no fluorescence and another quarter appeared bright green. Intensity of GFP signal thus seems to depend on the number of *Sall4-GFP* alleles inherited. The observed ratio is in accordance with expected Mendelian allele distribution in an interbreeding of animals heterozygous for *Sall4-GFP*.

To verify that *Sall4* knockout animals were present and developed comparably to wildtype until e3.5, blastocysts were genotyped with a PCR reaction using genomic DNA of blastocysts as template. Indeed, knockout animals were still present. Subsequently, I sorted embryos according to the intensity of the detected GFP signal and prepared RNA from single blastocysts. Reverse transcription and PCR showed that in embryos with strong GFP-fluorescence, no *Sall4* transcript can be amplified as presented in figure 8C. Actin mRNA could be detected at equally high levels in all embryos serving as control for successful RNA preparation and *in vitro* reverse transcription on single blastocysts. I therefore concluded that loss of *Sall4* does not affect development prior to implantation after e3.5. In addition this result clearly proofs, that no maternal mRNA is contributing to Sall4 protein in blastocysts homozygous for *Sall4-GFP*.

Secondly, immunostaining against Sall4 revealed that in one fourth of the embryos no Sall4 protein could be detected. While *Sall4* protein is detected in 16 cell wildtype littermates, no protein can be detected in *Sall4-GFP* embryos at this stage (figure 10 A-D) and thereafter. At e3.5 when GFP signal allows reliable sorting of animals, Sall4 protein was detected in all wildtype and heterozygous embryos while no *Sall4-GFP* homozygous animals show detectable staining (figure 10E-H).

Sorting embryos by fluorescence intensity prior to genotyping revealed that the GFP signal is indicative of the genotype. For further studies, from then on blastocysts were routinely sorted as described. No inconsistencies in the phenotype were observed in the analysis. Thus sorting embryos based on GFP intensity is a precise method to identify *Sall4* knockouts.

No Sall4 protein can be detected at 16 cell stage in a knockout. A phenotype due to the loss of Sall4 should thus manifest until the expanded blastocyst stage if Sall4 is required at preimplantation development. I asked whether these embryos show any effect due to the absence of Sall4 in expanded blastocyst stage by immunofluorescent staining.

Immunohistochemical Analysis of Expanded Blastocysts.

Blastocysts were collected at e3.5, sorted and kept in culture for 24 hours. In this period they will undergo one more round of cell division resulting in a total of around 128 cells and also hatch from the protective zona pellucida. Further internalization of fluid by the trophoblast results in an increased volume of blastocoel surrounded by a thin layer of mural trophectoderm. These blastocysts roughly resemble e4.0 blastocysts at the time of implantation *in utero*.

Three cell lineages can be found in an e3.5 blastocyst, trophoblast and the ICM containing embryonic stem cells or ES cells as well as primitive (extraembryonic) endoderm. Markers indicative of the different cell types present at this stage were used to screen for phenotypic aberrations in embryos depleted of Sall4. Counterstaining with DAPI revealed the morphology of blastocysts. This method also shows that overall cell number of expanded blastocysts does not differ from wildtype.

Oct4 protein is expressed in all cells of the early blastocyst but restricts to the inner cell mass after 24 hours of culture (Nichols et al., 1998). Staining Blastocysts with Oct4 antibody revealed no differences between *Sall4* knockout and control. This suggests that loss of *Sall4* does neither prevent inner cell mass specification nor *Oct4* transcription. Active translation of Oct4 also indicated viability of inner cell mass cells.

As opposed to Oct4, Nanog is expressed only in the epiblast but excluded from the primitive endoderm in the inner cell mass (Mitsui et al., 2003; Yamaguchi et al., 2005). Immunostaining *Sall4* negative blastocysts for Nanog expression also revealed no differences.

The trophoblast lineage was successfully established in *Sall4* knockout animals as seen by the morphology of the blastocyst (figure 8B). Only if cells form an epithelial sheet surrounding the inner cell mass, polarize successfully and form tight junctions can they pump in fluid to form the blastocoelic cavity. In accordance to this, TROMA, an antibody recognizing cytoskeleton of the trophoblast lineage (Nichols et al., 1998), labels the trophoblast also of *Sall4* deficient blastocysts without recognizable differences to wildtype blastocysts. In the inner cell mass, no TROMA epitopes can be recognized.

Another molecule expressed in the trophoblast is the homeobox transcription factor Cdx2 (Strumpf et al., 2005). It is the first known marker for the trophoblast lineage in mouse and starts to be differentially expressed already in compacting morulas. Immunostaining *Sall4*-deleted expanded blastocysts with an antibody against Cdx2 revealed that - like in the wildtype - Cdx2 is localized to trophoblastic nuclei and excluded from the ICM. Overall expression levels observed in the knockout were reproducibly higher than in wildtype littermates.

In conclusion, blastocysts deleted of *Sall4* do not show gross phenotypic effects, are seen at the expected ratio and are not developmentally delayed. All cell lineages are present based on the analysis of molecular markers suggesting that *Sall4* is not required for the establishment of early lineages in the mouse blastocyst.

One possibility is that maternal protein is still present until blastocyst stage and might prevent an earlier onset of a *Sall*4 deletion phenotype. To answer that question, a conditional knockout of *Sall4* specifically in maturing oocytes was generated using a Gdf9-Cre line (Lan et al., 2004). Oocytes from these females do not accumulate Sall4 protein. I intended to compare development of blastocysts from such a breeding to blastocysts from a GFP interbreeding. However, females with Sall4 deleted oocytes were completely sterile. Even after hormonal stimulation, plugged females had not ovulated. Sectioning ovaries revealed the absence of preanthral follicles showing that *Sall4* deleted oocytes fail to mature underlining the importance of *Sall* genes for development.

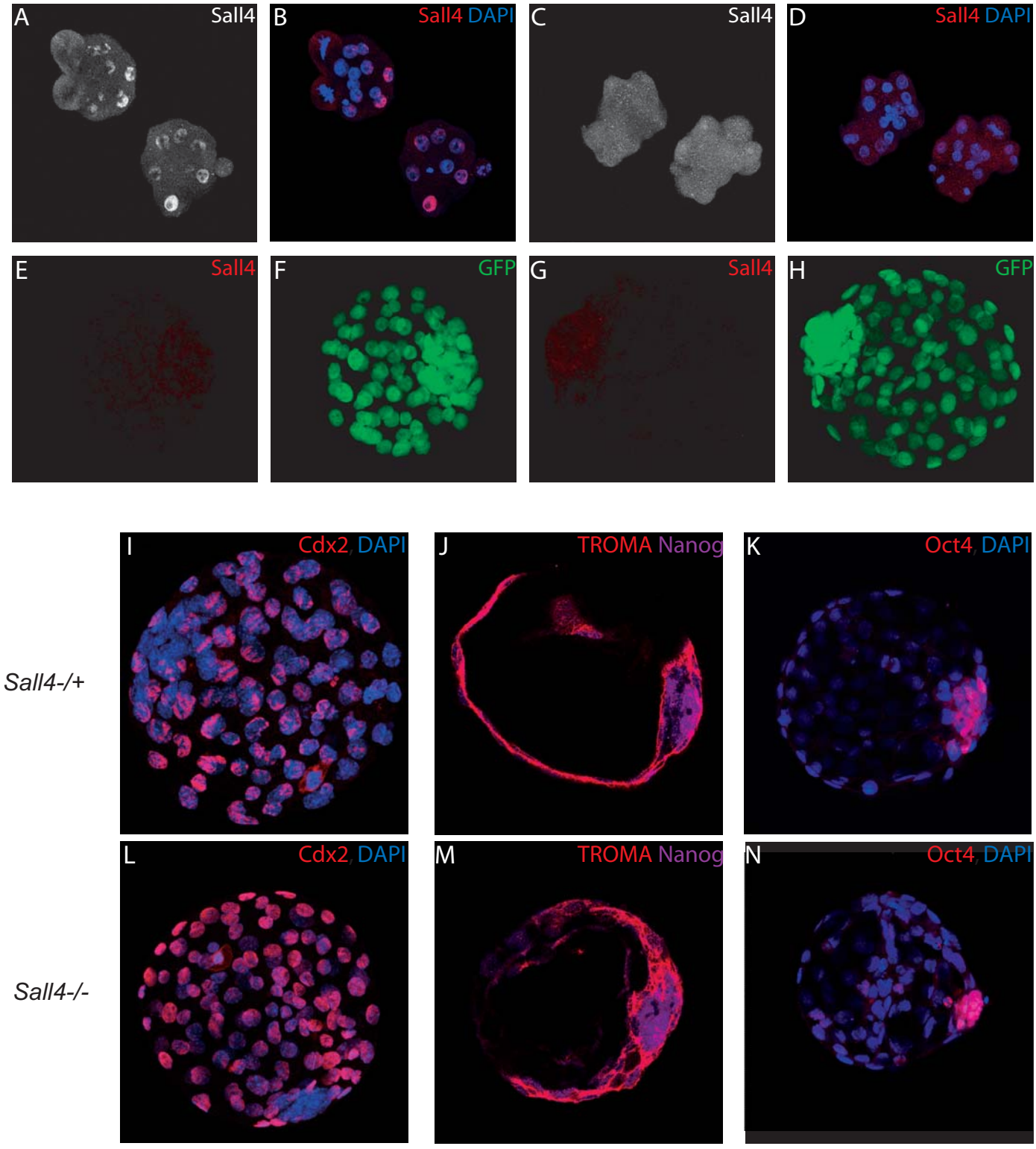


Figure 10: Analysis of expanded blastocysts deleted of *Sall4*.

A-H) No Sall4 protein can be detected in cleavge stage till blastcyst stage *Sall4* knockout embryos. While Sall4 protein can be detected in wildtype blastomeres (A, B), no signal is detected in immunostainings with an anti-Sall4 antibody on morulas or blastulas at 16 cell stage (C, D), early blastula (E, F), or expanded blastula (G, H). Nevertheless, blastocysts reach full size at e3.5 plus 24 hours in culture and show no delay compared to littermates.

I-N) Comparative immunostainings on littermates wildtype or deleted of *Sall4*. Blastocysts G-N were collected at e3.5 and kept in culture for 24 hours to allow full development to occur. No significant change of expression domains was detected for Cdx2, TROMA, Nanog or Oct4. In both genotypes, Cdx2 and TROMA positive cells are observed in the trophoblast while the inner cell mass labels for Nanog and Oct4. Expression levels of Cdx2 in the trophoblast were always ehanced in the knockout (L).

Histological Analysis at e6.5 and e7.5

Determining the developmental potential post implantation is complicated by the fact that even wildtype embryos can not easily be genotyped prior to e 7.5. It was known from the analysis at e10.5 that no embryos form in the knockout and thus no tissue could be retrieved for genotyping. To overcome this problem and to show that *Sall4* knockouts are absent post implantation rather than the negative result of a failure to identify knockout animals, blastocysts were collected at e3.5 from superovulated females. After sorting the embryos according to their GFP intensity, batches of 7-15 blastocysts for each genotype were transferred into separate pseudopregnant females 2.5 days post coitum. Litters were harvested at e6.5 and e7.5 and processed for paraffin sectioning. Periodic Acid-Schiff Reagent (PAS) staining several implantation sites for each genotype revealed a dramatic difference between knockout and control litters. While in control transfers all or almost all implantation sites carried an embryo of expected morphology, no such embryo was seen in any of the knockout litters analyzed. Instead, at e6.5 the implantation site was empty containing only scattered cells. Close Examination of the edges of the implantation site cavities revealed the presence of few very big cells with huge nuclei, characteristic of trophoblast giant cells (figure 11A, E: asterisk). This suggested that knockout blastocysts hatched successfully from the zona pellucida, implanted in the uterine wall and embryonic giant cells responded to the implantation adequately. At e7.5, hardly any embryonic cells could be identified from implanted knockout embryos.

Lineage Marker Analysis on Implantation Sites

To determine which embryonic cell types are still present in e6.5 implantation sites, *in situ* hybridization was used. *H19* is a non-coding RNA expressed at high levels in all embryonic (as opposed to maternal) cell types but the embryo proper (Avilion et al., 2003; Mitsui et al., 2003). Figure 11C illustrates *H19* expression in trophoblast and extraembryonic endoderm in the wildtype control. In implantation sites of *Sall4* knockout animals, *H19* staining is still seen surrounding the cavity. This is presumable due to the presence of primary giant cells and serves as a further proof that implantation occurred.

Oct4 expression is mutually exclusive with *H19* and found in all cells of the embryo proper at this stage (Avilion et al., 2003; Mitsui et al., 2003). *In situ* hybridization for *Oct4* revealed that no cells of embryo proper identity are present in the case of a *Sall4* knockout at e6.5 (F).

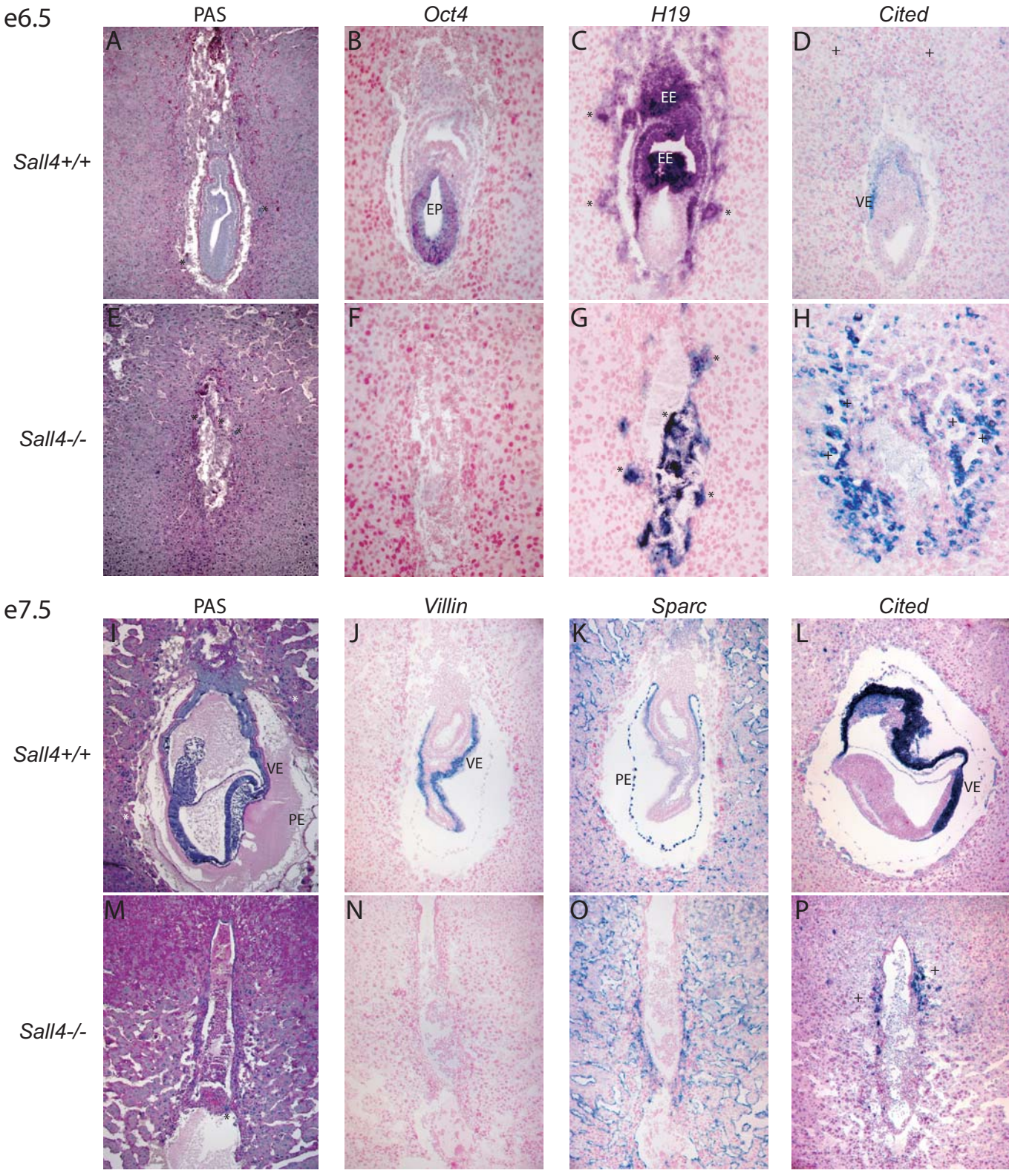


Figure 11: No inner cell mass (ICM) derivatives can be identified in *Sall4* knockout embryos post implantation.

PAS staining (A, E, I, M) on reimplanted, pregenotyped embryos wildtype and deleted of *Sall4* shows, that implantaion sites appear empty in foster mothers pregnant with *Sall4-/-* embryos. While residual cells (giant cells: asterisk) can be seen at e 6.5 (E), embryonic cells are rarely observed in e7.5 implantation sites (M). Reimplanted control embryos grow normally arguing that the sorting procedure does not disturb the development of the embryos.

In situ hybridization was used to label individual cell lineages expected at e6.5 and 7.5 to test which cell lines are affected in the knockout. No cells of the embryo proper (EP) expressing *Oct4* were observed in the knockouts (F) while strong *H19* staining (G) is visible in giant cells (F); *Cited* is expressed in secondary giant cells (+) and cells of the visceral primitive endoderm (VE) at e6.5 (D,H) and e7.5 (L, P). An overabundancy of secondary giant cells in knockouts is observed. Neither visceral (VE) nor parietal (PE) endoderm is present in *Sall4* deleted implantation sites at e7.5 as shown with in situ probes against *Villin*, *Sparc* and *Cited*. extraembryonic ectoderm (EE);

To distinguish between the two *H19* expressing cell lineages, trophoblast and primitive endoderm, I performed *in situ* hybridization with markers indicative of primitive endoderm. While *Hnf-4* (Taraviras et al., 1994) and *Villin* (Maunoury et al., 1988) are expressed in the visceral primitive endoderm located on the surface of the ICM cone, *Sparc* is expressed in the parietal endoderm giving rise to Reichert´s membrane (Mason et al., 1986). None of these markers could be detected in implantation sites at e7.5. It was thus concluded that all the staining observed with *H19* originates from cells of the trophectoderm lineage. In the implantation sites analyzed, signal could not only be detected in the surrounding primary giant cells but also in the center of the implantation site. *Cited* is expressed by cells of the visceral primitive endoderm as well as by secondary giant cells (Rodriguez et al., 2004). In a knockout for *Sall4,* expression of *Cited* can be detected in multiple cells around the implantation site. If primitive endoderm is absent in *Sall4* knockouts, expression must be due to secondary giant cells. The overabundance of *Cited* positive cells may be explained by differentiation of polar trophectoderm. The trophoblast differentiates into polyploid giant cells in the absence of Fgf4 secreted from the ICM (Tanaka et al., 1998). If indeed no primitive endoderm or embryo supplying Fgf4 can be found in knockout implantation sites it would be expected that the staining for *H19* observed in the center originates from giant cells.

As seen in PAS staining, giant cells are present in the center of a *Sall4* knockout implantation site while they are never found in such location in control animals. This analysis demonstrated that, *in vivo*, loss of *Sall4* leads to periimplantational lethality in mice. It also establishes that only trophoblast cells can be found at implantation sites of *Sall4* knockout embryos while all cell types originating from the inner cell mass are missing. To study the fate of ICM cells, I used an *ex vivo* system allowing the developmental observation of *Sall4* knockout embryos.

3.5. Ex vivo Analysis

Outgrowth of Blastocysts in Culture

In utero, blastocysts undergo implantation at around e4.0 unto e4.5. From then on up until e5.5 the blastocyst will hardly grow in size making identification hard. To circumvent this problem, blastocysts were harvested at e3.5 and cultured in ESC medium for up to ten days. In such conditions, blastocysts hatch from the surrounding zona pellucida and adhere to the culture dish at around day two in culture. Given the planar surface of the dish, the three dimensional structure of the developing embryo is altered. The trophectoderm flattens while

ICM cells grow on top in a shape similar to the *in vivo* situation. These cells might subsequently form embryoid bodies consisting of a primitive endoderm and an embryonic cell layer as well as a proamnion like cavity comparable to e6.5 (Wiley and Pedersen, 1977). Gastrulation does not occur. In other cases, the inner cell mass grows in a more undifferentiated fashion.

Blastocysts of an interbreeding of *Sall4-GFP* heterozygous animals were sorted at e3.5 and 24 of each of the three genotypic classes were placed in separate wells of a 24 well dish containing ES cell medium. Starting with day two in culture, the development was monitored daily. Figure 12A shows examples of such blastocyst outgrowth (BO) experiments. Control embryos behaved as expected and their ICM grew constantly. In contrast, ICMs of knockout blastocysts did not gain size after approximately four to six days. No embryoid bodies or large aggregations of cells were ever observed. Instead, the inner cell mass terminated growth, disintegrated and spherical multi-cellular aggregates detached from the underlying trophectoderm (asterisk). At this stage, the trophoblast has already undergone multiple rounds of endoreplication and formed the typical primary giant cells as judged by histological examination. The detaching aggregates of ICM derivative cells have lost the typical ESC appearance of small, round cells. Instead, cells appeared differentiated.

This suggested that the cells of the trophoblast were not altered in their developmental potential *ex vivo* while the ICM cells lost the ability of self renewal and proliferation leading to a complete loss of the ICM. The observed effect of *Sall4* deletion in blastocyst outgrowth supports the *in vivo* finding and suggests that *Sall4* is required specifically in the ICM to maintain the identity or a pluripotent ES cell.

Cell Types Present in Blastocyst Outgrowths

Observation of blastocyst outgrowths suggested that the identities of ICM stem cells may not be maintained in the absence of *Sall4*. To prove this, markers for the different cell lines present in the blastocyst were used. BOs were fixed and processed for immunofluorescent staining or *in situ* hybridization.

Apoptotic and necrotic cells can be identified in cell culture by fragmented nuclei in DAPI staining and large vacuoles. Judged by appearance in light microscopy, no cell death could be observed. To validate that apoptosis does indeed not account for the size reduction in the knockout outgrowth, apoptotic cells were identified with TUNEL staining. No differences were observed as presented in figure 12B.

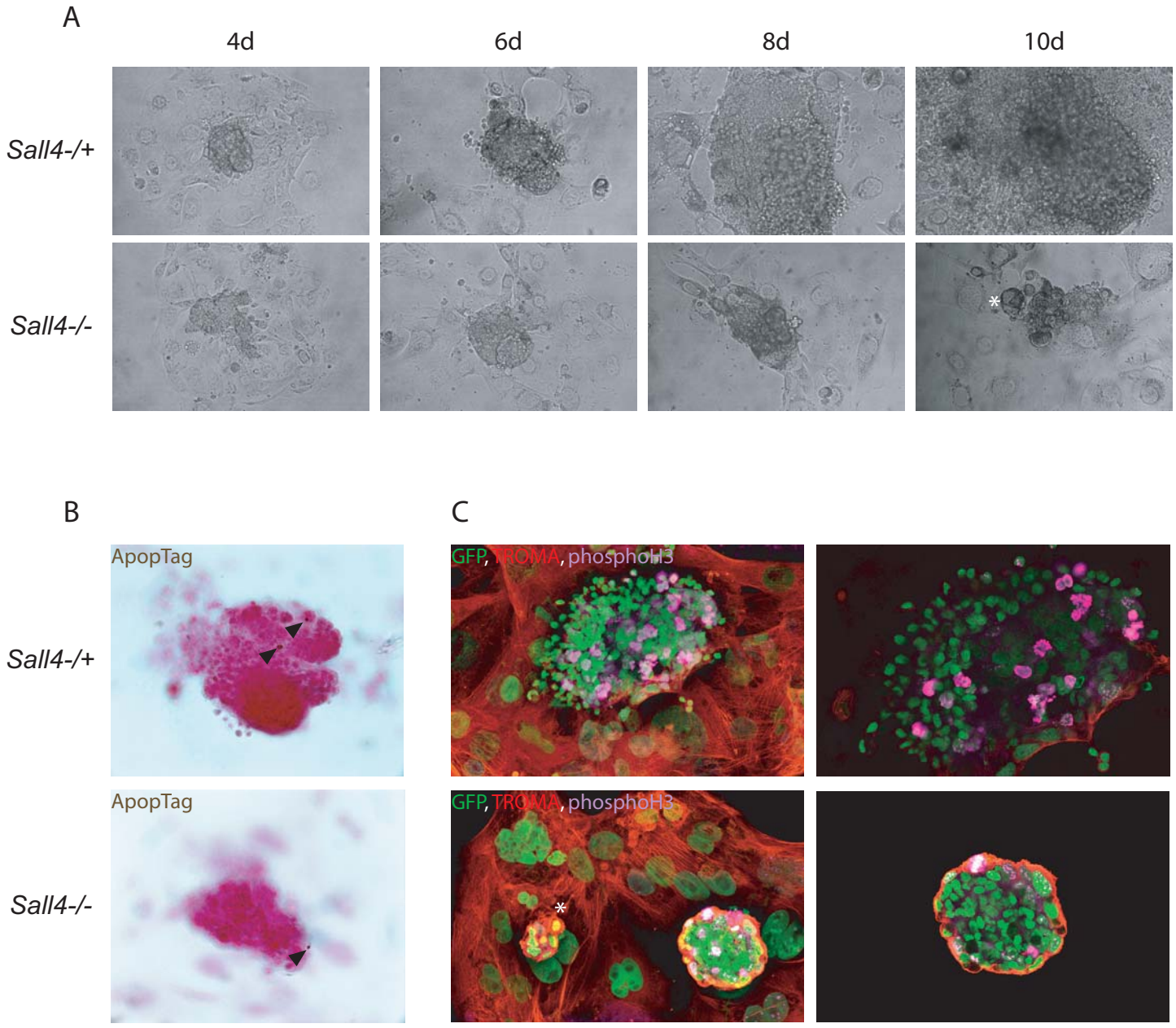


Figure 12: Sall4 is required for the ouwtgrowth of blastocysts in culture.

To study the dynamics of implantation and to follow the progress of the phenotype due to loss of Sall4, blastocysts were collected e3.5 and cultivated in ESC meduim containing LIF. Under these conditions, the trophoblast will adhere to the surface of the culture dish and the inner cell mass will grow on top. Progress of blastocyst outgrowth (BO) was documented daily, representative examples are shown in A. No significant differences between wildtype and knockout are observed until day four in culture. Subsequently, *Sall4-/-* BOs will terminate growth and the ICM disassembles and fractionates until only giant cells remain after 7-12 days.

B) Apoptag staining at day 6 indicates that the cell loss observed is not due to apoptosis since only very few cells stain positive both in wildtype and knockout (arrowhead).

C) Confocal immunofluorescence at day 8 illustrates observed changes. Pictures to the right are higher magnifications of a single optical section through the ICM. Strong TROMA staining recognizing intermediary filaments in epithelial cells around the mutant ICM suggests that ESCs differentiate in ESC culture medium. Detouching aggregates observed in A appear to be trophoblastic cells (asterisk in A, C). Cells undergoing division are marked by immunostaining against phosphorylated histone 3 and are present in wildtype and knockout BOs. In knockouts however, cell division is mainly seen in TROMA positive cells while almost no division is visible in GFP positive, TROMA negative cells.

To visualize trophectoderm, outgrowths were stained for TROMA immunoreactivity as seen in figure 12C, red channel. Giant cells stained strongly in knockout and control as expected but differences were seen in the inner cell mass. In wildtype, hardly any TROMA immunoreactivity is observed in and around the ICM since trophoblast cells settle on the surface of the slide. In contrast, cells surrounding the mutant ICM stain strongly for TROMA. In the overview of a knockout outgrowth in figure 12C one of the multicellular, spherical aggregates can be seen displaying TROMA immunoreactivity (asterisk). This suggests that cells at the surface of the inner cell mass are no longer ES cells since TROMA staining clearly shows presence of a cytoskeleton typical for epithelial cells.

Purple staining in figure 12C is indicative of dividing cells as visualized by an antibody against phosphorylated histone 3. It shows that ICM cells in both wild type and knockout BOs are still dividing. Nevertheless, the rate of proliferation is clearly reduced in knockout ICMs, especially amongst the TROMA negative cells. This result suggests that detaching cellular aggregates immunoreactive for TROMA together with reduced proliferation are the reason for a loss of cells at the site of the ICM.

Taken together, the analysis of cell lineage markers in blastocyst cultures further strengthens the hypothesis that *Sall4* is required specifically in the inner cell mass. To test the fate of ICM cells directly, I wanted to observe ICM growth in the absence of primary trophoblast cells.

Cell Culture After Immunosurgery

Trophoblast cells can be removed specifically from ICM by immunosurgery (Solter and Knowles, 1975). To do this, the zona pellucida is removed, blastocysts are incubated with an anti-mouse antibody for ten minutes, washed in cell culture medium and incubated in complement. Since only cells decorated with antibody will be attacked by the complement, solely those cells originally exposed to it are killed. The trophoblast forms a tight epithelium preventing access of antibody to the inner cell mass and thus the only cells surviving this procedure are the ICM stem cells. As can be seen in figure 13, size and appearance of the inner cell mass subsequent to immunosurgery is comparable in wildtype and knockout. Living cells form a round aggregate and separate from dead debris.

Cells were subsequently placed in culture medium and growth was monitored daily. The early inner cell mass is still plastic enough to contribute to trophoblast lineage if kept under suitable conditions. Several studies show independently that this potential is usually lost during 32 cell stage of a late expanding blastula (Handyside, 1978; Johnson and McConnell,

2004; Rossant and Lis, 1979). Other studies were able to identify growth of TSC colonies after immunosurgery even of e4.0 blastocysts (Hogan and Tilly, 1978). Estimation of cell numbers in blastocysts used for this study suggests that the trophoblast itself already consists of at least 32 cells (see figure 13), thus the blastocysts used for the experiments are consisting of well above 32 cells. Therefore, ICMs used in this experiment will usually generate ESCs and primitive endoderm. However growth of TSCs can not strictly be excluded.

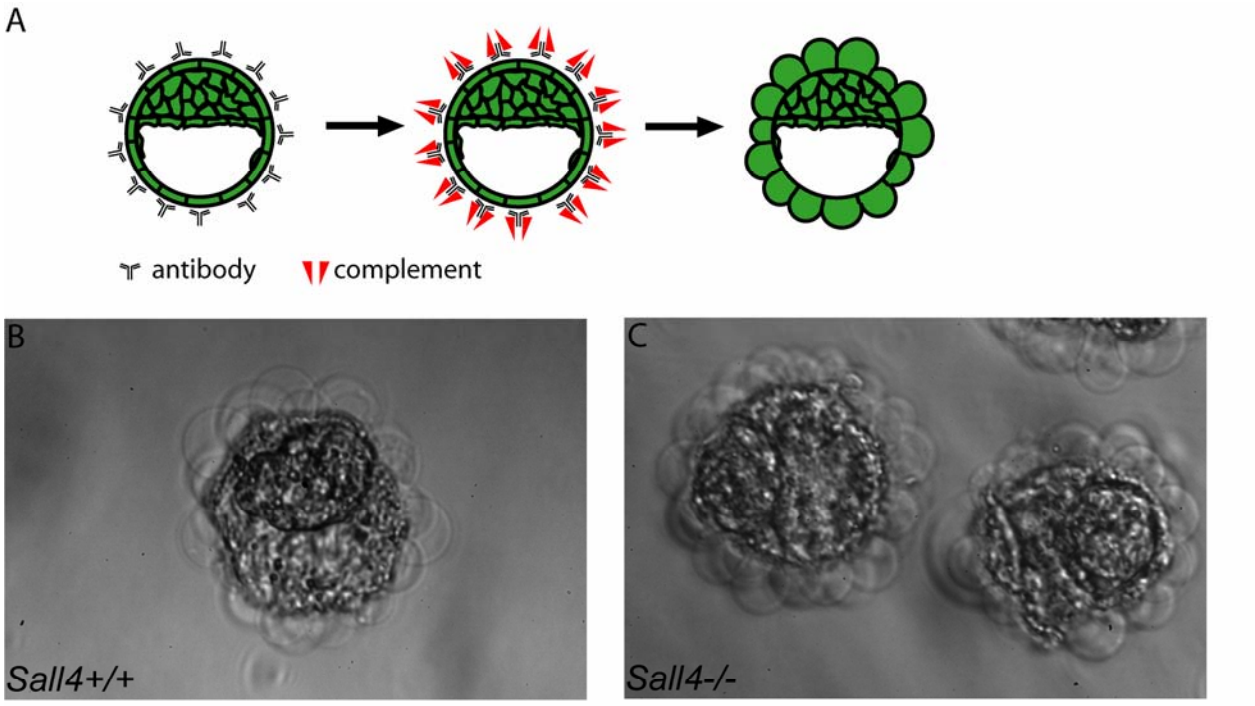


Figure 13: Immunosurgery of mouse blastocysts.
A) Blastocysts around 64 cell state were sorted, the zona pellucida was removed and they were incubated with rabbit serum immunized against mouse T cells for ten minutes to decorate all cells with surface exposed to the antibody solution. Subsequent washing in culture medium and incubation in guinea pig complement killed all antibody marked cells leaving only inner cell mass (ES cells and primitive endoderm) alive.
B, C) Blastocysts after immunosurgery showing swollen, dead trophoblast cells as well as a inner cell mass rounding up and seperating from the dead tissue. Cell number of blastocysts can be estimated based on dead trophoblast cells visible on the lateral surface of the blastocysts.

Obtained ICMs were cultured at two different conditions, in ES cell medium on gelatinized plates as well as in TS cell medium on culture dishes containing a monolayer of mouse embryonic fibroblasts.

In ESC medium, growth of knockout ICM cells is arrested. Cells form stable planar or spherical aggregates while control ICMs undergo rapid cell cycle giving rise to huge, loose aggregates of ESC like cells as well as to primitive endoderm cells migrating on the surface of

the culture dish. Initially, primitive endoderm cells (PE) can be identified clearly also in knockout colonies. After eight days, trophoblast giant cells can be observed in knockout ICMs while no giant cells are ever observed in cultures of wildtype inner cell mass. This experiment shows that ES cell deleted of *Sall4* lost the capacity to grow and self-renew in ES cell medium. Instead, differentiation into trophoblast giant cells occurred.

To test the capacity of ICMs to generate primitive endoderm or trophoblast, I cultured ICMs in TSC medium on a feeder layer of mouse embryonic fibroblasts subsequent to immunosurgery. Feeder cells were treated with mitomycine C prior to seeding the ICM so that feeder cells were inhibited in their cell cycle. Under TSC medium conditions, ES cells do not grow while primitive endoderm and trophoblast stem cells can be cultured for many passages in TSC medium in the presence of mouse embryonic fibroblasts and FGF4 (Kunath et al., 2005; Tanaka et al., 1998).

Of the 24 ICMs of the control group, 21 had settled down and formed colonies of primitive endoderm after four days while only little growth was observed in *Sall4* deleted ICMs. Nevertheless, the slow growth observed shows that cells originating from an inner cell mass that lost *Sall4* are viable and in principal capable to proliferate in trophoblast stem cell medium. This supports the notion that loss of Sall4 does not lead to cell death.

As shown in figure 14, epithelial cells reminiscent of TS cells grow on a feeder layer of mitomycine C treated fibroblasts in the knockout after seven days while cells in the wildtype appear more roundish in shape. TS cell medium supplies optimal conditions for the generation of primitive endoderm cell lines (Kunath et al., 2005). To identify the two types of cells undoubtedly, colonies for each genotype were cultivated seven days more, now in the absence of FGF4. Under these conditions, wildtype colonies disassociated and cells started migrating. Typical network like patterns of cell contacts allow identification of primitive endoderm cells in culture (XEN cells) (Kunath et al., 2005). Upon withdrawal of FGF4 in the knockout group, trophoblast giant cells differentiated in all wells, where growth had originally been observed. No giant cells could ever be observed in the control plates. Therefore it was concluded, that knockout ICM cells gave rise to trophoblast stem cells that differentiated into giant cells upon deprivation of FGF4. Due to the enormous size and polyploidity of giant cells, cell identity can be addressed by appearance. Differentiation into giant cells is the expected reaction of TSC in the absence of FGF4 and shows that TS cells derived from the inner cell mass of a *Sall4* knockout blastocyst are competent to sense FGF4.

Culture of ICMs in TSC medium proofs the competence of *Sall4* knockout ICM derived cells to survive and grow though the competence to maintain ICM cell types is lost. Both cell

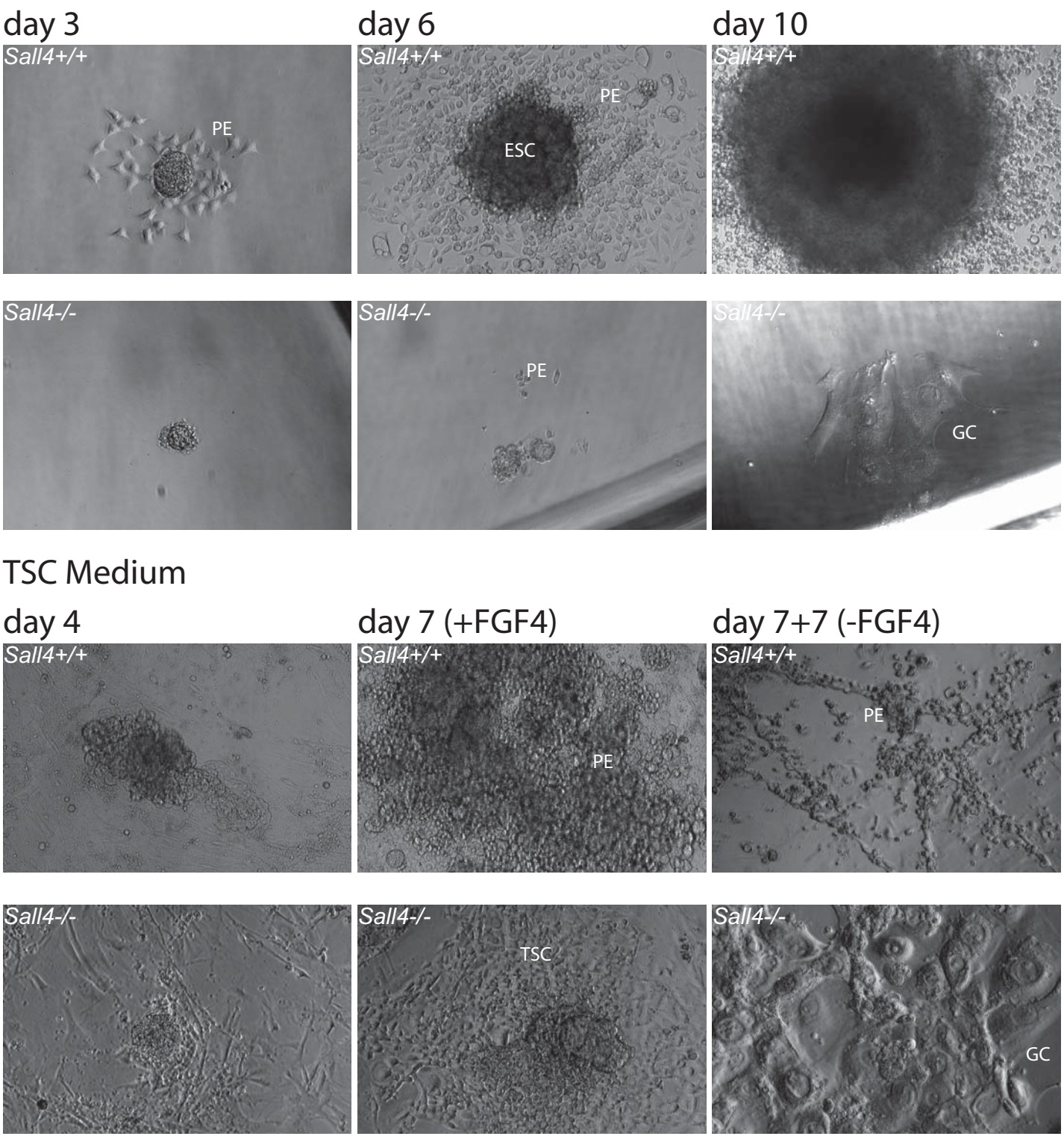


Figure 14: Cell culture of inner cell mass (ICM) after immunosurgery.
Based on the observation that *Sall4-/-* inner cell masses express TROMA in a blastocyst outgrowth, isolated ICMs were observed in culture. Growth was documented by photography.
When kept in ESC medium, wildtype ICMs grow rapidly. ES cells (ESC) either proliferate as loose aggregations (top) of cells or differentiate into embryoid bodies (data not shown). The second lineage of the ICM, the primitive edoderm (PE) migrates and spreads on the culture dish surface (PE). Single PE cells typically form vacuoles and die. In *Sall4* knockouts only very few cells and no ES cells are observed. Instead, Giant cells form from the inner cell mass (GC).
In trophoblast stem cell (TSC) medium supplied with FGF4, ESCs do not grow in either wildtype or mutant. While PE cells grow in the wildtype, no PE can be observed in *Sall4-/-* ICMs. TSCs grow from the inner cell mass culture of *Sall4-/-* blastocysts. These TSCs differentiate into giant cells upon withdrawl of FGF4. Thus *Sall4-/-* ICMs are competent of growth in an undifferentiated manner but can not maintain inner cell mass fates by self renewal.

types usually generated by the ICM, ES cells and primitive endoderm cells, fail to grow suggesting that *Sall4* is required to maintain inner cell mass identity. The formation of functional TS cells that are able to differentiate into giant cells illustrates that the trophoblast lineage is not severely affected by loss of *Sall4*.

Taken together with the observation that giant cells form even in conditions favouring ES cell fate, it suggests that inner cell masses lost the capacity to maintain ES cell and XEN cell lineages. Instead, cells differentiate into trophoblast lineage.

Deletion of Sall4 from ES Cell Lines

Two cell types are present in the inner cell mass of an e4.0 blastocyst. Adjacent to the blastocoel, cells forming the primitive endoderm will assemble while cells surrounded by trophoblast and primitive endoderm give rise to the embryo itself as well as the extraembryonic mesoderm (Johnson and McConnell, 2004; Rossant, 2004). Both cell types are interdependent. This is also shown by the failure to form a wildtype blastocyst outgrowth observed for mouse knockout models that affect only one of the two cell lines (Cheng et al., 1998; Nichols et al., 1998). In culture, primitive endoderm cells grow as XEN cells while the undifferentiated apolar cells inside the ICM are known as embryonic stem cells (ESC). Both cell types express high levels of *Sall4* and so I wanted to know which cell type is responsible for the effects described in a *Sall4* knockout situation.

To study, which of the two cell types is primarily affected by a *Sall4* deletion, or if both cell types require the protein, I wanted to delete *Sall4* from an ES cell line in culture. In order to do that, an ES cell line had to be established, that was homozygous for *Sall4-loxP*. Mice homozygous for the floxed allele were bred to heterozygously floxed animals, blastocysts were collected and allowed to form a blastocyst outgrowth. At day four in culture, the inner cell mass was picked with a drawn glass capillary and dissociated in trypsin. Single cell suspensions for each clone was placed in one well of a 96 well plate and cultivated until ESC colonies formed. ESC lines were expanded and genotyped by southern blot. Figure 15A shows analysis of the established cell lines with a probe that recognizes the floxed exon of *Sall4*. Digestion with restriction enzymes BglII and XbaI results in a DNA fragment running with a size of 3.8kb in agarose gel electrophoresis. An additional XbaI site in the FRT sequence shortens the fragment to 2.4kb. A probe against the first part of *Sall4* exon two was used to screen the derived ES cell lines. Of the six newly generated ESC lines, 4 clones were heterozygous for *Sall4-floxed* and two were homozygous.

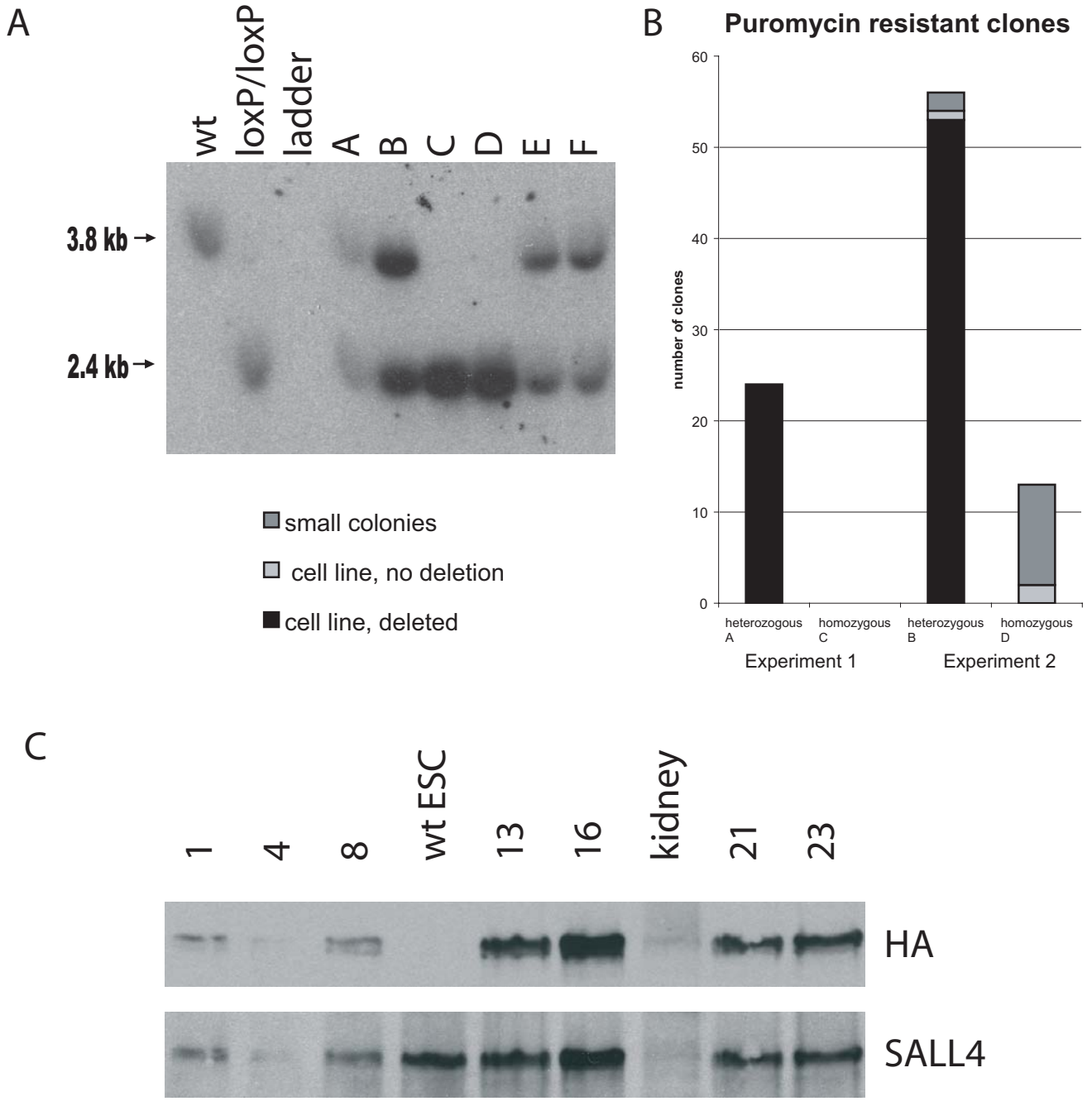


Figure 15: Sall4 protein level modification in ESC culture

A) Establishment of ESC lines carrying floxed *Sall4* alleles. In order to genotype the resulting ESC lines, southern blot analysis was carried out. The *Sall4* locus was marked with a probe recognizing the floxed exon two. DNA was digested with BglII and XbaI. The latter cuts in the FRT site next to the 5´loxP site and results in a shift of the DNA fragment containing *Sall4* exon two from 3.8 kb to 2.4kb. Clones C and D are homozygously floxed, clones A, B, E and F are heterozygous.
B) Twelfe million cells of clones D and E (experiment 1) as well as A and C (experiment 2) were electroporated with a linearized vector carrying Cre and IRES-puroR under the CAGGS promotor. Selection in cell culture, picking of single clones and subsequent genotyping of the *Sall4* locus showed that ESCs deleted of *Sall4* can not be expanded in cell culture under conditions suitable for growth of ESCs while control clones as well as clones that did not delete *Sall4* grew rapidly.
C) Overexpression of Sall4 in cell culture. ESCs heterozygous for *Sall4* (*Sall4+/Sall4-GFP*) were electroporated with a vector carrying an HA-tagged *Sall4* cDNA under the control of the CAGGS promotor. Selection for stable transformants was possible through a bicistronic transcript for Sall4 and puroR. Western blot analysis shows overexpression of Sall4 protein in clone 13 and 16 .

To delete *Sall4*, one heterozygous as well as one homozygous cell line were expanded separately and 12 million cells each were electroporated with a vector coding for a bicistronic transcript of Cre recombinase and puromycin resistance under the control of the CAGGS promoter (Sunaga et al., 1997). 14 days of selection for puromycin resistance resulted in stably transfected clones that should express Cre recombinase. By the enzymatic activity of Cre, recombination between loxP sites should have occurred in all alleles flanked by the recombination sites. Homozygously floxed clones should then loose Sall4 during the selection procedure. Resulting clones were picked, expanded and genotyped to test for successful deletion. To genotype the clones, a PCR strategy was devised in which the 5´ primer was positioned overlapping with the 5´ loxP site and two 3´ primers were located distally to each loxP site. While a wildtype locus can not serve as a template for a PCR reaction with these primers, floxed loci can result in two sizes of amplicons, dependent on whether Cre-excision of *Sall4* exon two had occurred.

The heterozygous ESC line yielded 24 stable clones while no clones grew in the homozygous cell line. To verify this result the experiment was repeated with two independently established ES cell lines. In the second round, 56 clones were established in the control while the homozygous cell line gave rise to only two normal growing clones from which a cell line could be established as well as eleven independent, very small colonies that stopped growth and could not be expanded even after two additional weeks in culture.

Genotyping revealed that the two normal growing clones of the homozygous cell line did not delete *Sall4* while the small colonies that terminated growth were identified as *Sall4* null clones. Of the heterozygous clones, only one had failed to delete the floxed allele indicating that excision should occur in almost all cases and that the absence of deleted clones in the homozygous ES cell line is due to requirement of Sall4.

The most likely explanation for the failure to delete *Sall4* in puromycin resistant colonies is the deletion of the Cre recombinase sequence of the electroporated plasmid through the integration into the genome. The absence of growing, *Sall4* deficient colonies in the homozygous *Sall4-floxed* ESC lines as opposed to 24 and 54 colonies in the control cell lines proofs the requirement of *Sall4* for growth of ESC lines in culture. Small colonies can be explained by the fact that homozygous deletion of *Sall4* may take several cell cycles and counter selection of that event will take place due to the requirement of Sall4. Complete loss of Sall4 protein might thus take a few days, sufficient to generate a small colony.

Loss of the transcription factor *Oct4* leads to differentiation of the inner cell mass into trophoblast, similar to the effects observed here. Loss of Oct4 in an ESC line can also not be

tolerated. Overexpression of Oct4 in ES cells leads to differentiation into primitive endoderm (Niwa et al., 2000). To test if Sall4 can be over expressed in ES cells, a vector coding for an HA-tagged version of SALL4 under the control of the CAGGS-promoter was assembled and electroporated into a cell line heterozygous for *Sall4-GFP*. As seen in Figure 15C, clone 16, overexpression of about 3-4 fold Sall4 compared to normal ES cells is tolerated. No phenotypic differences of these cells were observed compared to wildtype ES cells.

3.6. Sall4 is Required Cell Autonomously in vivo

Requirement for Sall4 in Inner Cell Mass Cells

The cell culture experiments indicated a requirement for *Sall4* specifically in ES cells. Further experiments were designed to confirm this finding *in vivo*. To test if *Sall4* requirement in ES cells is cell autonomous *in vivo*, studies on chimeric animals were performed. Morulas were collected at e2.5, sorted by GFP intensity and the zona pellucida was removed. In parallel, e2.5 morulas heterozygous for *ROSA-lacZ* (Soriano, 1999) were harvested and treated in the same way. Subsequently, each morula of a *Sall4-GFP* interbreeding was placed in small aggregation well together with one morula heterozygous for ROSA-lacZ. Under such conditions, blastomeres of both morulas attach, intermingle and form a chimeric morula in over night culture. In one third of the aggregation wells, chimeric blastocysts arose. In these blastocysts, GFP positive and negative cells had intermingled and each genotype had presumably given rise to all cell types. Immunostaining of such blastocysts revealed that the GFP positive cell line had indeed contributed to trophoblast and inner cell mass (figure 16A). The fact that *Sall4-GFP* homozygous cells are capable to contribute to the inner cell mass and do not sort exclusively to the trophoblast even in presence of competing wildtype cells indicates, that these cells are initially capable of ICM formation. This confirms earlier observations.

To address the question if cells homozygous for *Sall4* can contribute to the developing embryo, chimeric blastocysts of each genotype were transferred into pseudopregnant females. Embryos were collected at day 10, fixed and subjected to Bluo-Gal staining to identify the contribution of ROSA-lacZ cells in these embryos. As expected, aggregations of ROSA-lacZ morulas with cells heterozygous or wildtype for *Sall4* resulted in chimeric animals as identified by speckled blue staining (figure 16B). However, embryos obtained from aggregations of *Sall4* knockout morulas with *ROSA-lacZ* morulas, collected at e10.5 stained blue in all cells. Paraffin sectioning of embryos confirmed this result. Analysis of the

trophoblast in aggregations of knockout morulas with *ROSA-lacZ* revealed the presence of *Sall4* deleted cells in the placenta. Thus it was concluded that *Sall4* is required in a cell autonomous fashion in the embryo proper but not in the trophoblast.

In mice, tissue specific gene deletion allows dissection of phenotypes. To do so, mice carrying floxed alleles are bred to a tissue specific Cre line resulting in excision of the loxP flanked sequence in all cells where Cre recombinase is expressed. A *Sox2-Cre* strain was shown to mediate efficient recombination in the inner cell mass (Hayashi et al., 2002). Deletion starts at e3.5, by e6.5 all cells of the embryo proper have lost the floxed allele. At embryonic day 7.5, the visceral primitive endoderm will also have deleted efficiently. I decided to make use of this Cre-line and genetically ablate *Sall4* from the embryo proper without disturbing development by *ex vivo* manipulations. A mouse carrying a *Sall4-GFP* allele as well as one copy of the Cre transgene was crossed to mice homozygous for *Sall4*-loxP and litters were taken at e8.5 to allow development post deletion. 25% of mice carried a *Sox2-Cre* transgene as well as a *Sall4-GFP* and a *Sall*4-*loxP* allele as expected. In these mice, Sox2-Cre should have deleted the floxed allele of *Sall4* in all cells of the embryo proper between e3.5 and e6.5 while the primitive endoderm should delete only one day before harvest of the embryos In figure 16C, deletion is symbolized by darker tones of green. All trophoblast derived tissues will in this setup still be heterozygous for *Sall4*. Figure 16D compares the appearance of such an embryo with a *Sall4* heterozygous littermate. A developmental arrest of the embryo is clearly visible. Control embryos had undergone gastrulation and neurulation and are visible through the yolk sack. In the knockout population, embryo proper tissue was found as an unorganized clump. Deletion of *Sall4* mediated by a Sox2-Cre line takes until e6.5 in the embryo proper while visceral endoderm is still genetically wildtype at this stage. Embryos most severely affected do not even reach a state equivalent to e6.5. Since deletion in the primitive endoderm occurs only after e6.5, this experiment supports the finding that Sall4 is required in the embryo proper itself. Whether the requirement is cell autonomous can not be addressed in this experiment.

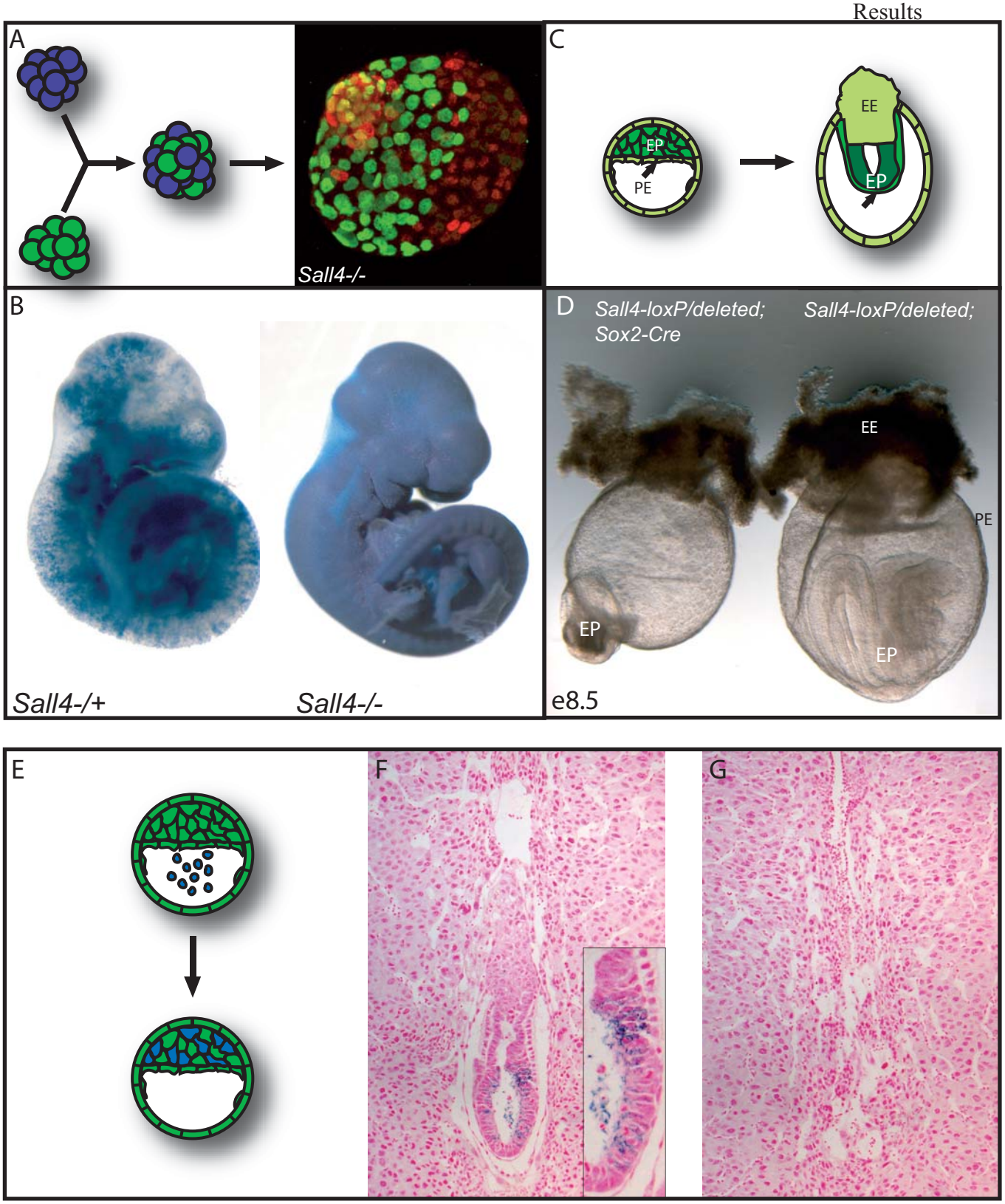


Figure 16: Chimeric analysis of *Sall4-/-* embryos

A) To test if Sall4 requirement is cell autonomous, morulas sorted for GFP intensity were aggregated with morulas heterozygous for ROSA-lacZ. Retransferred animals were collected at e10.5. Both morulas contribute to the inner cell mass stained for Oct4 (A, red staining). While aggregation of ROSA-lacZ with morulas wildtype or heterozygous for *Sall4* resulted in chimeric embryos at e10.5 as expected (B), aggregation with homozygous morulas resulted in embryos derived exclusively from ROSA-lacZ cells.

C) TheSox2-Cre line results in deletion of loxP flanked sequences first in the embryo proper (EP), from e7.5 on also in the primitive endoderm (PE). The extraembryonic ectoderm (EE) of the trophoblast does not delete.

D) Conditional excision of *Sall4* leads to developmental arrest of the embryo proper.

E) ESCs expressing cytoplasmatic beta galactosidase were injected into sorted blastocysts. Such ESCs can only contribute to the epiblast as seen in the control experiment (F). No beta galactosidase staining could be observed in any of the retrieved implantation sites of injected knockout blastocysts at e6.5 (G).

Green: *Sall4-GFP,* blue: *ROSAlacZ*, dark green:deletion of *Sall4* by Sox2-Cre;

Requirement of Sall4 in Primitive Endoderm

To address directly whether *Sall4* deficient cells can form a functional primitive endoderm *in vivo*, *ROSA-lacZ* ES cells were injected into blastocysts from a *Sall4-GFP* interbreeding after sorting the blastocysts by the intensity of the GFP signal. Injected ES cells can contribute to the embryo proper but do not contribute to the primitive endoderm or trophoblast (Tam and Rossant, 2003). Therefore, the defects due to *Sall4* deletion in the epiblast may be rescued by the presence of *ROSA-lacZ* cells while the primitive endoderm and the trophoblast will only consist of cells originating of the blastocyst. As shown before, *Sall4* is not required for trophoblast differentiation. Developmental potential of the chimeric animals resulting from this experiment will thus be determined by the integrity of the primitive endoderm.

Blastocysts were transferred into pseudopregnant females 2.5dpc and allowed to develop until e6.5. Whole mount BluoGal staining was used to distinguish between *Sall4*-/- cells of the blastocyst and injected *ROSA-lacZ* cells. In Figure 16E-G, sections through e6.5 embryos illustrate a contribution of *ROSA-lacZ* cells to the epiblast. Such contribution of wildtype *ROSA-lacZ* cells to the epiblast and not to any other cell lines was observed as expected in injections of control blastocysts. Injection into *Sall4* knockouts did not rescue the phenotype, these ESCs could not even be located at e6.5, suggesting that they did not survive. This experiment argues that defects outside the embryo proper, most likely the primitive endoderm, contribute to the *Sall4* phenotype.

4. Sall1, Sall2 and Sall3 Function in Neurogenesis

In contrast to *Sall4*, expression for *Sall1*, *Sall2* and *Sall3* largely overlaps as has been presented earlier. Based on the finding that loss of Sall function leads to severe defects in early development, I wanted to investigate if additional, unpublished phenotypes might be visible upon loss of one or several members of the SALL protein family.

4.1. Loss of Sall1 or Sall3 Results in Perinatal Lethality

In order to obtain animals homozygous for *Sall1-lacZ*, heterozygous mice were intercrossed. Shortly after birth, litters were visually inspected to screen for phenotypes. Subsequently pups were genotyped to correlate any observations to the genotype.

Litters from *Sall1* intercrosses regularly contained pups that had no milk in their stomach. This could easily be observed due to the translucent nature of the body wall at this age. Genotyping of the pups of such litters showed that lack of milk correlated with homozygosity for the *Sall1* knockout allele. Animals deleted of *Sall1* die within the first six hours after birth appearing dehydrated but without showing any further obvious abnormalities that can be observed. Breathing and reflexes are normal.

Loss of *Sall3* results in perinetal lethality (Dr. G. Schütz, personal communication). Because of similar expression patterns with *Sall1*, *Sall3* litters were specifically analyzed at P0 for their ability to suckle milk. Genotyping litters directly after birth could show the presence of knockout animals and again these mice had not suckled milk. Dehydration is a likely cause for lethality within six hours after birth (Hongo et al., 2000). Thus, the failure to suckle milk might contribute to the lethality both for *Sall1* and *Sall3* knockout animals. Previously, severe kidney defects ranging from absence of kidneys to severe hypomorphia upon loss of *Sall1* have been described and will also lead to lethality (Nishinakamura et al., 2001).

4.2. Macroscopic Analysis

Anatomical defects in mouth or oesophagus could result in the inability to suckle milk. Examination of pups showed no major defects. A recent report observed minor deformations in the palate in *Sall3* knockouts (Parrish et al., 2004). Upon re-examination of *Sall3* based on

this information still I could not confirm defects of the kind described. This might be due to genetic background effects.

Sall1 and *Sall3* are expressed predominantly in the central nervous system. Expression in the palate has not been observed. Therefore I hypothesized that a defect in the CNS might explain the lethality of mice homozygous for *Sall1* or *Sall3*. This is a plausible hypothesis since several mouse models link neuronal defects to the failure to drink milk.

4.3. Histological Analysis of the Olfactory Bulb in Mice Mutant for Sall1 and Sall2

Rodent models for anosmia result in postnatal lethality due to the failure to suckle milk. Amongst them are mechanical ablation of the olfactory epithelium in rats (Risser and Slotnick, 1987) and genetic targeting of a cyclic nucleotide gated cation channel (Brunet et al., 1996).

As described earlier, *Sall1*, *Sall2* and *Sall3* are all strongly expressed in the olfactory bulb. All receptor neurons in the olfactory epithelium project their axons onto the surface of the olfactory bulb. Mammalian olfactory epithelia harbour up to 1000 different types of receptor neurons, each expressing only one odorant receptor (Mori et al., 1999). For each kind of receptor, axons of all neurons expressing it converge at the surface of the olfactory bulb in so-called glomeruli. Dendrites of underlying mitral and tufted cells invade each glomerulus to form excitatory synaptic connections with the olfactory receptor neurons. In mice, each glomerulus receives converging axonal inputs of several thousand sensory neurons while only around 20 mitral and tufted cells innervate it (Royet et al., 1998). Axons of all mitral and tufted cells project the information through the lateral olfactory tracts (LOT) to the olfactory cortex of the brain. Mitral cells are organized in a dense layer (MCL) between glomeruli and the high number of GABAergic inhibitory interneurons in the center of the olfactory bulb. These form local circuits through dendrodendritic inhibitory connections to mitral and tufted cells. A rostral extension of the lateral ventricle extends into each olfactory bulb.

To identify a possible defect in the olfactory bulb responsible for the inability of newborns to suckle milk, sagittal cryostat sections were prepared and nuclei were stained with nuclear fast red. In a wildtype mouse at P0 the mitral cell layer is clearly visible while this dense neuronal layer can not be observed in a *Sall1* knockout (figure 17A). In contrast, mice mutant for *Sall2* display a clearly identifiable MCL at e19.5. This was expected, as these mice show normal feeding behaviour. In the double mutant for *Sall1* and *Sall2*, the defect becomes

even more pronounced. In order to study the maximal defect caused by loss of *spalt* genes in the olfactory bulb I decided to study the phenotype in a *Sall2* deleted background.

4.4. Comparison of the Effects of a Loss of Sall1, Sall2 and Sall3 on the Olfactory Bulb

To analyze the contribution of *Sall1* and *Sall3*, mice heterozygous for *Sall1* and *Sall3* as well as homozygous for *Sall2* were intercrossed. Litters were taken at embryonic day 18.5 to exclude loss of pups. Figure 15B presents examples of observed phenotypes in different allelic combinations. Mice mutant for *Sall2* exhibit normal interneuron- and mitral cell layer organization. In coronal sections, layers appear as concentric rings around the ventricle. As shown previously, no mitral cell layer (MCL) was detected in *Sall1*/*Sall2* double mutants while no obvious effect on underlying granular neurons (GC) was observed.

In double mutants for *Sall2* and *Sall3* only mild effects on MCL organization could be seen. In contrast, the outer plexiform layer (bar in figure 17B, OPL) contains almost no cells bringing the MCL closer to the surface of the olfactory bulb.

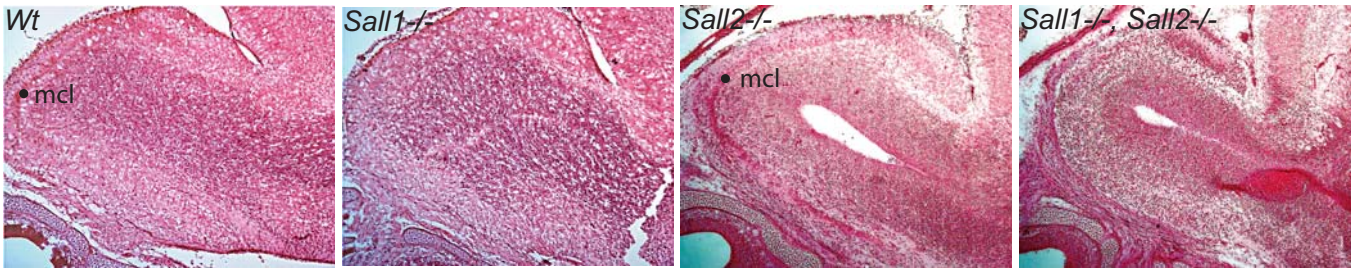
Loss of all three *Sall* genes, *Sall1*, *Sall2* and *Sall3* results in almost complete absence of the olfactory bulbs. No internal lumen could be observed and all dense neuronal structures are missing including granular neurons. Due to the minimal rostro-caudal dimension, the only sections obtained rather resembled superficial tangential sections of the cortex.

For further studies of the described histological defects with molecular markers I decided to focus on a comparison of *Sall2* knockout mice to *Sall1*/*Sall2* double knockouts for two reasons:

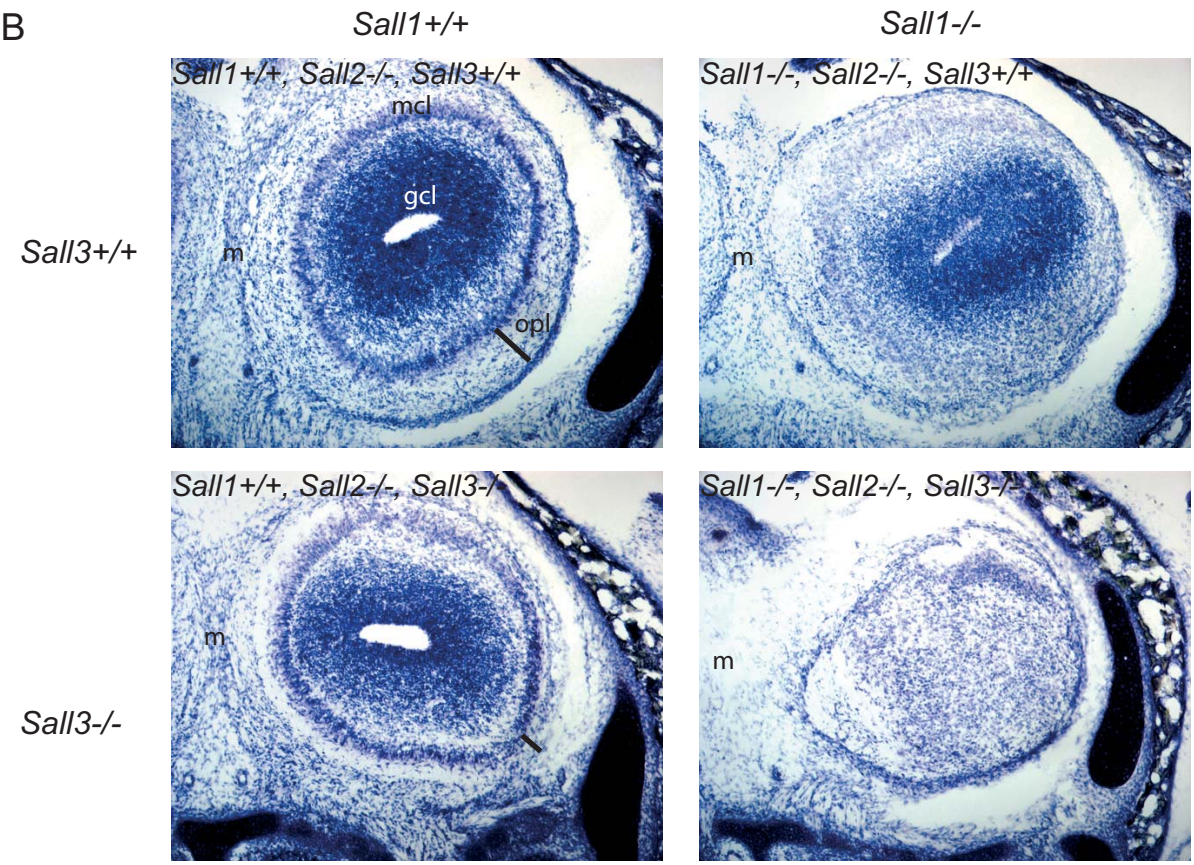
First, breeding for triple knockout mice is time consuming and inefficient. The parental generation was obtained from interbreeding *Sall1*-/+, *Sall2*-/-, *Sall3*+/+ males with *Sall1*+/+, *Sall2*-/-, *Sall3*-/+ females and results theoretically in one female double heterozygous for *Sall1* and *Sall3* in eight pups born. To analyze e18.5 litters that should contain a triple knockout in one of 16 pups, such females bred to males of the same genotype had to be sacrificed. Expected accumulation of *Sall* deletion was found in a sub-mendelian frequency in both generations and mating efficiency was very low.

Secondly, I wanted to study neurogenesis in the olfactory bulb using molecular markers. Olfactory bulb tissue is absent in a triple knockout and so I focused on the MCL in a *Sall1*/*Sall2* double knockout situation.

A



B



C

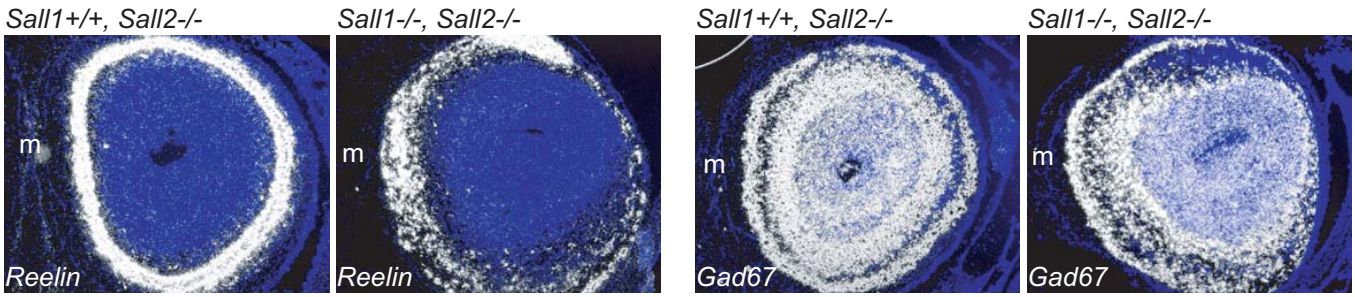


Figure 17: Loss of mitral cells in a *Sall1/Sall2* compound knockout.

A) H&E staining of sagittal sections of olfactory bulbs e19.5 (wt and *Sall1-/-*) or e18.5 (*Sall2 -/- and Sall1-/-, Sall2-/-*) shows developmental defects caused by loss of Sall1 in the mitral cell layer (mcl). These changes are more pronounced in a *Sall2* deleted background.

B) Coronal sections of olfactory bulbs e18.5 stained with thionine to compare the effect of Sall1 or Sall3 loss or both in a Sall2 deleted background. Loss of Sall1/2 results in absence of a visible mcl while loss of Sall2/3 causes a thinning of mainly the outer plexiform layer (bar, opl) containing horizontal processes of mitral and tufted cells as well as cell nuclei of local circuit neurons. In a triple knockout of *Sall1, Sall2* and *Sall3*, all OB structures are absent. midline (m)

C) In situ hybridisation on a *Sall1/2* double knockout in comparison with a wildtypic *Sall2* knockout to molecularly label the mitral cell layer (Reelin) as well as local circuit inhibitory (GABAergic) neurons (Gad67). Some mitral or tufted cells are still found in a *Sall1/2* knockout but they are no longer organized in a condensed, pallial layer. Thus, cells of the opl and gcl also appear less organized and the two layers are further apart. Lateral, the opl specific *Gad67* staining is absent.

To verify the absence of mitral cells from the olfactory bulbs, *in situ* hybridization was performed. *Reelin* is specifically expressed in mitral and tufted cells (Bulfone et al., 1998) and was used to detect these cell types. The concentric ring observed in histological stains labels intensely for *Reelin* in all control animals (figure 17C). In contrast, staining was strongly reduced and dispersed in the mutant. *Reelin* positive cells are still present in the olfactory bulb of *Sall1*/*Sall2* deleted mice and are probably tufted cells, typically present in a scattered distribution (Bulfone et al., 1998). To verify this result, *Gad67*, expressed in the most superficial periglomerular layer (PG) as well as the underlying granule cell (GC) layer (Long et al., 2003), was used as an *in situ* probe. These layers were not severely affected by loss of *Sall1* and *Sall2*. They also did also not collapse as would be the case if mitral and tufted cells were missing completely.

These results demonstrate that cells of the mitral cell layer are almost absent in the described mutant and their organization is severely impaired. It is still possible though that few mitral cells remain functional. To investigate this, further analysis was necessary.

4.5. Remaining Mitral Cells do not Project to the Olfactory Cortex in a Sall1/Sall2 Double Knockout

To further characterize the neurons in a *Sall1/Sall2-/-* olfactory bulb, the cellular extensions of neurons were studied. Rhodamine red coupled phalloidine was used to visualize f-actin in cytoplasmic extensions of neurons. Figure 18A shows a coronal view on the medial olfactory bulb region. GFP expression in the olfactory bulb outlines the mitral cell layer in a heterozygous mouse. Phalloidine indicates the presence of the nerve layer of the sensory neurons (NL), the glomeruli (dashed circles), dendrites of the mitral cells in the outer plexiform layer (OPL) as well as the axonal extensions of the mitral cells to the olfactory cortex running through the inner plexiform layer (IPL).

While the nerve layer as well as rudimentary glomeruli can be observed in the knockout situation, dendrites and axons of the mitral cells are absent. In contrast, the glomerular focal points, in which axons of the olfactory epithelium converge, are still present. This is not surprising as it had been shown earlier, that axons of the olfactory epithelium sort independently of mitral cells and GABAergic interneurons (Bulfone et al., 1998).

To follow the neuronal projections of sensory neurons in the olfactory epithelium (OE), grains of the lipophilic dye DiI were placed on the olfactory epithelium of mice collected and fixed at e19.5. Samples were stored at room temperature in 2% PFA for six weeks to allow

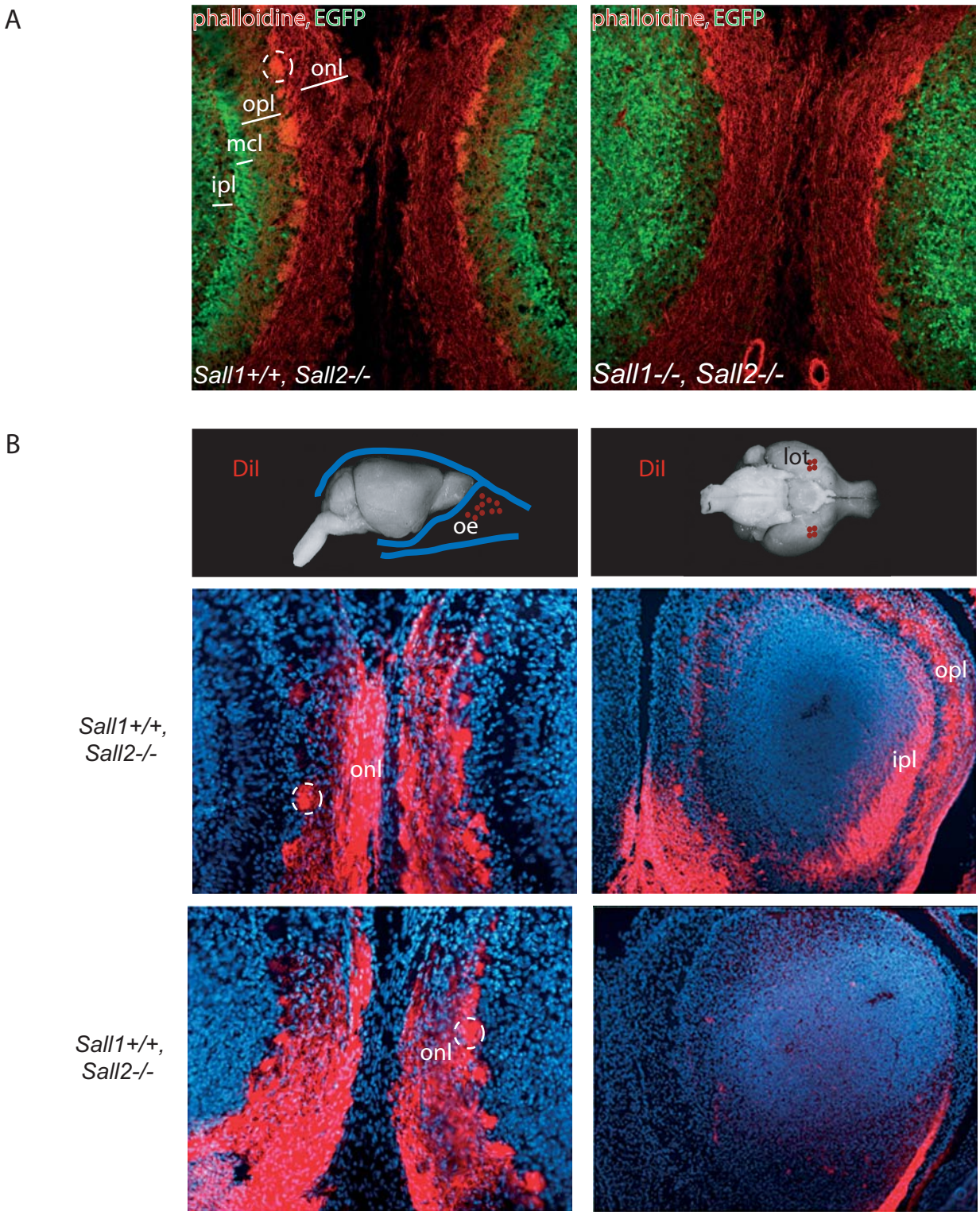


Figure 18: Axonal projections in the *Sall1-/-, Sall2-/-* olfactory bulb.

A) shows actin staining with rhodamine-coupled phalloidine to visualize cytoplasmic extensions like dendrites and axons. Together with EGFP expression from the *Sall1* locus, the outer nerve layer (onl), glomeruli (dashed circle), outer plexiform layer (opl), mitral cell layer (mcl) and inner plexiform layer can be distinguished. Axons from the olfactory epithelium (oe) reach the glomeruli through the onl and connect in each glomerulus with dendrites of mitral cells underneath. lateral extensions of mitral cells in the opl as well as axons collectingin the ipl to reach the cortex are seen as red staining within the olfactory bulb. In the mutant, onl and glomeruli-like structures are still visible while no mitral cells or mitral cell extensions can be observed.

B) DiI labeling was used to follow projection paths in the olfactory bulb. DiI grains were either placed in the olfactory epithelium (oe) or on the lateral olfactory tracts (lot) to allow incorporation into cells or celular extensions present and subsequent lateral diffusion in the membrane. Forward labelling of neurons of the olfactory epithelium did not reveal changes between control and knockout while retrograde labelling of mitral cells by placement of DiI to the site, where axonal projections should be found (lot), shows that the remaining reelin positive cells do not project to the olfactory cortex. Thus it can be assumed that Sall1/Sall2 knockout mice are anosmic.

incorporation of the dye into membranes of the olfactory epithelium and lateral diffusion into axonal projections. In accordance with the results obtained with phalloidine, DiI labelling and subsequent sectioning the olfactory bulb revealed no differences in the projectory behaviour of OE neurons. In both genotypes, olfactory receptor neurons had reached the olfactory bulb and formed focal points of axonal convergence visualized by DiI staining.

Next, I tested the projections of mitral cells by placing DiI grains on the lateral olfactory tracts (LOT) to allow for retrograde diffusion into the olfactory bulbs. In control animals, axons in the IPL as well as dendrites of the MCL in the OPL strongly label while no signal is detected in *Sall1*/*Sall2* double mutant animals (figure 18B).

This observation shows that scattered cells expressing *Reelin*, thus presumably remaining mitral cells or tufted cells, do not project to higher cortical regions in mice mutant for *Sall1* and *Sall2* and are therefore not functional. Absence of responsive mitral cells leads to anosmia in mice (Mori et al., 1999). The defect described in the olfactory bulb is thus sufficient to explain the inability of newborn *Sall1*/*Sall2* mutant mice to suckle milk and will therefore lead to the observed lethality.

4.6. Histological Analysis of the Neocortex in a Sall1/Sall2 Double Knockout.

The olfactory bulb is the most rostral extension of the forebrain and in continuum with the neocortex. To test if loss of *Sall* genes also affects neocortical neurogenesis, I collected embryos at different stages of development and compared the brain histology. Figure 19 shows thionine staining of neurons in coronal sections of *Sall2* and *Sall1/Sall2* knockout littermates at embryonic days 16.5 to 19.5. Figure19A exemplifies the differences in the appearance of the brain in a *Sall1*/2 double knockout and a *Sall2* littermate. The ventricles of the double knockout are always larger due to a less developed neocortex. Cells in the superficial, cortical layer of the neocortex appear compacted and display a more clearly defined boundary to the underlying regions. The corpus callosum (CC), the main connection between the two hemispheres, was less developed in all observed animals.

To describe the extent of this defect during development, mice were collected at different stages of neurogenesis. Figure 19B summarizes examples showing a dorsal section of the neocortex in littermates at e16.5 to e19.5. At all stages, the cortical layers in the knockout are thinner and contain fewer cells (bars). The compact structure of the neocortex observed in the control animals starts disappearing with differentiation of neurons at e17.5 as opposed to the double knockout where even at e19.5 a higher compaction of the cortical layers is observed when compared to the wildtype. Neuronal stem cells and precursor cells are located within the ventricular and subventricular zone while differentiating neurons migrate radially towards the cortex and differentiate (Takahashi et al., 1999). The ventricular zone of the neocortex adjacent to the lateral ventricles does not show significant changes in radial dimension in this staining, thus a reduction of the number of stem cells can not be shown. In conclusion, neurogenesis and differentiation of precursors does occur in the absence of *Sall1* and *Sall2*, the two *spalt* orthologs most prominently expressed in the neocortex. However, neurogenesis is reduced as shown by the smaller number of post mitotic neurons present in the cortical layers.

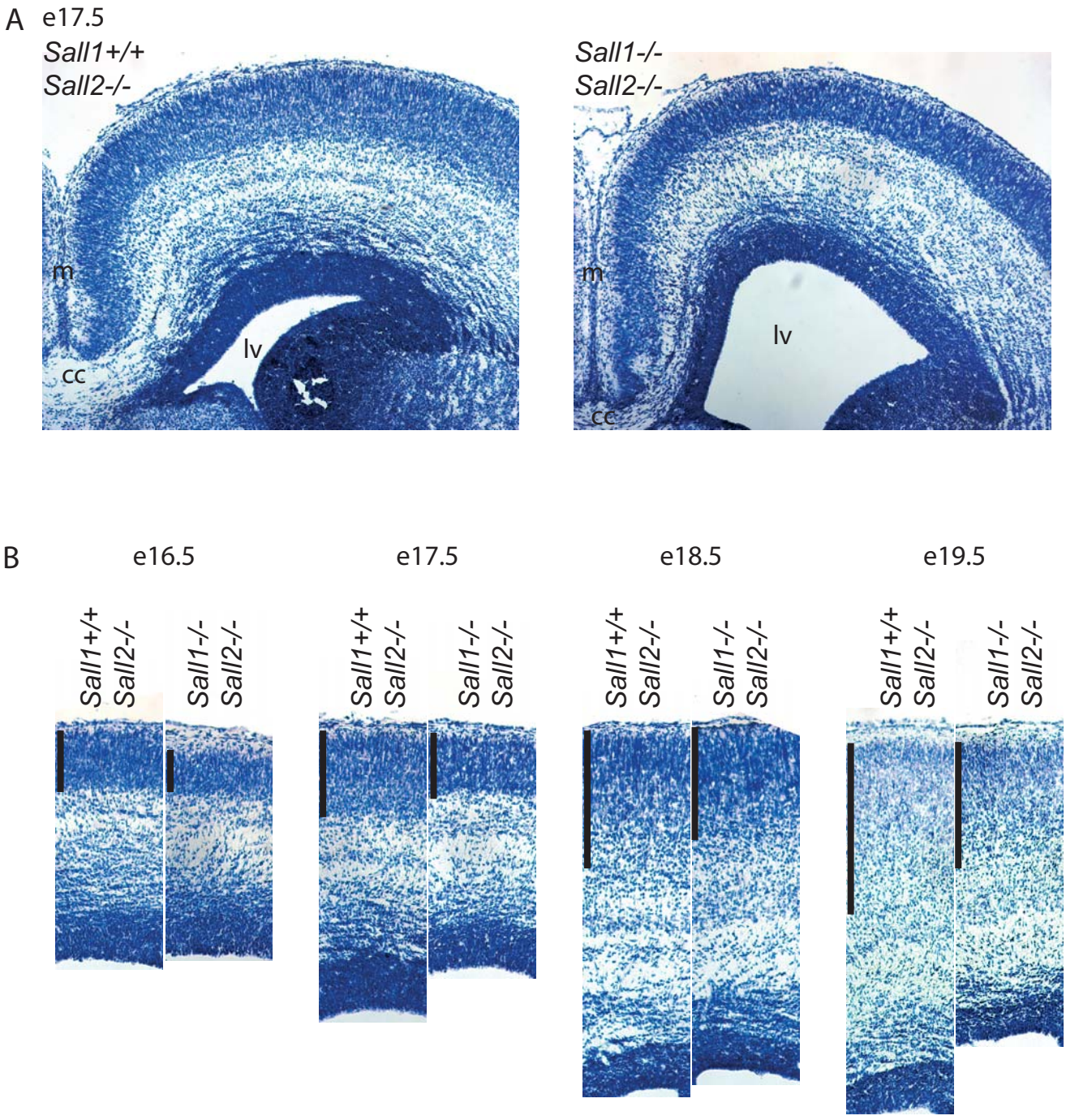


Figure 19: Neurogenesis of mice lacking *Sall1* and *Sall2*

Histological analysis of coronal sections through mouse brains. Shown are sections of one hemisphere of the neocortex from midline (m) on. For comparison, all sections were taken at the rostro-caudal height of the corpus calosum (cc) in the most caudal region of the lateral ventricle. Figure A shows the enlarged lateral ventricle of the mutant accompanied with a thinner cortex. In addition, the corpus callosum appeared thinner in the mutant.

In figure part B, only a small region of the dorsal part of the cortex is shown. The extent of the cortex is highlighted with bars. All sections compared were obtained from littermates. Thionine staining reacting with the nissl substance was used to label cell nuclei in blue.

4.7. Birthdating of Neurons with BrdU

Next, I wanted to quantify the described defects by labelling cells with BrdU. BrdU (bromo-deoxyuridine) is an analogon of deoxy-thymidine and will be incorporated in the DNA of cells in S-phase of cell cycle upon administration. To label embryonic cells that are currently undergoing DNA synthesis, timed pregnant dams were intravenously injected with BrdU dissolved in PBS.

In cortical neurogenesis, neuronal stem cells reside in the ventricular zone and will incorporate BrdU. These progenitors typically divide eleven times during neurogenesis. Each round of division, some of the daughter cells will leave the ventricular zone, migrate into the cortex and initiate differentiation. The number of cells remaining is above 50% in the beginning leading to an increase of the progenitor population and will drop below 50% later on. Neurogenesis occurs in two distinct steps. During the first rounds of cell cycle, neurons are generated that form the preplate (Takahashi et al., 1999) (see also figure 20A). Neurons generated later will subsequently migrate into the preplate and thus split it into a superficial marginal layer (ML) and the subplate (SP). Newly generated neurons will now always pass the last population and reside adjacent to the marginal layer or layer one, thereby generating new layers of neurons. By this, a correlation between the date of birth and the location of the neuron is established, with the neurons generated last residing most superficial. This process gives rise to a six layered cortex above the axonal tracts of the white matter.

By injection of BrdU at day 14.5 and 15.5 of pregnancy, neurons born at that age will be labelled while incorporated BrdU in the progenitor population will be diluted out again by several rounds of cell division (McEvilly et al., 2002). Figure 20B illustrates the relative location of neurons at day 19.5. While neurons born at e14.5 predominantly reside in superficial layers of the cortex in the control animals, almost equal contribution to all cortical layers can be observed in a *Sall1*/2 double knockout. This effect is more pronounced when labelling neurons at e15.5. In the control, neurons almost exclusively reside in the most superficial layers while a contribution to deeper layers is observed in the mutant. Since the cell cycle in lateral regions of the neocortex is slightly ahead of the most dorsal part, the effect is best seen laterally. Black bars outline the dimension of the neocortex to enable the evaluation of migratory behaviour.

This experiment indicates that at any specific point in time, identity and/or localization of the neurons born in the knockout is shifted and delayed compared to wild type littermates. Whether this is due to a general delay in generation of progenitor cells from the subventricular

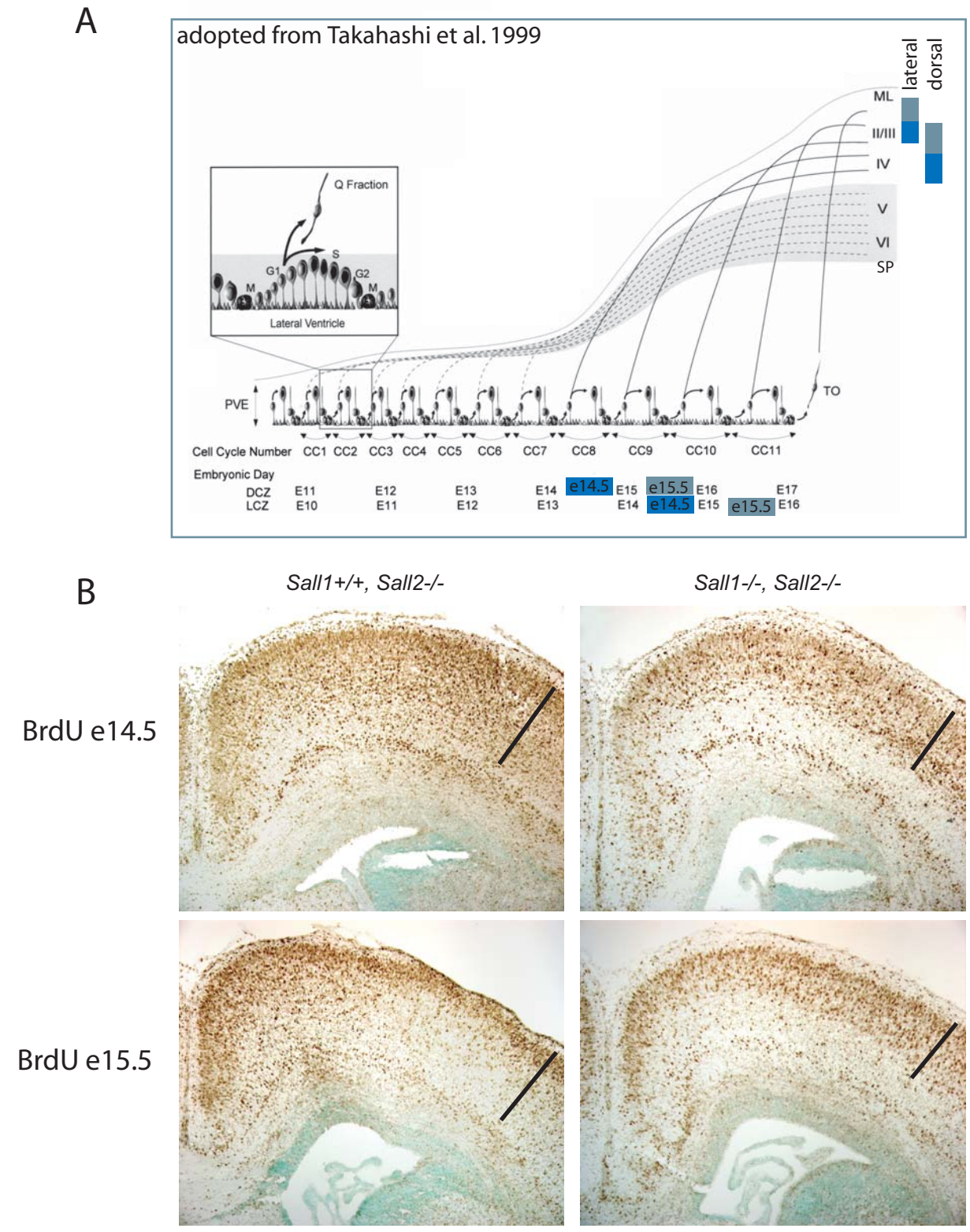


Figure 20: Neocortical neurogenesis is delayed in mice mutant for *Sall1* and *Sall2*.

A) Neurogenic precursors residing at the walls of the lateral ventricles undergo eleven rounds of cell division to give raise to neurons of the neocortex. During each cell cycle, its nuclei migrate to and from the ventricular walls undergoing mitosis directly at the ventricle. A fraction of daughter cells will initiate differentiation after each round of division and migrate into the cortex. Subsequent to preplate formation, all six cortical layers will migrate in and settle between cells of the preplate splitting it into subplate (SP) and marginal layer (ML). The neuronal layers in the cortex are differentiated in an inside-to-outside fashion with cells born first will settle in layer VI while subsequent cohorts of cells always migrate through the last layer defined and stop at the edge of the marginal layer. Neurogenesis is slightly more advanced in the lateral cortical zone (LCZ) compared to the most dorsal part (DCZ).

B) BrdU administration to dams pregnant for 14.5 and 15.5 days with subsequent analysis of the pups at e19.5. BrdU incorporates into cells undergoing DNA-synthesis and can be detected by immunohistochemistry. Labelled cells settle in earlier/deeper layers in mutant littermates compared to control animals. The effect is even more pronounced laterally and when taking the thinner cortex (bars) of mutant mice in account.

zone or to misspecification of neurons can not be answered in this experiment. Having shown earlier that *Sall* genes are expressed in the ventricular zone of the neocortex, not in the maturing neurons, it was tempting to speculate that the observed delay in neurogenesis could be explained by a reduction in the neuronal stem cell population. This would be in accordance with the histological observation that the cortex appears thinner in a knockout animal at all stages analyzed.

4.8. Immunohistochemical Analysis of the Subventricular Zone

To test for the rate of neurons generated in the subventricular zone, immunostainings against BrdU and phosphorylated histone 3 were performed on adjacent sections. Histone 3 is phosphorylated during mitosis and can be recognized with a specific antibody. In contrast, BrdU labels cells in S-phase. Figure 21 presents the results of such an experiment at e11.5 and e15.5. Dams were injected with BrdU 30 minutes prior to harvest of the embryos. At e11.5, no significant differences in the numbers of cells staining positive for phosphorylated histone 3 or BrdU incorporation were observed.

An alternative explanation for the reduced number of newly generated neurons can also be increased numbers of cells undergoing apoptosis. Apoptotic events were detected by immunostaining for cleaved caspase 3 without detecting any differences between knockout and control animals (figure 21). At e11.5, *Sall4* is still strongly expressed in the neocortex and might mask the phenotype of a *Sall1*/*Sall2* double knockout. Therefore, BrdU incorporation was also carried out at e15.5, when no expression of *Sall4* is detectable. Figure 21B shows comparable sections of the lateral ventricle after pulse labelling with BrdU for 30 minutes. For the mutant, stained sections contain a reduced number of cells in S-phase. Counting revealed a reduction by approximately 25%. Also, the width of the stripe harbouring the stained nuclei seems reduced. Due to the shorter duration of mitosis relative to S-phase, fewer nuclei stain positive for phosphorylated histone than for BrdU at any given time. The difference in the numbers of cells in M-phase is thus less pronounced but present.

Nuclei of progenitor cells undergo a phased migration to and away from the surface of the ventricles with mitosis occurring directly at the ventricular walls while nuclei in S-phase are found in the subventricular zone (Takahashi et al., 1999). Since BrdU was injected only 30 minutes prior to fixation, no migration will occur after S-phase and location of nuclei currently undergoing DNA synthesis can be addressed. In mouse mutants with disturbed cell cycle duration, uncoupling of cell localization and cell cycle phase can be seen

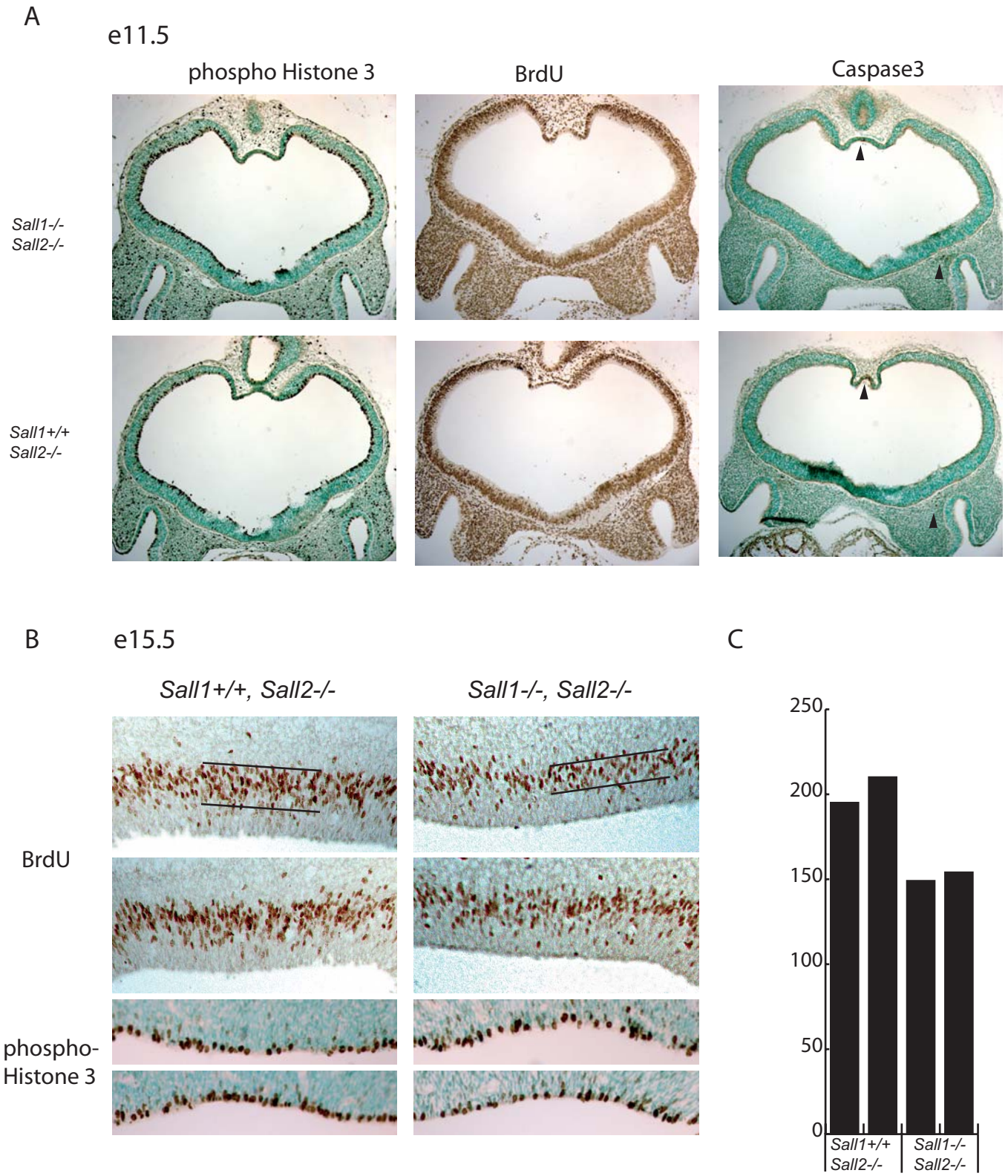


Figure 21: Cell cycle and apoptosis in mice mutant for *Sall1* and *Sall2.*
Immunohistochemistry was used to determine differences in cell cycle and apoptosis between *Sall1/2* double mutants and control animals. DNA synthesis (S-phase) was identified by injection of BrdU 30 minutes prior to collection of specimen. Mitosis was visualized with an antibody against phosphorylated histone 3, apoptosis through detection of cleaved caspase 3. Cell cycle and apoptosis seemed normal at e11.5, when *Sall4* is still expressed ubiquitously. It would be expected from a delay in cell cycle, that M- and S-phase are no longer localized to the compartments at and anway from the ventricular walls. This is not the case in *Sall1/2* double mutants arguing for a normal duration of cell cycle both at e11.5 and e15.5. Apoptosis was not more frequent than in wildtype (A, arrowhead) not observed at e15.5 and can thus not account for smaller numbers of cortical neurons.
B) reduction both in number of staining cells as well as in width of the stripe of BrdU incorborating cells in mice deleted of *Sall1* and *Sall2*
C) Quantification of neuronal progenitors in S-phase at e15.5 shown in B);

(Estivill-Torrus et al., 2002). In such a case, nuclei in M-phase would not be located exclusively at the ventricular wall and some S-phase nuclei will in turn be found at the ventricle. At e11.5 and e15.5, no ectopically positioned nuclei were seen.

Taking the observed thinner cortex in account, the result that less cells label BrdU positive allows two explanations. Either the number of progenitor cells is reduced or they undergo cell cycle in a slower fashion. The latter is unlikely since no progenitors were found to uncouple cell cycle and nuclear position. This makes a delay in cell cycle unlikely and argues for a reduced number of cells in cell cycle.

The time in embryogenesis and number of the cell cycle in the progenitor population has been tightly linked to the cortical fate of the resulting neuron (Takahashi et al., 1999). Here I show that mice mutant for *Sall1* and *Sall2* produce deeper neuronal layers compared to the wildtype at any given point in time and show reduced cell cycle activity. The differences in localization of neurons at e19.5 that were born e14.5 and e15.5 can be explained in two different ways. Either, a delay in neurogenesis without any effects on final cortical architecture upon completion of neurogenesis or a fate change in neuronal cohorts exiting the precursor zone. In both cases one might predict that late generated superficial layers lack cells at e19.5, either because they are not yet specified due to the delay or because neurons determine to the wrong the fate of neurons in a deeper layer.

4.9. Analysis of the Specification of Cortical Layers I-VI

Markers indicative for all cortical layers have been described previously (McEvilly et al., 2002). Genetic deletions are published leading to aberrations from the organization observed in wildtypic mice (Gupta et al., 2003). *Tbr1* deficient mice fail to develop layer six and the preplate (Hevner et al., 2001), in the naturally occurred mutation of the *reeler* mouse, all layers are specified but *Reelin*, the protein disrupted by the mutation, prevents normal migratory behaviour (Rice and Curran, 2001). This results in a failure of early born cortical neurons to invade and split the preplate resulting in a so-called superplate. Subsequent cohorts of neurons will also fail to pass their predecessors resulting in an outside-in pattern of the cortex (Caviness and Sidman, 1973).

To test for the presence and correct localization of cortical layers, *in situ* hybridization and immunostainings were utilized. Analysis was performed at e19.5, just before birth, to ensure full differentiation of the cortical layers.

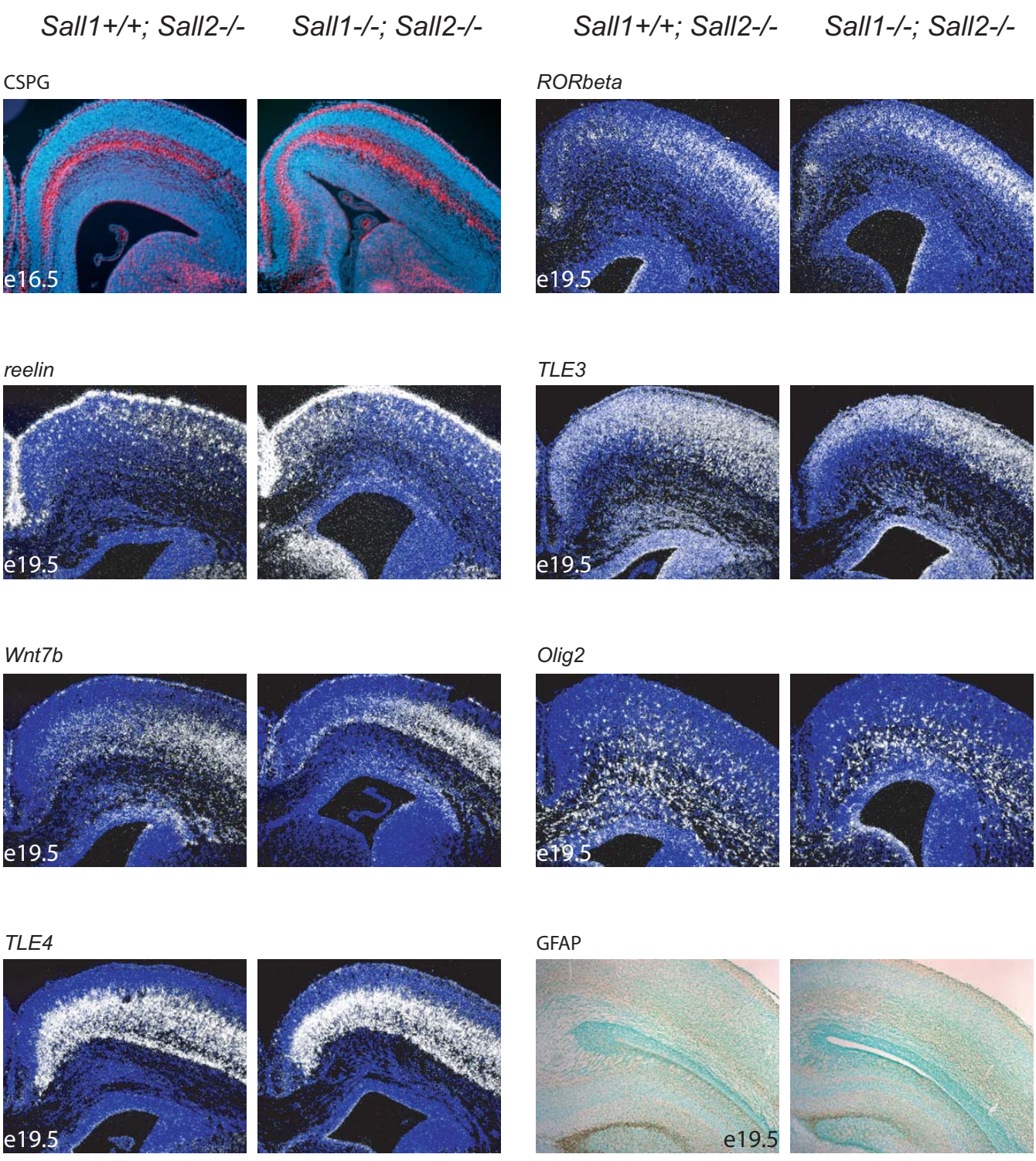


Figure 22: Molecular analysis of the effect of loss of Sall1 and Sall2 in neurogenesis.
In situ hybridization and immunostaining on littermates. Early generated layers show no or almost no change due to the loss of *Sall* alleles as visualized in staining for CSPG, *Reelin* and *Wnt7b*. In contrast, superficial layers are found closer to the surface (*RORbeta*) and thinner (*TLE3*). Oligodendrocytes are unaffected in a *Sall1/2* knockout, *GFAP* positive astrocyte are reduced in number. All images are coronal views, the midline is always left. CSPG appears red in immunofluorescence, DAPI is blue. For GFAP signal is brown, counterstain green.

Chondroitin sulphate proteoglycan (CSPG) is deposited in the extracellular matrix of the preplate and will be found in its derivatives, marginal layer and subplate (McEvilly et al., 2002). Figure 22A shows that cortical plate neurons successfully invade the preplate, splitting it in two layers as observed at e16.5. The more intense signal of CSPG in the mutant is probably due to the reduced number of intercalating neurons. An especially thick marginal layer can also be observed histologically at e16.5 (figure 11B). Presence of the marginal layer is also shown by *in situ* hybridization to *Reelin*. The signal in the knockout is again more intense than in the wildtype. The marginal layer thus consists at least of the cell number observed in wildtype neocortex.

The next layer that is specified is layer VI. It is marked by expression of *Wnt7*. Even though the cells in the knockout labelling positive for *Wnt7b* are more compacted, cell number appears approximately equal. As seen in wildtype, cells are positioned in the deepest portion of the cortical plate.

TLE4 is expressed in the subplate as well as in layer V. No significant differences can be seen between control and double mutant, the width of the layer is comparable.

Superficial to the expression of *TLE4*, *RORβ* is expressed in layer IV. While expression is similar in intensity and superficial relative to layer IV, expression reaches much closer to the surface indicating that layers positioned even later might be hypoplastic. This observation is underlined by *TLE3*, a marker for layers II to V. The *Sall1*/2 double mutant shows fewer cells staining positive for this marker.

Late in embryonic development, neurogenesis switches to gliogenesis and progenitor cells give rise to oligodendrocytes and astrocytes (Kessaris et al., 2001). To test if glia cells are present, *in situ* hybridization against *Olig2* was used to detect oligodendrocytes (Yokoo et al., 2004) and Immunostaining against GFAP revealed astrocyte presence. No significant differences were seen for oligodendrocytes originating from the ventral telencephalon (Qi et al., 2002) In contrast, astrocyte numbers were reduced (figure 22).

Taken together, analysis of all neuronal layers and glia cells with molecular markers indicates, that layers deposited later are more severely affected by loss of Sall activity than layers forming early. The astrocyte population is generated via differentiation of neuronal progenitors (Hunter and Hatten, 1995). Reduced numbers of astrocytes thus support the finding that the neurogenic precursor population is reduced. This strengthens the hypothesis that *Sall1*/*Sall2* deletion leads to a delay in neurogenesis. This delay occurs at the expense of late, more superficial cortical layers. Low numbers of astrocytes suggest that the progenitor pool is reduced in size and thus causes such retardation. Indications for a loss of progenitor

cells have already been obtained by BrdU labelling. In contrast to the observation in the *reeler* mouse and related phenotypes, cortical layers were found to be in the correct relative position to each other.

4.10. Expression of the Inserted lacZ Marker from the Genomic Locus of Sall1, Sall2 and Sall3

As described previously, *Sall1*, *Sall2* and *Sall3* are strongly expressed around the ventricular zones of the brain as well as in the olfactory bulb. Delayed neurogenesis might be explained by deletion of *Sall* genes normally expressed in neuronal progenitor and stem cell populations residing in the ventricular zone of the neocortex. In absence of good antibodies against these Sall proteins it was decided to study the expression of the markers replacing the coding sequence to gain additional information to the data obtained by *in situ* hybridization.

EGFP and *lacZ* had been introduced into the second exon of *Sall1*. Both markers reported expression. As seen in figure 23, A-F, Bluo-Gal staining faithfully recapitulates the expression pattern observed by *in situ* hybridization in the adult central nervous system. A very high number of cell nuclei in the olfactory bulb show expression of both the inserted *lacZ* gene in a mouse heterozygous for *Sall1-lacZ* as well as expression of *Sall1* mRNA in a wildtype mouse. In most other regions of the adult brain, expression is seen in a small fraction of cells giving rise to a speckled pattern. In the cerebellum, almost all nuclei located in the deep white matter of the cerebellum harbouring axon bundles stain. Only nuclei of glia cells are found in this region. The good correlation with the *in situ* is therefore especially conclusive here (figure 8C). Comparison of β-galactosidase and *in situ* data in the CNS and elsewhere strongly suggests that all cells expressing *Sall1* also transcribe the bicistronic transcript of EGFP and lacZ inserted into exon two. Expression of these markers can thus be used to locate *Sall1* expressing cells.

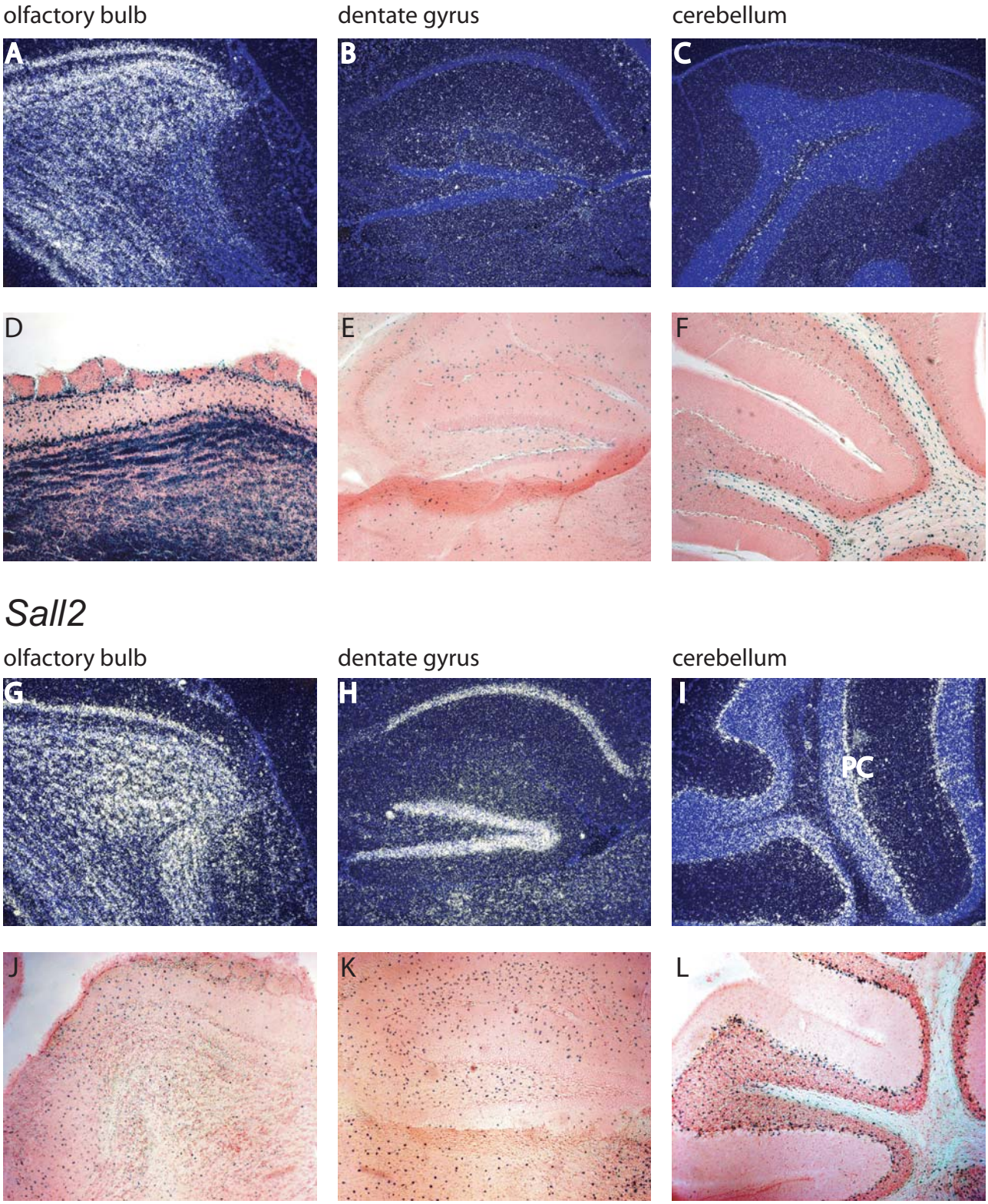


Figure 23: Expression of beta galactosidase from the *Sall1* and *Sall2* locus compared to in situ hybridization of the original transcript.

In situ hybridization of *Sall1* (A-C) and *Sall2* (G-I) in sagittal sections of adult brains is compared to beta galactosidase staining. Examples shown are olfactory bulb, hippocampus and cerebellum.

Sall1 in situ hybridization resulted in a speckled pattern of expression throughout the brain as well as high expression in the olfactory bulb interneurons (A). This expression is entirely recapitulated by beta galactosidase.

In contrast, *Sall2* expression is not faithfully reported by beta galactosidase. Expression of *Sall2* mRNA in the olfactory bulb is indistinguishable from *Sall1*, while neurons in the hippocampus express only *Sall2* (H). Expression in the cerebellum is specled but localized different from *Sall1*. Bluo-Gal staining only reports the speckled pattern observed (K;L) while interneurons of the olfactory bulb (J) and hippocampus (K) do not express lacZ. PC: purkinje cells;

To generate a *Sall2* knockout, the second exon had been replaced with a *lacZ* allele. Figure 23G-L illustrates, that β-galactosidase activity can not be visualized in all cells where *Sall2* transcript is detected. While the speckled pattern seen in the *in situ* and also for *Sall1* is reported by *Sall2-lacZ,* expression in the neurons of hippocampus and olfactory bulb is not recapitulated by the inserted cassette. In the hippocampus, expression does report as expected. It has to be concluded, that the genetically modified transcript does not fully recapitulate expression of *Sall2* and can not be used to study expression in detail. While part of the ih.gov/entrez/query.fcgi?cmd=Retrieve&*db*=PubMed&dopt=Citation&list_uids=11687494 </url></related-urls></urls></record></Cite></EndNote> (Blau et al., 2001; Weissman et al., 2001)

*Howe*ver, a definitive proof that some adulfor *Sall2* studied by *in situ* hybridization is reminiscent of *Sall1* expression both in olfactory bulb and the speckles seen throughout the brain (with exception of the cerebellum). Therefore, cells expressing *Sall1* might also express *Sall2*.

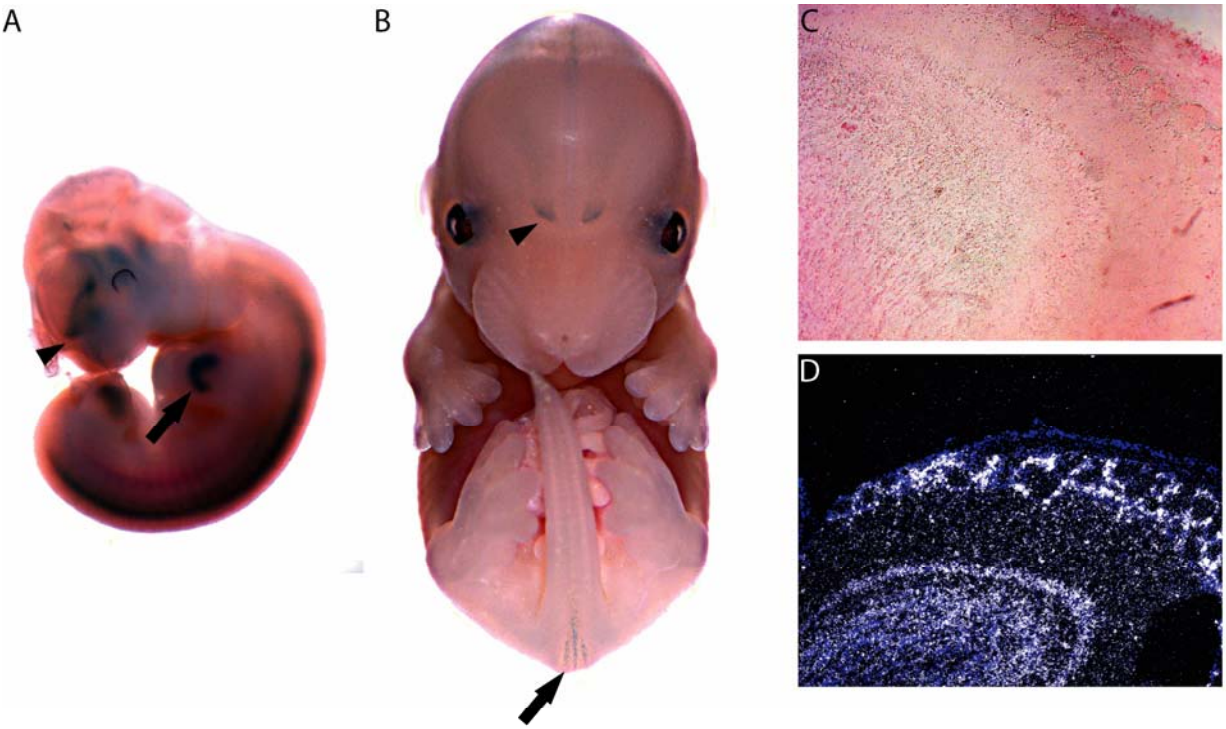


Figure 24: Expression of a beta galactosidase from the *Sall3* locus is very weak.
Whole mount beta galactosidase staining at e10.5 (A), e13.5 (B) and on a section of the adult olfactory bulb (C). While signal is still clearly detectable at e10.5, only weak staining is seen at e13.5 and no signal can be detected on sections of adult brain. This is probably due to lower levels of beta galactosidase expression compared to SALL3. The level of translation is too low to verify and study beta galactosidase as a reporter for *Sall3*. Expression of *Sall3* mRNA in the adult olfactory bulb is shown in D for comparison with C.
Observed expression includes olfactory bulb (A and B, arrowhead) and diencephalon at e10.5 as well as spinal cord e13.5 (B, arrow). Outside the CNS, strong expression can be observed in the growing limb bud at e10.5 (arrow). This expression overlaps with expression of *Sall1*.

Expression of *Sall3-lacZ* is very weak and can only be detected at embryonic stages in a whole mount staining with Bluo-Gal as shown in figure 24. While the expression observed is roughly reminiscent of transcription observed through *in situ* hybridization, staining was too weak to allow cellular expression analysis based on β-galactosidase as exemplified in a comparison of *in situ* and Bluo-Gal visualization in the olfactory bulb. Unlike *Sall2-lacZ*, β-galactosidase is expressed in the olfactory bulb from the *Sall3* locus (figure 24A, B, arrowhead). Nevertheless, barely any expression is visible on a cryostat section of an adult mouse heterozygous for *Sall3-lacZ.* Arrows point to expression in the limb (figure 24A) and spinal cord (B).

Since only *Sall1* showed strong β-galactosidase activity precisely reporting transcription of the endogenous genomic locus, further expression analysis was restricted to *Sall1*. Nevertheless, *in situ* data suggest a considerable overlap between *Sall1*, *Sall2* and *Sall3* expression also on cellular level for example in the olfactory bulb.

4.11. Precise Expression Pattern of Sall1 in the CNS

In situ data for *Sall1* had shown strong expression in the subventricular zones of the lateral ventricles. It had been reported previously that neuronal stem cells reside in this location giving rise to the interneurons in the olfactory bulb (Gritti et al., 2002). To test if *Sall1* is expressed in the SVZ and the differentiating neurons migrating into the olfactory bulbs, β-galactosidase staining was performed on lateral sections of the brain seven days and eight weeks after birth. Blue labelling is seen around the ventricle, in the rostral migratory stream (RMS) as well as in the olfactory bulb itself at P7 (figure 25A). This indicates that *Sall1* is indeed expressed in or around the undifferentiated neuronal progenitor cells migrating into the olfactory bulb to give rise to granular cells (GC) (Gritti et al., 2002). It an adult mouse only few cells that stain positive for Bluo-Gal are observed in the ventricular zone. As this was suggestive of an involvement of *Sall1* in undifferentiated neuronal stem cells or the stem cell niche also in adult brains, I next wanted to identify the cell types expressing *Sall1* in the different regions of the brain.

Fluorescent coimmunostaining with an antibody against β-galactosidase or EGFP, both indicative of *Sall1* expression, and markers for different cell types was performed. Consistent with *in situ* data, EGFP expression was detected in almost all cells of the ventricular and subventricular zone at e16.5 while cortical neurons (top half) do not express *Sall1-lacZ*. Cytoplasmic staining of GFP shows long extensions of the GFP positive cells. These radial

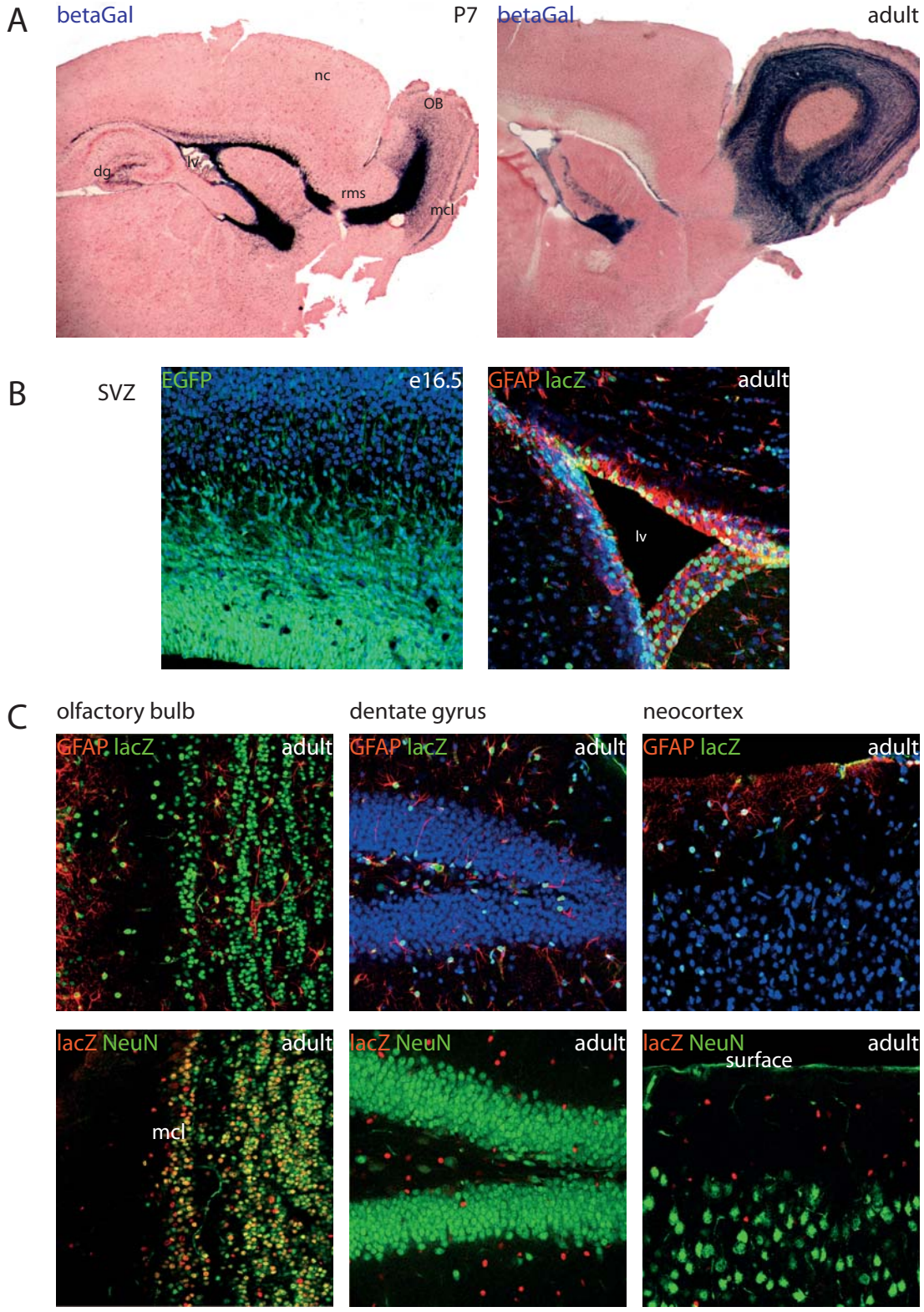


Figure 25: Expression analysis of *Sall1* at cellular resulution

Based on the assumption that the reporter cassette including an EGFP fusion as well as transcription of lacZ faithfully reports *Sall1* transcription, expression was analyzed on cryostat sections of brains heterozgous for *Sall1*.

In part A, an overview is shown at postnatal day seven (P7) as well as in an adult brain. In addition to the speckled expression observed, in the whole CNS including the neocortex (nc), higher density of cells presumably expressing *Sall1* can be found in the neurogenic regions of the dentate gyrus (dg) as well as the lateral ventricle (lv). The rostral migratory stream (rms) consists on newly generated neurons migrating from the walls of the lateral ventricle into the olfactory bulb (OB) and minimizes in adulthood. Expression is also present in the mitral cell layer (mcl) of the OB.

Within the young lateral ventricle, expression is seen in almost all cells of the subventricular zone including cells forming long radial extensions towards the cortical surface (B). Reduced expression in adult brain colocalizes in cells staining positive for glial fibrillary acidic protein (GFAP) in all regions analyzed (B,C). GFAP is a marker for the glial celltype of astrocytes. While immunostaining for lacZ and NeuN (neuronal nuclei) is exclusive in dentate gyrus and cerebellum, colabelling in yellow can be observed only in the olfactory bulb interneurons.

extensions are indicative of radial glia and migrating neuronal precursors (Coskun and Luskin, 2001). Radial glia serves as neuronal progenitor population during embryogenesis (Tamamaki et al., 2001). At this age, the SVZ still consists of several layers of neuronal precursors and undifferentiated neurons suggesting that the staining observed is indeed located in the precursor population. The high cell density makes interpretation of colabelling studies difficult. Therefore, cells positive for *Sall1-lacZ* were costained on coronal sections of an adult mouse with an antibody against glial fibrillary acidic protein (GFAP). GFAP is an intermediary filament expressed in astrocytes (Garcia et al., 2004), one type of glial cell. Several recent reports identified astrocytes as neuronal stem cells in mouse (Doetsch et al., 1999; Garcia et al., 2004). As illustrated in figure 25B, cells with nuclei labelling positive for β-galactosidase also express *GFAP* showing that *lacZ* is expressed in astrocytes of mice heterozygous for *Sall1-lacZ.*

New evidence suggests that GFAP positive neuronal stem cells are not restricted to the lateral ventricles but are also found in the olfactory bulb, dentate gyrus of the hippocampus and neocortex (Gritti et al., 2002; Song et al., 2002a; Song et al., 2002b). I thus analyzed these regions of the adult brain for their expression of *Sall1* and *GFAP*. As presented in figure 25C, immunoreactivity is located in cells positive for GFAP in all regions. To test whether *Sall1-lacZ* is also expressed in neurons, fluorescent colabellings with an antibody raised against neuronal nuclei (NeuN) were performed. *Sall1-lacZ* is expressed in neurons of the olfactory bulb (mitral cells and granular cells) but absent from all other neurons tested. Since neuronal progenitor cells are residing at the lateral ventricle while differentiating neurons are found in the cortex, the brain offers the possibility to illustrate that *Sall1* is expressed in neuronal precursors and astrocytes, identified as neuronal stem cells.

These studies confirmed that *Sall1* is expressed in neuronal progenitors and stem cells during development as well as in adult mice. Absence of olfactory bulb neurons and delay in cortical neurogenesis in a *Sall1/Sall2* double knockout is thus best explained by requirement of *Sall1* and *Sall2* in the neuronal progenitor population.

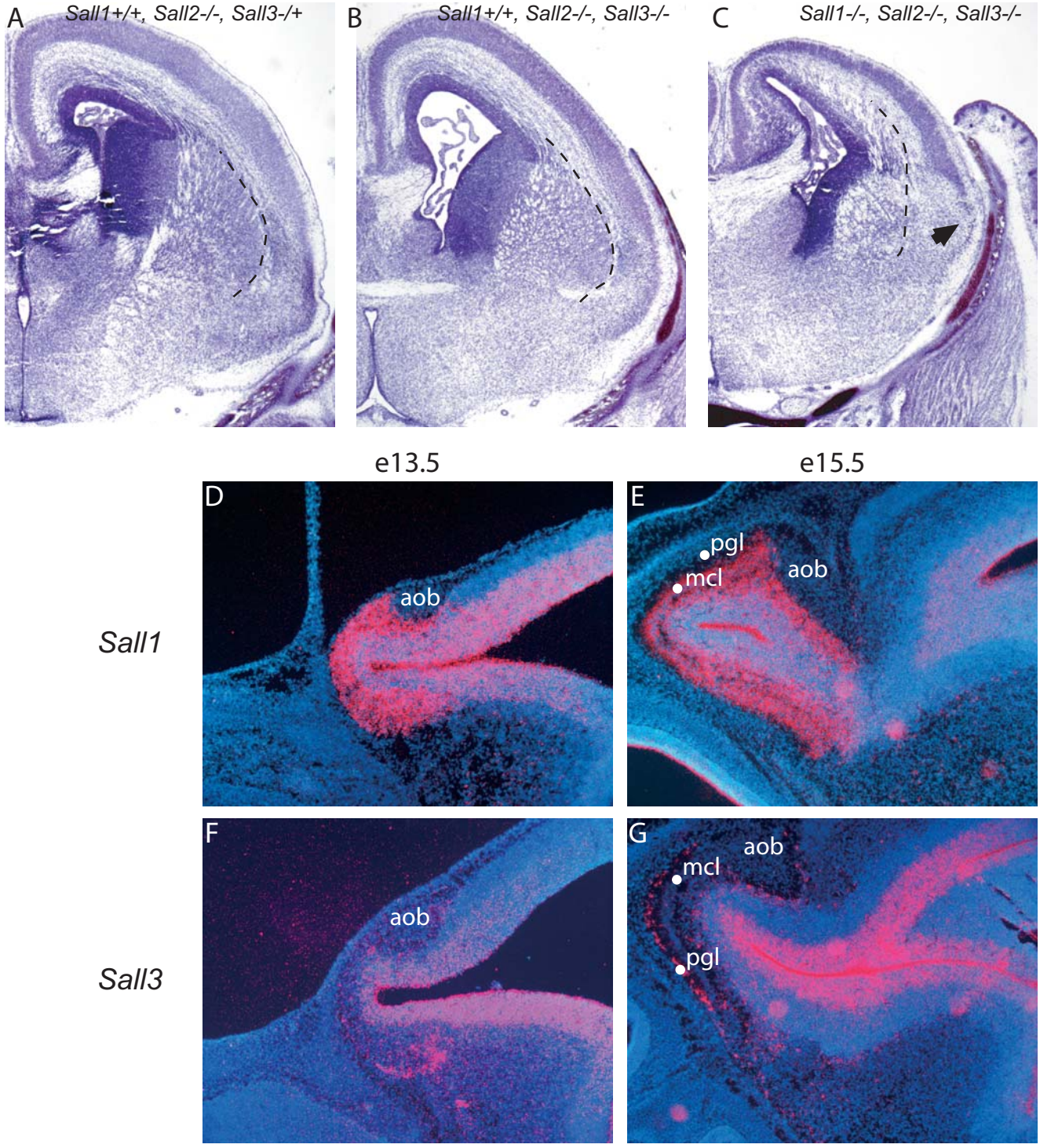


Figure 26: Effects of a *Sall1-3* triple knockout on brain architecture and comparison of *Sall1* and *Sall3* expression in forebrain and the developing olfactory bulb.

A-C) Comparative thionine staining to study the effects of a *Sall3* compared to a *Sall1* and *Sall3* deletion in a *Sall2* deficient background. Deletion of *Sall3* does not lead to major effects in the neocortex. Even a triple knockout of *Sall1-3* still retains a neocortical structure similar to a *Sall1-/-, Sall2-/-* deletion. In conrast, deletion of *Sall1-3* leads to almost complete absence of cortical structures from the ventral forebrain or paleocortex (arrow) where expression of *Sall1*, *Sall2* and *Sall3* overlap. Due to overall changes in brain morphology, sections look slightly different.

D-G) Sagittal sections through mouse telencephalon and olfactory bulb at e13.5 (D, F) and e15.5 (E, G). At e13.5, *Sall1* is expressed in the ventricular and subventricular zone as well as mitral cell progenitors, at e15.5 expression in the mitral cell layer (mcl) is already apparent. In contrast, *Sall3* is not expressed in the mitral cell progenitors, only in the ventricular and subventricular zone at e13.5. Later expression is detected in the periglomerular layer, a region not transcribing *Sall1* while *Sall3* is not detected in the *Sall1* expressing mitral cell layer. While the ventricular zone gives rise to inhibitory local interneurons at e13.5, mitral cells are generated in the subventricular zone. (aob: accessory olfactory bulb)

4.12. Synergistic Effects of Sall Proteins in Neurogenesis

Considerable overlap of expression patterns amongst murine *Sall* genes was described for the central nervous system. In the developing neocortex, expression of *Sall1* and *Sall2* is most prominent during stages, in which neurogenesis takes place (e.g. e13.5). *Sall3* transcript is considerably less abundant but present (figure 4). In contrast, *Sall1*, *Sall2* and *Sall3* are strongly expressed in the olfactory bulb all through embryogenesis (figure 3, 4). A deletion of *Sall1* and *Sall2* leads only to minor effects in the neocortex while in the olfactory bulb, functional mitral cells are completely absent. Can expression of *Sall3* account for these differences?

OB projection neurons (mitral and tufted cells) originate from the ventricular zone of the developing bulb, are not renewed postnatal and are already present at embryonic day 13 while interneurons (granular and periglomerular cells) are of subventricular origin and mostly generated late in embryonic development or after birth (see figure 25A), (bulfone1998). *Sall1* is expressed in (figure 26D-G) progenitors as well as interneurons and projection neurons at e13.5 as well as at e15.5. *Sall3* transcript, also found in the ventricular and subventricular zone of the olfactory bulb, is absent from developing neurons of the mitral cell layer at e13.5 and e15.5. Additional expression is observed in periglomerular cells at e15.5.

In accordance with the notion of redundancy, deletion of *Sall1* manifests in loss of mitral cells that express only *Sall1*. Loss of *Sall3* results in loss of cells in the outer plexiform and periglomerular layer where only *Sall3* is present and in mice deleted for *Sall1*, *Sall2* and *Sall3* all cell types are missing. This suggests that both, *Sall1* or *Sall3* alone, are sufficient for the formation of granular cell precursors and deletion of the genes leads to defects in cell types in which they are exclusively expressed.

In the dorsal telencephalon, additional deletion of *Sall3* does not result in a major enhancement of the phenotype observed in a *Sall1*/*Sall2* knockout while the ventral portion of the telencephalon lacks almost all cortical structures (figure 26C, arrow). *Sall3* is expressed at high levels in the ventral telencephalon at e13.5 and e15.5. Expression in the dorsal telencephalon is only seen at birth, a stage where neurogenesis in a wildtype mouse is almost completed.

5. Loss of Function of Sall1 and Sall3 Causes Oligodactyly

As mentioned earlier, *Sall1* and *Sall3* are expressed in developing limb buds (Kiefer et al., 2003; Ott et al., 2001). Mutations of *SALL* factors in humans cause limb deformations. It has been suggested that these deformations may be explained by haploinsufficiency (Kohlhase et al., 2002b; Kohlhase et al., 1998), gain of function or dominant negative effects (Kiefer et al., 2003). Clean deletions of *Sall* coding sequence through knockout technology in mouse enabled me to study effects of Sall proteins on limb development in a well defined loss of function situation.

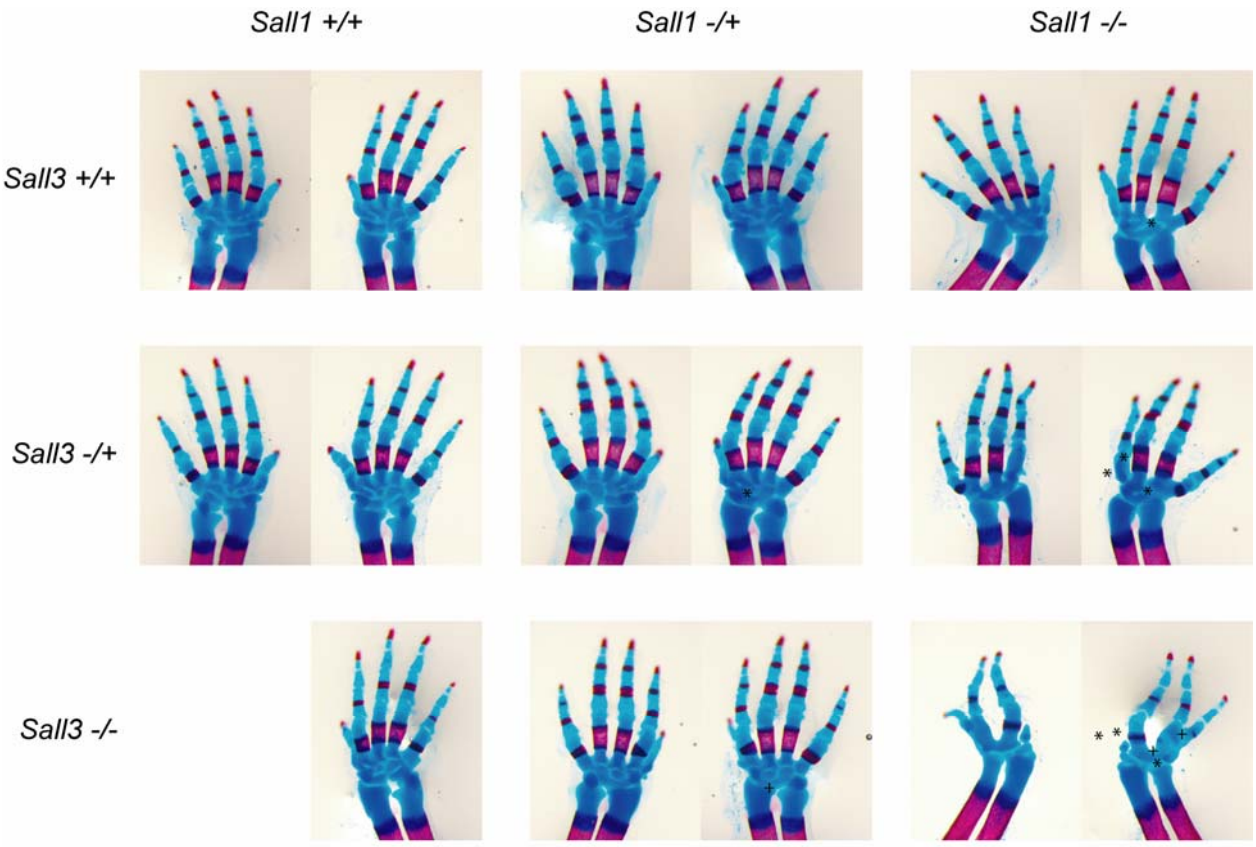


Figure 27: Quantitative effect of SALL loss in limbs.
Alizarin red/alzian blue staining on mice e19.5, bone stained in red, cartilidge in blue. Dorsal view of the forelimb autopod showing loss (*) and fusion (+) of elements. *Sall1* and *Sall3*, the only two *Sall* genes expressed at high levels in the developing forelimb at later stages show a quantitative effect upon allele loss. Defects get more and more pronounced the more *Sall* alleles are lost and even one single wt allele of *Sall1* or *Sall3* can significantly rescue the defects observed in a double knockout.
Phenotypes are generally loss of elements or whole phalanges causing oligodactyly. Polydactyly as described in humans heterozygous for *Sall1* or *Sall4* was never observed arguing that at least some human mutations are gain of function rather than of dominant negative nature.

Intercrossing animals heterozygous for *Sall1* and *Sall3* generated nine classes of progeny at e18.5. Embryos were collected and processed for alcian blue alicarin red staining. Figure 27 illustrates a clear genetic dosage effect. Developmental malformation correlates with progressive loss of *spalt-like* alleles that are expressed in the limb bud. Observed defects included loss and fusion of skeletal elements. Oligodactyly can be seen upon loss of at least three of the four alleles provided by *Sall1* and *Sall3*. Like for the human syndromes of TBS and DRRS, anterior phalanges were affected most by loss of *Sall* genes in mouse.

For humans, polydactyly as well as oligodactyly has been described. Polydactyly could never be observed in this loss of function study arguing that at least some of the dominant mutations described for TBS and Okihiro/Radial Ray Syndrome are due to gain of function effects, not of haploinsufficient or dominant negative nature. This is consistent with the finding that truncations of *sem-4* containing the first zinc finger act as gain of function proteins (Toker et al., 2003).

Additional loss of *Sall2* did not enhance the observed phenotype. Expression of *Sall4* in early limb buds had been reported previously (Kohlhase et al., 2002a), whether *Sall4* function is required for murine limb development was not addressed.

Discussion

The *spalt* family of transcription factors is highly conserved in all functional domains not only between different species but also amongst the four known SALL paralogs of human and mouse. Previous work in several model organisms including *Drosophila*, *C. elegans*, fish, chicken and mouse has established a crucial role for *Sall* genes in development. Nevertheless, no systematic genetic approach in vertebrates has been undertaken till date to analyze the effects of a loss of *Sall* genes. Due to functional redundancy amongst the four *Sall* family members, such an approach is required. I chose a systematic genetic approach in mouse to gain new insights in the role of SALL transcription factors in vertebrate development. Generation of knockouts for all four murine *Sall* genes enabled a combinatorial genetic loss of function analysis of several developmental processes.

1. Sall Transcription Factors are Expressed in Proliferative Cell Populations

Some information on the expression of *Sall* genes in mouse could be retrieved from previous studies. Typically, expression was studied by whole mount *in situ* hybridization or northern blot analysis. *Sall1* expression is described in brain, limb buds and kidney (Takasato et al., 2004). For *Sall2,* northern blot analysis (Kohlhase et al., 2000) showed high expression levels in adult brain, and low transcript abundance in lung, kidney and ovary. No expression was observed in adult heart, liver and testis. A study in 2003 (Sato et al., 2003) stated an overlap to *Sall1* expression.

Sall3 expression is described for neural tube as well as the midbrain-hindbrain boundary, the diecephalon and the ventral forebrain (Ott et al., 2001). Expression of *Sall3* in the palate and subsequent loss of function analysis demonstrated a function of *Sall3* in the palate (Parrish et al., 2004).

Cloning of *Sall4* was reported in 2002 (Kohlhase et al., 2002a) and expression analyzed by whole mount *in situ* hybridization. Transcript was detected in the embryo proper at e7.5, in the primitive streak, isthmus as well as in limb and tail buds at later stages. So far, no comparative analysis for all four *Sall* members has been undertaken to compare expression in space and time. In this study, systematic radioactive *in situ* hybridization was utilized to compare the expression of SALL factors in mammals. *Sall* genes code for highly conserved

transcription factors. The *in situ* hybridization protocol used included digestion of single stranded RNA post hybridization and thus prevented cross-hybridization to closely related transcripts.

Sall1 and *Sall3* expression is almost exclusively localized to the central nervous system (figures 3 and 4). Additional expression outside the CNS was observed in heart and kidney. Expression of *Sall1* and *Sall3* overlaps largely except for the neocortex where mainly *Sall1* is transcribed at stages of neurogenesis. Later stages also show expression of *Sall3* in neocortical regions. In the kidney, expression in the cortex is unique to *Sall1* while epithelializing mesenchyme also expresses *Sall2* and *Sall3*. Cells expressing transcripts of *Sall1* and *Sall3* within the CNS are found adjacent to ventricles. Strong expression in adult mice is restricted to the olfactory bulb. Utilizing the *lacZ* knockin into the *Sall1* locus revealed that expression was seen in almost all cells of the ventricular and subventricular zone of the lateral ventricles in the neocortex at e16.5 and restricts to astrocytes along the ventricular walls, hippocampus as well as intermingled within neurons in adult mice. Thus it seemed as if *Sall1* was expressed in cells capable to serve as neuronal progenitors. The only neuronal population found to express *Sall1* are olfactory bulb interneurons.

In contrast to that, transcripts for *Sall2* and *Sall4* are not restricted to the neuroectoderm. Rather, these genes are expressed in a fairly wide range of cell types in all germ layers. Detailed expression analysis revealed that *Sall4* is expressed from the 16 cell stage on, restricts to the inner cell mass of the blastocyst at the time of implantation and remains localized to ICM derivatives until decline of protein levels around day 9.5 (figure 9G). Expression of *Sall4* in later embryonic development is restricted to the liver (figure 4J), in adult mice, only *Sall4* was found in maturing oocytes.

Interestingly, the similarity of expression patterns correlates with protein homology amongst *Sall* genes in mouse and human (figure 3 and 4, compare to figure 1) *Sall1* and *Sall3* are the closest related members judged by protein similarity and their strongly overlapping expression suggests that ORF and promoter region duplicated together during evolution. Both genes express mainly in the central nervous system. *Sall2* and *Sall4* express in all germ layers with *Sall4* expressing in an earlier time window than *Sall2.*

As we and others show, *spalt-like* transcription factors are in general developmentally expressed and only few adult tissues retain *Sall* transcription. A systematic expression analysis of all four *Sall* genes revealed that embryonic transcription is found in cells well described to act as self-renewing, proliferating stem- and progenitor cell populations. Totipotent germ cells and cleavage stage embryos (Johnson and McConnell, 2004; Nichols et

al., 1998) express *Sall4*, pluripotent epiblast cells express *Sall2* and *4*, cells of the CNS ventricular zones express all *Sall* genes, especially *Sall1* and *Sall3*. Expression of *Sall1* in the neocortex peaks at the time of maximal progenitor population e12.5-e14.5. Adult expression of *Sall* transcription factors is found in astrocytes serving as neuronal stem cells (Doetsch et al., 1999), interneurons of the olfactory bulb that underlie constant renewal (Kaplan et al., 1985) as well as oocytes of the female germline. Upon differentiation of progenitor cell populations in the neocortex, *Sall1* expression declines (figure 25B). In cell culture, induced differentiation of ES cells leads to a reduction of *Sall4* expression (figure 9M).

Specific localization of *Sall* gene products to progenitor compartments was not previously recognized and served to evaluate the various phenotypes described herein. It lead to the hypothesis that SALL function is specifically required for progenitor cells. In the *Drosophila* sensory organ precursors, loss of *spalt* expression is a prerequisite for differentiation (de Celis et al., 1999) suggesting a function of *spalt* genes in maintenance of an undifferentiated state.

2. A Role for Sall Factors in Stem and Progenitor Cell Populations

A stem cell, by definition, is an undifferentiated cell that can produce daughter cells capable to either remain a stem cell (self-renewal) or enter a differentiation pathway. The signals that induce self renewal can be intrinsic, external (the stem cell niche) or both (Tada and Tada, 2001). Often, differentiating cells will undergo several additional rounds of cell division. They can not produce a self-renewing stem cell any more, rather all daughters will ultimately be terminally differentiated. Hence, these cells are referred to as progenitor cells. Besides self-renewal, proliferation, differentiation and tissue regeneration capacity, some definitions for stem cells include that stem cells must be able to self renew indefinitely, at least for the lifetime of the organism. The capacity to generate multiple different cell types is not included in the term stem cell. Rather a distinction between totipotent (the fertilized egg), pluripotent (ES cells) multipotent (radial glia cells) or cell type specific stem cell (e.g. in the skin) is made. However, the developmental potential of non-totipotent stem cells is under debate (Raff, 2003).

Till date, ES cells are by far the best characterized stem cell population due to well established culture conditions and because protocols for ES cell derivation are relatively simple and remain unchanged to the present day (Rastan and Robertson, 1985; Smith, 2001)

Intense studies during the last many years have lead to a well characterized system. Knockout analyses have supported such studies and identified *Oct4*, *Sox2* and *Nanog* as core players in a transcriptional circuitry required to maintain the developmental pluripotency of stem cells. These three factors are believed to have distinct but related roles (Chambers, 2004) in embryonic stem cell populations.

Deletion of *Oct4* leads to differentiation of ICM into trophoblast within four days in cultures of blastocyst outgrowths and immunosurgically exposed inner cell masses. This could be shown by morphological changes and upregulation of TROMA (Nichols et al., 1998). Subsequent studies of *Oct4* in ES cell culture showed that deletion of Oct4 results in transdifferentiation of ES cells into trophoblast. In contrast to that, overexpression of *Oct4* leads to differentiation into primitive endoderm. However, this was only shown by increased levels of *Gata4* expression measured by RT-PCR (Niwa et al., 2000). Oct4 is thus believed to play an important role in the separation of the inner cell mass lineages from the trophoblast.

In a knockout of *Sox2*, implantation sites are empty at e6.0 and only *H19* can be detected while the *Oct4*-expressing embryo proper is missing. Blastocyst outgrowth experiments revealed that the mutant ICM failed to grow. After immunosurgery, few giant cells were observed in ESC medium (Avilion et al., 2003) leading to the assumption that Sox2 is required for ICM lineage maintenance.

Nanog is only expressed in epiblast cells forming the embryo proper. Knockout ES cells differentiate into primitive endoderm (Mitsui et al., 2003) judged by the upregulation of multiple markers for primitive endoderm. Additionally it could be shown that ES cells ectopically expressing *Nanog* remain undifferentiated even in the absence of LIF (leukaemia inhibitory factor). LIF is required for ES cell culture to maintain cells in a pluripotent state. These observations placed Nanog directly into a cascade required to maintain pluripotency.

Common to all three studies is a developmental arrest immediately after implantation as well as the presence of trophoblast tissue accompanied with an absence of cells expressing the respective transcription factor. The *Oct4* phenotype establishes already at preimplantation stage. In an *Oct4* and *Sox2* knockout, giant cells replace ES cells upon immunosurgery. Similar effects are seen in the *Sall4* knockout. How might they relate?

2.1. Sall4 is Required for Inner Cell Mass Fate

Sall4 is expressed in preimplantation embryos and will remain expressed in inner cell mass and its derivatives until mid-gestation at decreasing levels. No other *Sall* genes are expressed that early and at such high levels (figures 3, 4 and 9). In absence of additional Sall factors, deletion of *Sall4* causes dramatic effects. Implantation sites of *Sall4* deficient blastocysts lack embryonic structures or cell types at e6.5 as it was observed for *Sox2-/-* and *Nanog-/-* embryos. Neither embryo proper nor primitive endoderm can be identified either histologically or by marker analysis such as *in situ* hybridization against *Oct4*. Implantation was verified by detection of *H19* and *Cited* transcript expressed in primary and secondary giant cells (figure 11). Unlike all prior analyses of periimplantation lethality (Avilion et al., 2003; Hanna et al., 2002; Mitsui et al., 2003; Nichols et al., 1998) in mouse, this study is able to receive positive confirmation for the genetic identity of the blastocysts leading to empty implantation sites due to the possibility of non-invasive genotyping of living e3.5 blastocysts prior to implantation (figure 8). Chimeric studies in vivo present evidence that *Sall4* requirement is cell autonomous in the epiblast (figure 16). The failure to rescue a *Sall4* deficient blastocyst by injection of labelled wildtype cells suggests that defects in primitive endoderm prevent survival of injected ES cells, since results of *in vitro* studies argue against a defect in the trophoblast (see also below).

In vivo studies were supplemented by the observation of blastocyst development *ex vivo* to study the dynamics leading to complete termination of development at e6.5. Culture of blastocysts in ES cell medium supported the finding that ICM cell lines but not trophoblast is affected by loss of *Sall4*.

Intense TROMA expression in a tight layer surrounding the ICM in a *Sall4-/-* blastocyst outgrowth suggests that ES cells can not maintain identity in culture. Instead of proliferation in a self-renewing fashion, cells differentiate and form a tight epithelium. Cultivated in ESC medium, wildtype ESCs maintain their identity and grow even on the surface of a blastocyst outgrowth (figure 12C) or in cell culture of ES cell lines.

Immunosurgery was used to separate ICM cells from trophoblast cells to then follow the developmental potential of ES and primitive endoderm (XEN) cells upon deletion of *Sall4*. If lineage commitment is already established by transcription factor expression, cells will continue growth as ES cells and XEN cells. Indeed, in ICM cultures of control blastocysts cells grew as ES and XEN cells. In contrast to that, no such cells could be observed from *Sall4-/-* ICMs after a few days in culture.

The absence of growing ES and XEN cells that would be expected in ICM culture confirms observations *in vivo* where no ICM derivatives could be observed at e6.5. I conclude that Sall4 is required for maintenance of both cell types of the ICM, ES cells and primitive endoderm. Trophectoderm replaced cells of the inner cell mass. Whether TS cells or giant cells are formed from the ICM depends on presence of Fgf-4 like it does for cell culture of wildtype TS cells indicating normal responsiveness.

The two cell types present in ICMs are interdependent. Thus, experiments were designed to distinguish effects in ES cells and primitive endoderm. Deletion of *Sall4* by Cre recombinase mediated excision in two independently established ES cell lines showed that *Sall4* is required specifically for ESCs in culture. Growth of ES cells in culture is cell autonomous. Therefore Sall4 is cell autonomously required for growth of ES cells in culture.

Establishment of XEN cells and subsequent targeting with Cre remains to be done in order to ask if XEN cells also depend on the presence of *Sall4* in a cell autonomous fashion. Absence of primitive endoderm cells in culture of ICMs however suggests that both ICM cell types require *Sall4* cell autonomously. *In vivo*, morula aggregation confirmed that *Sall4* is required cell autonomously in the epiblast.

Taken together, this study was able to show that Sall4 is required for maintenance of embryo proper and primitive endoderm in a cell autonomous fashion both *in vivo* and *in vitro*.

Cell lines of trophoblast and inner cell mass separate at the compacting morula stage (Rossant, 2004). The lineage separation is accompanied by differential expression of several transcription factors. *Oct4*, *Sox2* and *Nanog* are expressed in the inner cell mass as discussed earlier. *Cdx2* is the first known transcription factor to be specifically expressed in trophoblast cells. How the two separate lineages of ICM and trophoblast are established and different sets of transcription factors expressed is not known.

One hypothesis is that the location position of the cells serves as the initial signal for lineage determination and will induce the transcriptional differences. By the eight cell stage, blastomeres start compaction and transform from an apolar to a highly polarized cell leading to the first proto-epithelium of development (Fleming and Johnson, 1988). This polarization occurs at all levels of organization: cytocortical, cytoplasmic and cytoskeletal (Johnson and McConnell, 2004). If subsequent cell division occurs in a radial axis relative to the morula, it will result in a distal, polarized cell as well as a proximal cell in the center of the 16 cell morula. The centric cell will derive exclusively from the basolateral side of the dividing polar cell and thus be apolar again. Further cell divisions generate a 32 cell compacted morula with

apolar ICM cells in the center as well as highly polarized cells on the surface giving rise to the trophoblast lineage. This lineage commitment is still plastic in an experimental setup. Removal of all outer cells will initiate an additional round of cell division with subsequent differentiation of superficial cells into trophoblast (Johnson and Ziomek, 1983). Thus the location of each cell might induce a transcriptional profile generating different cell identities in the mouse morula.

Further support for this model comes from the observation that initial polarization and compaction is independent of protein synthesis (Kidder and McLachlin, 1985). Differential expression of transcription factors like *Sox2*, *Oct4* and *Nanog* in the inner cell mass as well as *Cdx2* in the trophoblast thus might be a consequence of cell localization and serve to reinforce differences to a developmental commitment. In accordance to a rudimentary ability of preimplantation embryos to establish the two cell lineages prior to transcription factor expression, no knockout of transcription factors described till date interferes with initial lineage separation. The most severe phenotype described has been found in *Cdx2* knockouts. *Cdx2-/-* morulas can still transform into a blastula like structure, a process that requires polarization and epithelialization of trophoblast cells (Strumpf et al., 2005). Nevertheless, these blastulas fail to implant, demonstrating that no full trophoblast commitment can be reached. Deleting ICM specific factors like *Oct4* (Nichols et al., 1998), *Sox2* (Avilion et al., 2003) and *Nanog* (Mitsui et al., 2003) allows developmental progression until and including implantation.

This model of position-dependent lineage commitment has interesting implications for the interpretation of the results presented on *Sall4*. I show that ES cells and primitive endoderm cells fail to self-renew in a *Sall4-/-* blastocyst outgrowth experiment. Nevertheless, TROMA does not label all cells of the inner cell mass but rather is only detected on cells at the surface indicating that superficial cells are epithelialized. The presented model might explain the absence of TROMA staining in the center. While all ICM cells loose the ability to self-renew, differentiation and epithelialization of ES cells might require the exposure to the surface. Hence, only superficial cells in a *Sall4-/-* blastocyst outgrowth express TROMA while the whole ICM differentiates upon immunosurgery. As shown before, TROMA positive cells detach from a blastocyst outgrowth. This will ultimately lead to the complete loss of ES cells. Therefore, the interpretation that Sall4 is required for ICM maintenance is in accordance with the present understanding of cell lineage determination and maintenance in early embryos.

The definition stem cell includes the capacity to proliferate, self renew and differentiate. Growth of TS cells as well as subsequent differentiation into giant cells does not fulfil requirements for embryonic stem cells since both cell types are extraembryonic and not a natural progeny of ES cells. Therefore, ES cell fate can not be maintained in absence of Sall4 function.

2.2. Explanations for Appearance of Trophoblast in Sall4-/- ICMs

As discussed earlier, giant cells appear upon culture of ICMs in ES cell medium while TS cell medium leads to the formation of trophoblast stem cells in a *Sall4* knockout but not in wildtype. How can the appearance of trophoblast be explained?

Two alternative explanations can account for the presence of trophoblast lineage in culture of ICM. A transdifferentiation model suggests fate change of ICM cells upon loss of *Sall4*. The maintenance model suggests that endogenous plasticity of ICM cells allows for TS cell growth from ICMs while ES and XEN lineage are not maintained.

As discussed earlier, cell lineages of trophoblast and inner cell mass separate during morula compaction but final lineage commitment based on differential expression of transcription factors occurs at some point in the blastula in wildtype (Johnson and McConnell, 2004). Several reports find plasticity until the 32 cell stage (Handyside, 1978; Rossant and Lis, 1979) while trophoblast formation was seen even from ICMs of blastocysts collected e3.5 and cultured for 24 hours (Hogan and Tilly, 1978) by other groups. Inefficient killing of trophoblast cells by the immunosurgery procedure is unlikely as has been discussed in earlier studies utilizing immunosurgery (Hogan and Tilly, 1978). Blastocysts used in this study contained approximately 64 cells. The interpretation of the results obtained from cultivating ICMs depends on an assessment of the plasticity in the ICM.

According to a transdifferentiation model, Sall4 is required for the identity of the inner cell mass and acts as a lineage switch between trophoblast and ICM. *Sall4* is expressed at higher levels in the inner cell mass compared to the trophoblast lineage and will restrict to ICM derivates after implantation. In a *Sall4* knockout, ICM lineage commitment can not be maintained and cells transdifferentiate into trophoblast upon surface exposure as they would do prior to the 32 cell stage. Loss of Oct4 is thought to have such an effect. However, to prove this model, it has to be formally shown that trophoblast cells arose from a population that was or should be fully committed to ICM fate. Whether lineages in a 64 cell blastocyst are already

committed is controversial (Hogan and Tilly, 1978; Rossant and Lis, 1979). *Oct4* depletion from an established ES cell line lead to differentiation into trophoblast and thus fulfils these criteria. Also, differentiation occurs more rapidly than in case of *Sall4*.

A maintenance model assumes that lineage commitment is not complete at the 64 cell blastocyst stage (Hogan and Tilly, 1978). ICM cells might still be able to give rise to trophoblast, even though under normal conditions ES cells and XEN cells will grow. Supportive to that is the notion that ICM will also give rise to totipotent germline, ultimately able to form trophoblast. If cells are not able any more to grow as ES or XEN cells, the remaining plasticity to generate trophoblast might divert some cells. Trophoblast is not affected by loss of Sall4 and will thus be able to grow.

To distinguish between the two models, a transdifferentiation model versus a plasticity based model, it is crucial to assess the developmental potential of the cells undergoing the switch in cell types. To do so, one might make use of the floxed *Sall4* ES cell line that has been described herein. Deletion of *Sall4* in the presence of TS cell medium supplemented with FGF4 can challenge a transdifferentiation model because the plasticity of the floxed ES cell line is restricted. Murine ES cells never differentiate into trophoblast cells even under TS cell conditions. Subsequent to deletion of *Sall4* by Cre transfection, presence of TS cells would thus argue for a transdifferentiation model. For *Oct4*, such a switch could be established.

In conclusion, cell culture with *Sall4* deficient blastocyst outgrowth, inner cell mass and ES cells served to establish that cells of the ICM depend on *Sall4* expression to maintain their self-renewal capacity. Expression of TROMA in superficial cells of the inner cell mass together with absence of ES cells and XEN cells in culture of ICMs supports a model suggesting that *Sall4* is required for commitment of ICM cells and only trophoblast cells can grow upon loss of Sall4.

2.3. What Could be a Molecular Role for Sall4 in ES Cells?

Chromatin co-immunoprecipitation with *Oct4*, *Sox2* and *Nanog* transcription factors from human ES cells and subsequent microarray analysis (Boyer et al., 2005) identified binding sites of all three transcription factors in the promoter region of several other transcription factors. Amongst them, Oct4, Nanog and Sox2 each bind to their own promoter as well as to the promoter of the other two genes. It is assumed, that mutual regulation of these three transcription factors serves as a core regulatory circuitry to define stemness.

Recent studies come to the same conclusion for the murine orthologous for *Oct4*, *Sox2* and *Nanog* (Robson, 2005). While a regulatory loop can maintain expression of these three stem cell factors, it would be expected that they also employ additional transcriptional regulators to maintain, proliferation, self-renewal and pluripotency. Interestingly, human Oct4, Sox2 and Nanog bind to the *Sall1* promoter (Boyer et al., 2005), while all three mouse orthologs bind to the genomic locus of *Sall4* (Robson, 2005). This suggests that *Sall* genes -*Sall4* in mouse- are directly downstream of a core circuitry defining the properties of stem cells.

Here I present data suggesting that ES cells loose the capacity to self renew in *Sall4* knockout embryos and ES cell lines. It is therefore tempting to speculate that *Oct4*, *Sox2* and *Nanog* act in part through *Sall4* to maintain stem cell identity and commitment. Down regulation of Sall4 protein levels as presented in a differentiation assay with ES cells might thus be an indirect effect caused by reduction in *Oct4*, *Sox2* or *Nanog* expression. Based on present expression of *Oct4* and *Nanog* in expanded *Sall4-/-* blastocysts it can be concluded that *Sall4* is downstream of these two factors. An alternative explanation would be that Sall4 also bound the promoter of *Oct4*, *Sox2* and *Nanog* and is part of a transcriptional network controlling stemness. Loss of *Sall4* might then not be sufficient to inactivate expression of *Oct4* and *Nanog.* The ability to activate expression has been shown for *spalt* family transcription factors (Onai et al., 2004), though several studies suggest that Spalt proteins act as transcriptional repressors.

What might be the molecular role of *Sall* genes in progenitors, more specifically *Sall4* in ES cells? Drosophila *spalt* interacts genetically with *polycomb* complexes, important to maintain a transcriptionally silenced state (Landecker et al., 1994) arguing for a role of *spalt* related genes in general transcriptional silencing. Genetic deletion of members of the polycomb complex in mouse have been shown to lead to periimplantation lethality and growth defects in the epiblast (Donohoe et al., 1999; O'Carroll et al., 2001) Studies also establish protein interaction of vertebrate SALL factors to histone deacetylase as well as its localization to heterochromatin (Kiefer et al., 2002; Netzer et al., 2001).

While the underlying process by which a stem cell chooses to self renew or differentiate must be molecular in nature, the pathways remain elusive. A critical and under-investigated area is chromatin organization and epigenetic control in ES cells (Smith, 2001). In part, historical reasons contribute because the field of stem cell research and epigenetic research developed independently. With increasing knowledge, both fields have started to converge recently (Tada and Tada, 2001).

Two important means of transcriptional silencing, histone acetylation and DNA methylation, could be functionally linked in the last years (Jones et al., 1998; Nan et al., 1998). Especially DNA methylation during early embryonic development received much attention. DNA of somatic cells is methylated in a distinctive pattern that is stably inherited by the action of Dnmt1, a maintenance enzyme that methylates hemimethylated CpG dinucleotides in the nascent strand of DNA after DNA replication (Bestor, 1992). In contrast, developing gametes as well as early embryos undergo almost complete demethylation. It is achieved partially actively but also passively through tethering of Dnmt1 to the cytoplasm followed by a demethylation through cell division in preimplantation embryos. In both cases, gametogenesis as well as preimplantation development, erasure of methylation marks is believed to reset the genome and will be followed by de novo methylation (Li, 2002; Rideout et al., 2001). Mature gametes are again highly methylated (Rideout et al., 2001). In addition, especially the female pronucleus also contains relatively deacetylated histone as a means of transcriptional repression (Jackson-Grusby et al., 2001). After the second round of demethylation in embryogenesis, *de novo* methylation can be observed in the epiblast from implantation stage onwards (Li, 2002; Santos and Dean, 2004), while the trophoblast will remain demethylated. New methylation marks are believed to establish cell lineage restriction. ES cells in culture are generally hypomethylated even though these cells do possess a *de novo* methylation activity (Tada and Tada, 2001) and e.g. imprinted genes will remain methylated. Deletion of Dnmt1 does not affect ES cell growth.

Linked to the methylation status of the genome, both X chromosomes are active in oocytes, female preimplantation embryos and female ES cell lines and serve as a model to understand chromatin organization during periimplantation stages. Two active X chromosomes are a disadvantage for propagation of ES cells and thereby one reason why most mouse ES cell lines are derived from male blastocysts (Rastan and Robertson, 1985). If stemness is a state of general transcriptional permissiveness (Smith, 2001) what ensures stable inheritance of that state in absence of common mechanisms for transcriptional silencing?

Methylation is believed to be typically the last step in a silencing cascade functionally linked to histone deacetylation (Jones et al., 1998; Nan et al., 1998). Unlike methylation, inhibition of transcription through histone deacetylation plays an important role in ES cells (Rideout et al., 2001).

As mentioned earlier, SALL proteins bind to histone deacetylase through their N-terminal Zn-finger and localize to heterochromatin. One could envision a scenario whereby absence of *Sall4* at a time where methylation marks are erased will lead to loss of

transcriptional repression in stem cells and thus the failure to maintain an undifferentiated state. In accordance to this, *Sall4* is not required in early cleavage stages when methylation is still present but absence of *Sall4* in demethylated embryos and ES cells is deleterious. Loss of Sall4 also leads to developmental arrest during oogenesis, the other period of global de- and remethylation. This model could explain the paradox that no phenotype was observed until expanded blastocysts but ES cells fail to grow in absence of Sall4. Post implantation, the repressed state of many loci in the epiblast will be re-established by DNA methylation and can be visualized with a specific antibody. In contrast, trophoblast cells do not methylate DNA until e7.5. Whether Sall4 is required for silencing specific promoters or involved in a general silencing complex remains to be established.

To test if Sall4 is part of a general silencing-machinery essential for periimplantation embryos it would be interesting to purify Sall4 with its interaction partners from ES cells. Complementary to that, inducible deletion of *Sall4* from an ES cell line can be achieved by stable expression of Cre fused to a modified estrogen binding domain (Metzger et al., 2005). Subsequent transcriptional profiling by microarray technology can address changes in the transcriptional landscape and would elucidate the molecular role of Sall4 in ES cell maintenance. These two complementary experiments will also address, if Sall4 is involved in a general mechanism of transcriptional silencing in ES cells.

3. Sall Factors in Adult Stem Cells

Adult stem cells are undifferentiated cells capable of proliferation, self-renewal, production of a large number of differentiated progeny and regeneration of tissues (Blau et al., 2001). Especially the latter feature has caught interest recently. A general distinction is made between embryonic stem cells (ES cells) that can give rise to all cell types and tissues of the embryo including the germline (Bradley et al., 1984) and are therefore called pluripotent as opposed to adult stem cells residing in many or most tissues. Adult stem cells are generally believed to be more linage committed though this view has been challenged recently through a number of experiments (Blau et al., 2001; Weissman et al., 2001)

However, a definitive proof that some adult stem cells retained pluripotency is challenging due to the technical difficulties of identification of cell fate before, during and after a possible switch. Also, results could not always be reproduced (Raff, 2003). The

interpretation of the presence of trophoblast cells upon culture of *Sall4-/-* ICMs discussed here suffers from similar ambiguities.

If mesenchymal stem cells of adult rodent bone marrow, muscle or brain are cultured at low density for many passages, a population of cells might arise that shares many properties with ES cells (Jiang et al., 2002a; Jiang et al., 2002b) and are termed multipotent adult progenitor cells. Such cells will contribute to all mouse organs when injected into blastocysts. It remains to be seen whether these cells pre-exist at very low density or if they gradually acquire ES cell like properties due to culture conditions in the presence of LIF. Multipotent neuronal stem cells have been reported especially for the dentate gyrus as well as the lateral ventricles.

Are SALL factors required for maintenance of adult stem cell populations? The genetic study of a requirement for *Sall* genes is complicated by lethality prior to adulthood in case of conventional knockout analysis.

Expression analysis for *Sall* transcription factors suggested an intriguing possibility, namely a general function of SALL factors in stem and progenitor cell compartments. To test this hypothesis, the effect of a loss of *Sall* genes in domains of late expression was studied.

3.1. Redundancy of Sall1, Sall2 and Sall3 in Olfactory Bulb Development

Mammalian olfactory bulbs (OB) develop around a rostral extension of the lateral ventricles in the forebrain. Projections of the olfactory receptor neurons converge in glomeruli on the surface of the olfactory bulbs. Dendritic extensions of subcortical mitral and tufted cells reach the glomeruli and form synaptic connections with the axons of the sensory neurons of the olfactory epithelium (OE). Most of the neurons present in the olfactory bulbs are inhibitory interneurons in the granular cell layer. In rodents, these interneurons are constantly replaced. Less than 10% of the granular cells born in young adult rats survive 21 months (Kaplan et al., 1985). *Sall1*, *Sall2* and *Sall3* are expressed in the developing olfactory bulbs and remain expressed in adulthood. An analysis thus included all three genes since *Sall4* is only expressed early in development and caused lethality prior to olfactory bulb development.

In a comparative histological analysis of the olfactory bulb at e18.5 I show severe effects of *Sall* factor deletions on OB development: *Sall2* deletion has no impact on the histology of the OB. Additional deletion of *Sall1* results in a disturbance of the mitral cell layer (MCL) while additional deletion of *Sall3* leads to cell loss in the outer plexiform layer.

In a triple knockout of *Sall1-3*, no specific olfactory bulb structures can be recognized. *In situ* hybridization against *Reelin* revealed only a scattered pattern as opposed to a dense ring structure in the olfactory bulbs of *Sall1*/*Sall2* knockout mice. Scattered *Reelin* expression might be due to presence of tufted cells (Bulfone et al., 1998) or remaining mitral cells. Granular and periglomerular cells are less affected by loss of *Sall1* as could be shown through *in situ* hybridization against *Gad67*. Retrograde DiI tracing of axonal projections of mitral and tufted cells showed an absence of projections to the olfactory cortex indicating anosmia in *Sall1*/*Sall2* deficient mice. Previous studies (Belluscio et al., 1998; Bulfone et al., 1998; Hongo et al., 2000) show, that anosmic mice fail to find their mothers nipples and die due to the inability to suckle milk. Thus postnatal lethality in *Sall1*/*Sall2*, possibly also *Sall2*/*Sall3* knockout mice may be due to anosmia. It is likely that proliferation of progenitors at the time where mitral cells are born is reduced. BrdU labelling as well as immunohistochemistry for phosphorylated histone 3 at e12.5 will clarify this point.

3.2. Loss of Sall1 & Sall2 Leads to Defects in Cortical Neurogenesis

Can neurogenic defects be attributed to defects in the progenitor cell population? For the developing neocortex, it seems likely. The progenitor population is locally separated from maturing and mature neurons and *Sall* genes are expressed in the progenitor population, but absent from their progeny. In adult mice, *Sall1* is expressed in astrocytes (figure 25B, C), shown to function as adult neuronal stem cells.

Sall1 and *Sall2* are the two orthologs of the Spalt family, prominently expressed in the subventricular zone of the developing neocortex between e12.5 and e19.5. The neocortex of double knockouts of *Sall1* and *Sall2* is hypomorphic as shown by histological analysis e16.5 until e19.5. Overall, organization of the neocortex remained normal but the thickness and appearance of the cortex resembled that of a wildtype cortex approximately 24 hours earlier. BrdU labelling of cells undergoing DNA synthesis showed that serial differentiation of the cortical layers is also delayed to a similar extend. The observed delay could not be attributed to apoptosis but rather is due to 25% less cell divisions occurring in the ventricular zone of the lateral ventricles (figure 21C), the expression domain of *Sall1* and *Sall2*.

Delay in neurogenesis does not lead to a disturbance in cortical layering. Rather, all layers are present and located in the correct relative position (figure 22) to each other. In accordance with a neurogenic delay, superficial layers of the neocortex that develop late in embryogenesis are most severely affected by loss of *spalt-like* genes in the ventricular zone.

Two mechanisms could account for a reduced number of neuronal progenitors originating from the subventricular zone in the absence of *Sall1* and *Sall2*. Either the cell cycle of neuronal progenitors is extended or the number of progenitor cells is reduced. For the Pax6 mouse mutant, a cell cycle delay has been shown. Nuclei of neuronal progenitors migrate to and away from the ventricular walls (Takahashi et al., 1999), (figure 20A). In a murine knockout of *Pax6* extended cell cycle periods result in an uncoupling of nuclear migration and cell cycle (Estivill-Torrus et al., 2002). Upon BrdU administration for 30 minutes, cell nuclei in S-phase are found also on the ventricular walls in a *Pax6* mutant and mitotic nulclei are found away from the ventricle in turn. In the *Sall* mutants, no mislocalization of S-phase nuclei or mitotic nuclei was accompanied with the 25% reduction in cells incorporating BrdU suggesting that not cell cycle duration but rather population size may account for delayed neurogenesis.

Whether daughter cells remain in the ventricular zone of the developing neocortex as progenitors or exit to the neocortex has to be tightly regulated to maintain a correct progenitor population. Over-expression of stabilized β-catenin leads to an increase in the number of cells re-entering the cell cycle after mitosis leading to an enlarged progenitor population and thus enlarged brains (Chenn and Walsh, 2002; Zechner et al., 2003). On the other hand, deletion of β-catenin from the brain leads to loss of progenitors resulting in a very thin neocortex (Zechner et al., 2003). Similar effects have also been described for Pax6 (Estivill-Torrus et al., 2002). In case of the *Sall1*/*Sall2* double knockout, a shift from re-entering towards exit of the cell cycle is therefore the most likely interpretation for the reduced numbers of cells found to be in S-phase. *Sall1* and *Sall2* might be required for the self-renewal of the stem/progenitor cell population in the neocortex.

Neurogenesis does proceed in the *Sall1* and *Sall2* double knockout and brains reach almost full size at birth. This may be due to expression of *Sall4* in the ventricular zone of the forebrain until approximately e12.5. *Sall3* is expressed only very weakly in the neocortex of the telencephalon at e13.5. In accordance to that, defects in a *Sall1*/*Sall2*/*Sall3* triple knockout are only mildly enhanced compared to the *Sall1*/*Sall2* double mutant. In contrast, the paleocortex of the ventral forebrain expressing all three *Sall* genes is strongly affected by the additional deletion of *Sall3* (figure 26C, arrow). Here, almost no cortical cells can be identified further strengthening the hypothesis that SALL transcription factors share redundant functions where coexpressed.

3.3. Could Progenitor Loss Account for Other Late Phenotypes?

Oligodactyly might also be explained by progressive loss of proliferative cells. According to the progress zone model (Wolpert, 2002; Wolpert et al., 1979), limb elements are defined in a proximal to distal sequence. Successive loss of dividing progenitor cells in the limb bud would thus lead to cell loss in the most distal part. Anterior posterior patterning in the limb is achieved by a morphogen gradient (*Shh*) with anterior elements being established at lowest Shh concentration. If fewer cells are present in a primordium to be patterned by Shh, it could be envisioned that anterior elements, normally receiving lowest Shh levels, are most dramatically affected. However, direct interference with cell signalling in the limb bud could also explain the observed loss of digital and wrist elements. It would be interesting to quantify size and proliferative index of limb primordium at stages of distal patterning.

Further, metanephric mesenchyme expresses *Sall1* during the mesenchymal-epithelial transition (Nishinakamura et al., 2001), during a process that has been described as a stem cell phase in renal development (Bard J. B. L., 1996). As seen in figures 3 and 4, *Sall2* and *Sall3* are also present during mesenchymal to epithelial transition. In a *Sall1* knockout, mesenchyme is still competent to condense and undergo transition but kidneys are severely hypoplastic. The phenotype was interpreted as a failure of the ureteric bud to invade the mesenchyme (Nishinakamura et al., 2001). However, it is also consistent with the idea of a role for SALL transcription factors in progenitor population maintenance. The metanephric mesenchyme already consists of fewer cells at e11.5 and fails to grow upon induction, yet some nephrons do develop. Consistent with the notion of redundancy, kidneys are severely hypoplastic upon loss of *Sall1*, the only *Sall* member expressed in the cortex of the kidney. In mice deficient for *Sall1*, *Sall2* and *Sall3*, renal tissue is completely absent, possibly due to complete loss of renal progenitors.

The maturing oocyte can also be regarded as a totipotent stem cell, even able to reprogram somatic nuclei to totipotency upon nuclear transfer and is used in cloning (Wilmut et al., 1997). Female mice with an oocyte-specific deletion of the only expressed *Sall* gene, *Sall4*, are sterile. It will be interesting to see if the maturation block observed in an oocyte specific knockout for *Sall4* is due to a differentiation process of the oocyte.

The summarized observations do allow for an explanation according to a progenitor population based phenotype but do not prove it. To prove that Sall factors are required to maintain an undifferentiated state in adult stem cells, further experiments need to be carried

out that directly address identity of *Sall* expressing cells as well as their potential upon loss of Sall function.

To formally show that *Sall1* is expressed in the neuronal stem cell compartment, it would be required to genetically label *Sall1* positive cells in adulthood and trace their fate. Expression of an inducible version of Cre recombinase (CreERT2, (Metzger et al., 2005) from the *Sall1* locus could be employed. Upon activation of Cre in adult mice, excision of a transcriptional stop cassette in the 5´UTR of a ubiquitously expressed reporter will lead to irreversible labelling (Novak et al., 2000) of all cells expressing *Sall1* as well as all progeny. In this setup one may ask if neurons that do not express *Sall1* but the activated reporter can be identified in the adult mouse. While such a study could show presence of Sall1 in cells giving rise to newly generated neurons in adult mice, requirement of *Sall1* would not be addressed. Disassociation of postnatal mouse brain and subsequent neuronal stem cell culture in mice null for *Sall1*, *Sall2* and *Sall3* would address if self renewal potential is reduced.

Recently, a study showed the upregulation of Xenopus Sall4 during regeneration of limbs (Neff et al., 2005) fitting with the function I propose of *SALL* genes in vertebrates. While most mammalian tissues have lost the capacity of regeneration, it is maintained in the liver. Since only *Sall4* is expressed in the liver, redundancy is not a concern in this system. An inducible Cre line has been reported for the adult liver (Schneider et al., 2003). Excision of the conditional *Sall4* allele generated in this study can deplete the adult liver of Sall activity. Breedings to establish the required genotype are currently on the way. I intend to delete *Sall4* from the adult liver and then induce liver regeneration by surgical removal of liver tissue. In this setup I will be able to ask if regeneration of adult tissues requires the expression of *Sall* genes.

4. Murine Compound Knockouts Partially Recapitulate Human Syndromes Caused by Mutations in Sall Genes.

Townes-Brocks Syndrome (TBS) and Duane Radial Ray Syndrome (DRRS) are caused by autosomal dominant mutations in *SALL1* and *SALL4* respectively.

Until date, no mouse model recapitulating the defects observed in humans suffering from mutations in *SALL* homologs could be established. Expression of a truncated version of SALL1 (Kiefer et al., 2003) while enhanced the phenotype of a *Sall1* knockout mouse. Nevertheless, observed malformations do not all appear similar to TBS and are also not

penetrant even in homozygosity. The compound knockout strategy used in the presented study recapitulates anterior digit loss as well as neurogenic defects. Like for TBS, a hypoplastic corpus callosum was observed in *Sall1/Sall2* double knockout animals. Further, *Sall1/Sall2* knockouts display hypoplasia of uterus and vas deferens (preliminary observations).

Mutations in humans include insertions, deletions, nonsense mutations as well as frameshift mutations and result in highly variable phenotypes. In a clean genetic deletion of exon two as it was used for this study, phenotypes appear reproducible and invariable e.g. in CNS and limbs. Also, phenotypes are typically seen only upon deletion of several genes expressed in a specific tissue. This supports previous work stating that expression of truncated protein rather than loss of one allele is causal for the observed defects in TBS and DRRS. Since phalanx duplication was never observed and skeletal elements were lost upon SALL reduction, it can be argued that some mutations in *SALL1* and *SALL4* are gain of function mutations. The high variance in clinical manifestations of TBS and DRRS may thus be due to different functional effects of different molecular lesions.

The identification of the progenitor pools as sites of *Sall* expression and action sheds new light on the underlying nature of TBS and DRRS as well as on the general function of *Sall* genes in vertebrate development. Interpretation of defects observed in Townes-Brocks Syndrome and Okihiro Syndrome should consider the finding that SALL transcription factors are required for mammalian progenitor populations.

Acknowledgements

I am very grateful for the help and support I received during the years of my PhD.

Primarily I would like to thank Mathias for the opportunity to work in his lab. He was always available, invested a lot of time and taught me a great deal. Thank you, Mathias!

Steve Cohen and Claus Nerlov were on my thesis advisory committee and I am thankful for all suggestions, support and constructive discussion.

Lieber Stephan, danke, dass Du mit mir hier schon durch die zweite Akademische Abschlussarbeit gehst! Vielen Dank, Herr Professor Dr. Witzgall, dass Sie diese Arbeit als Gutachter einsehen. Sehr Geehrte Herren Professoren Seufert, Kunzelmann und Warth, ich danke Ihnen dafür, als Prüfungskommission für den Schlussakt zur Verfügung zu stehen.

Für die Arbeit an *Sall4* hatte ich in Christian einen tollen Mitstreiter! Ohne sein Wissen, sein Geschick und seine Geduld im Umgang mit Blastulas (und mir) wäre vieles nicht möglich gewesen. Dankeschön!

Also, big thanks must go to past and present members of the Treier-lab, Vitor, Lars, Katrin, Eve, Dirk, Maria, Sandra, Henry, Thomas, Anna-Corina, Claudia, Anne, Jana, and Tobi for constructive and distracting discussions, help, play and a lot of little anecdotes to remember. Especially, thank you, Katrin and Sandra, for all the help with ES cell culture and and big thanks to Anne, Thomas and Henry for critical reading of the manuscript!!!

Thanks to all the people in the animal house for taking care of my mice and thank you to the EMBL transgenic service for such successful ES cell injections and all the germline transmissions.

The great years at EMBL and in Heidelberg were also largely due to the exciting international community. I would like to thank all friends for the wonderful time I could spend with them, their support and friendship. Especially Julius, Uli, Lars, Thomas and, of course, Natascha; Mikko, you were like a brother to me! I will miss you at home, in the bars and on the bike...

Liebe Familie, Vielen Dank für Eure Unterstützung auf verschiedenen Ebenen!

Liebe Edith, Dir gebührt der allergrösste Dank! Du hast mich immer unterstützt, und mir die Möglichkeit gegeben, hier in Heidelberg zu promovieren, Danke für die vielen Besuche und die schöne Zeit mit dir. Ich freue mich darauf, in Zukunft mehr davon mit Dir zu haben.

Materials and Methods

1. Materials

1.1. Chemicals

Acrylamide	Biorad
L-Adenine	Sigma
Agar, Bacto	Becton-Dickinson
Agarose	Invitrogen
Ampicillin	Sigma
Ammonium Chloride	Merck
Ammonium-Persulfate (APS)	Biorad
Bovine Serum Albumin, Fraction V (BSA)	Sigma
BrdU	Sigma
Bromophenol blue	Sigma
Calcium Chloride	Merck
Carrier DNA (Herring testis)	Clontech
Chloramphenicol	Sigma
Dextran Sulfate	Amersham Pharmacia
Dextrose	Merck
DiI	Molecular Probes
Dithiothreitol (DTT)	Biomol, Germany
DMSO (Dimethylsulfoxide)	Sigma
DPX	Agar Scientific
EDTA	Merck
Ethanol	Merck
Ethidium Bromide	Sigma
Eukitt mounting medium	Fluka
Fetal Calf Serum	Invitrogen
N,N-Dimethylformamide	Sigma
Formalin	Electron Microscopy Sciences
L-Glutamine	Invitrogen
Glycine	Merck

HEPES	Invitrogen
Isopropanol	Merck
Kanamycin	Sigma
Lithium Acetate	Sigma
Lithium Chloride	Merck
Magnesium Chloride	Merck
MEM Non-essential amino acids	Invitrogen
β-Mercaptoethanol	Sigma
Methanol	Merck
P^{32}-dCTP	Amersham
Paraformaldehyde (PFA)	Sigma
Penicillin/Streptomycin	Invitrogen
Peptone	Difco
PEG4000	Sigma
Potassium Chloride	Merck
Protease Inhibitors Complete	Roche
Puromycine	Sigma
^{35}S-UTP	Amersham
Sodium Acetate	Merck
Sodium Hydroxide	Merck
Sodium Pyruvate	Invitrogen
TEMED	Biorad
Thionin powder	Sigma
Triethanolamine	Sigma
Triton X-100	Sigma
Trypsin	Invitrogen
Tween20	Gerbu
Xylene	Sigma

1.2. Equipment, Plastic Ware and other Materials

Acid washed beads (0,5 mm)	Sigma
Aquamount, Aqueous Mountant	LERNER Laboratories
Bacterial Petri dishes	Greiner Bio-One

High-Density mouse BAC membranes	Research Genetics
Cell counter chamber	(Neubauer improved) BRAND
Cell culture dishes	NalgeNUNC International
Cell incubator Hera cell	KendroLaboratoryProducts
Centricon YM-10	Amicon
Centrifuge 5417C	Eppendorf
Centrifuge RC5C	Sorvall, Instruments
Centrifuge RC-3B	Sorvall, Instruments
Cryotome CM 3050S	Leica Microsystems
Coverslips (24x60 mm)	Superior Marienfeld
Electrophoresis chambers	PeqLab
Electroporation system Gene Pulser II	BIO RAD
Electro-SeparationSystem S&S BIOTRAP	Schleicher&Schuell Bioscience
Filtration bottles STERICUP	MILLIPORE
Fine forceps and scissors	ROBOZ
Gene Pulser , Electroporation cuvettes	BIO RAD
Hybridiser incubator Techne HB-1D	TECHNE
Hybridization bottles	TECHNE
Hybridization Transfer Membrane NEN	Life Science Products
Microlance hypodermic needles	Becton Dickinson
Microscope Slides Super Frost ,Plus 25x75mm	ROTH
Microscope Axiovert 25	Zeiss
Mini Protean 3 system	Bio-Rad
Mini Trans-Blot Cell	Bio-Rad
Nitrocellulose Membranes PROTRAN	Schleicher & Schuell
Nylon Membranes Biotrans	ICN Biomedicals Inc.
OCT Coumpound Embedding Medium	Tissue-Tek,
Optical Microscope	Zeiss
Paraffin Microtome RM2165	Leica Microsystems
PCR thermal cycler PTC-200	Biozym
PCR thermal reaction tubes	ABgene
Plastic tubes (50 ml and 15 ml)	FALCON
Power supply Power Pac 300	BioRad
Reaction tubes (1.5 ml)	ROTH

Serological Pipets (25 ml; 10 ml; 5 ml; 2 ml)	FALCON
Spectrophotometer Ultraspec 3000	Pharmacia Biotech
Ultracentrifuge L8-M	BECKMAN

1.3. Enzymes

Proteinase K	MERCK
Hot Star Taq DNA Polymerase	QIAGEN
Restriction endonucleases	New England BIOLABS;Roche,
Pwo DNA polymerase	Roche
Expand High Fidelity system	Roche
Klenow	New England Biolabs

1.4. Molecular Weight Markers

1kb DNA ladder	Invitrogen
Broad Range protein ladder	Biorad

1.5. Oligonucleotides

Mouse Genotyping:

Sall1tailF:	5’ act aag agc aag gat gcc cat gtc tg 3’
Sall1internR:	5’ gat gtt ggt gac acc gct cag agt gc 3’
Sall1gfpR:	5’ gtg cagatg aac ttc agg gtc agc 3’
Sall2internF:	5’ gtg aca gat gaa gat tcc cta gca gg 3’
Sall2lacZ-F:	5’ gtc gct acc att acc agt tgg tct gg 3’
Sall2tailR:	5’ agc caa cag cat gtg ctt ctt gag gg 3’
Sall3tailF:	5’ att ctg ggg cac acc cag cct gta gc 3’
Sall3internR:	5’ ccg cag atg tta cat ttg aag ggc cg 3’
Sall3KO-R:	5’ ggt caa agt aaa cga cat ggt cac gtg g 3’
Sall4tailR:	5’ tca gtagtg gcc ttg tct atg ttc tcc 3’
Sall4internF:	5’ cca agg ttt tcg gga cag ata gct cc 3’
Sall4gfpF:	5’ atg gac gag ctg tac aag taa agc gg 3’
S4loxP1F:	5’ ctc cac caa ctc tag ctg cta atg gc 3’

S4loxP1R:	5’ gtt aca gca ata cgg aga tac aca gc 3’
S4loxP2F:	5’ cct ctcacg tag ctg caa tga taa gc 3’
S4loxP2R:	5’ atg gcc ttt gca cat gtg ttc tgg ag 3’
S4onloxP1F:	5’ gct atc tag tgg gac gta gtg cca caa ac 3’
Cre1:	5’ gcc tgc att acc ggt cga tgc aac ga 3’
Cre2:	5’ gtg gca gat ggc gcg gca aca cca tt 3’
hCre1:	5’ gct gcg cat tgc cga aat tgc 3’
hCre2:	5’ ggc aga aca ggt agt tgt tgg 3’
FlpE1:	5’ cta atg ttg tgg gaa att gga gc 3’
FlpE2:	5’ ctc gag gat aac ttg ttt att gc 3’
lacZ-F:	5’ tac cca act taa tcg cct tgc agc ac 3’
lacZ-R:	5’ tca gac agc aaa cga ctg tcc tgg 3’

Southern Blot Probes

Sall2-5´F:	5’ gct tca aga cag tgc ttg cat tct aca c 3’
Sall2-5´R:	5’ agc cca cag cac aca aga aga ggc 3’
Sall4-5´F:	5’ acg ttg cag tat ttg gac agg atc ag 3’
Sall4-5´R:	5’ ttc tca aca gta agt cta ggg gat gg 3’

RT-PCR

Sall4-RTF:	5’ cga gca gcc tca gca gct acc gag 3’
Sall4-RTR:	5’ ttc aag tac agg atg gag tcc gtg cc 3’
Actin-RTF:	5’ ggc cca gag caa gag agg tat cc 3’
Actin-RTR:	5’ acg cac gat ttc cct ctc agc 3’

1.6. Antibodies

Primary Antibodies:

αphospho-Histone H3	rabbit	Upstate Biotech
αCaspase	rabbit	Cell Signalling Technol.
αTROMA	rat	gift from Dr. R. Kemler

αnanog	rabbit	CosmoBio
αOct4	mouse	Santa Cruz Biotech.
αCdx2	mouse	BioGenex
αBrdU	rat	Abcam
α–βgalactosidase	rabbit	Cappel
αGFAP	rabbit	DAKO
αNeuN	mouse	Chemicon
αHA	biotinylated	Roche
α–Tubuline	mouse	gift from Dr. M. Taipale
CSPG	mouse	Sigma
α mouse T cells	rabbit	Cedarlane

Peptide Antibodies against SALL4

N- and C-terminal peptides were synthesized commercially and coupled to KLH. The sequence for the N-terminal peptide is KQAKPQHINWEEGQGEC, the C-terminal peptide sequence is CAKHQFPHFLEENKIAVS. Each peptide was injected into two rabbits (Animal IDs N-terminus: SG-2671, 2672, Animal IDs C-terminus: SG-2673, 2674) using Freud's adjuvant to induce antibody production against these epitopes.

In a final bleed after the 5th boost, 76 and 80 ml of serum were obtained from SG-2671 and SG-2672 respectively. A dilution series on SG-2671 showed that the signal to background ratio is best in a 1:5000 dilution for western blotting. Testing the sera for immunohistochemistry revealed good immunoreactivity in a 1:1000 dilution for serum of SG2671. After purification, a 1:1000 to 1:3000 dilution works best in immunohistochemical applications. All immunostainings shown subsequently were thus done diluting the antibody 1:1000 in PBS.

Secondary Antibodies:

HRP-coupled αRabbit (western)	Goat	Chemicon
HRP-coupled αMouse (western)	Goat	Chemicon

All biotinylated antibodies were purchased from Vector Laboratories

All fluorophore-labelled antibodies came from Jackson ImmunoResearch

Standard guinea pig complement Cedarlane

1.7. Plasmid Vectors

Bacterial vectors:	
pBluescript SK+	Stratagene
pSP73	Promega
Mammalian vectors:	
FLPe cassette	Stewart, A.F.
pKS-DTA	Tajbakhsh, S.
Mouse BAC genomic clones	Research Genetics

1.8. Commercial Kits

ABC kit	Vectastain
DAB Peroxidase Substrate System	Vector
Geneclean Spin kit	Q-BIO gene
Large Construct kit	Qiagen
QIAquick Gel Extraction kit	Qiagen
Maxiprep	Qiagen
Prolong Antifade Kit	Molecular Probes

1.9. Generally Used Solutions

Acidic Tyrode Solution	0.800 g/100 ml NaCl, 0.020 g/100 ml KCl, 0.024 $CaCl_2 \cdot 2H_2O$, 0.010 g/100 ml $MgCl_2 \cdot 6H_2O$, 0.100 g/100 ml Glucose, 0.400 g/100 ml Polyvinylpyrrolidone. Prepare at room temperature and adjust to pH 2.5 with HCl.
BAC Church:	7 % SDS, 25 mM Na_2HPO_4 pH 7.2, 1 % BSA, 1 mM EDTA
BAC Wash:	1 % SDS, 20mM Na_2HPO_4, 1mM EDTA
Gel-loading Buffer:	0.25 % bromophenol blue, 15 % Ficoll, in TE
PBS:	137 mM NaCl, 2.7 mM KCl, 6.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4
20xSSC:	3M NaCl, 0.3M Na_3citrate$\bullet 2H_2O$ (pH to 7.0 with 1M HCl)
50xTAE:	0.04 M Tris-Acetate, 0.001 M EDTA

LB:	1 % Bacto-tryptone, 0.5 % Bacto-yeast extract, 0.5 % NaCl
Tail Buffer:	0.1 M EDTA, 0.1 M NaCl, 1 % SDS, 0.05 M Tris-HCl, pH 7.6, ProteinaseK, 0.5 mg/ml
5xTBE:	45 mM Tris-borate, 1 mM EDTA
TBF1:	30 mM Potassium acetate, 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$, Glycerol to 15 % (v/v), pH 5.8 with acetic acid
TBF2:	10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl, Glycerol 15 % (v/v), pH 6.5 with KOH
TBS:	100 mM Tris-HCl pH 7.5, 150 mM NaCl
TE	10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)

1.10. Cells

Bacterial Strains

E.coli XL-10 strain: Tetr, Δ(mcrA)183, Δ(mcrCB-hsdSMR-mrr)173, endA1, supE44, thi-1, recA1, gyrA96, relA1, lac Hte (F' proAB, lacIqZ DM15 Tn10 (Tetr) Amy Camr)a

ES Cells

ES cells used for gene targeting were from the IB10 and R1 line, both mycoplasma-free. IB10 was used for targeting of *Sall2*, both *Sall4* alleles described were targeted in R1.

2. Methods

2.1. DNA - Plasmids

Preparation of Plasmid DNA from Bacteria

Mini preparation of plasmid DNA was performed according to standard protocols. For large scale preparation of plasmid DNA, the QIAGEN Plasmid Maxi Kit was used according to the manufacturer's instructions.

Purification of Supercoiled DNA by CsCl Gradient Centrifugation

For purification of supercoiled plasmids, 4 g Caesium Chloride were dissolved in 3.8 ml of TE containing the DNA to be purified. 0.2 ml of Ethidium Bromide solution (10 mg/ml in

water) were added and the mixture was transferred to a 4 ml centrifuge tube (Beckman) for centrifugation in a Beckman vertical Vti65 rotor. After sealing, the tubes were spun O/N at 60000 rpm and 25°C. One visible band at the center of the gradient was collected into a glass tube using a hypodermic needle. Ethidium Bromide was extracted from the DNA by a series of equal volumes of 1-Butanol extractions until no pink colour was detected. The DNA was centrifuged in a centricon filter device at 4000 rpm (SS34 - Sorvall) and 25°C and washed 3 times with TE to remove CsCl.

Spectrophotometric Determination of DNA and RNA Concentration

Spectrophotometric measurements of DNA solutions were done at wavelengths of 260 nm. An OD=1 at 260 nm corresponds to a concentration of 50 μg/ml for double-stranded DNA and 33 μg/ml for single-stranded oligonucleotides and 40 μg/ml for RNA.

DNA Restriction and Klenow Treatment

To a solution of DNA the appropriate 10xbuffer and restriction enzyme were added according to the manufacturer's recommendations, mixed in a reaction tube and typically incubated O/N at the appropriate temperature. Klenow treatment was done in accordance with the NEB protocol.

Electrophoresis of DNA

Analysis and preparation of DNA were performed using agarose gels containing 0.5 μg/ml Ethidium Bromide. 1xTAE-buffer was used as gel and electrophoresis buffer. DNA samples were loaded into the gels in 1xDNA loading buffer and run at 5 V/cm of gel. The DNA samples were generally run in parallel to a DNA molecular weight marker (1kb ladder).

Isolation and Purification of DNA from Preparative Agarose Gels

DNA was isolated from agarose using two different commercial kits according to the fragment size. For fragments up to 6 kb, the QIAquick® Gel extraction Kit (Qiagen) was used according to the manufacturer's instructions. For fragments bigger than 6 kb, the Gene Clean Spin Kit was used according to the manufacturer's instructions. For purification of large amounts of DNA from agarose gels, the S&S BIOTRAP Electro-Separation-System for elution and purification of charged molecules was used according to the manufacturer's instructions.

DNA Ligation

For ligation of purified DNA fragments into linearized plasmid vectors, an appropriate amount of each DNA species was incubated O/N in a total volume of 10 µl with 1 µl of 10xbuffer and 1U DNA ligase. Ligation reactions were incubated O/N at 16°C.

Preparation of Chemocompetent Escherichia coli XL-10 Cells

5 ml of LB medium (+tetracycline) were inoculated with a single colony of E.coli XL-10 and incubated with shaking (300 rpm) O/N at 37°C. 250 ml of LB medium were inoculated with 3-4 ml of the overnight culture and incubated with shaking at 37°C until an OD600 of 0.5 to 0.6. The cell suspension was incubated on ice for 5 minutes, poured into 50 ml Falcon tubes and centrifuged at 3000 rpm (GSA - Sorvall) and 4°C for 10 minutes. The supernatant was discarded and the bacterial pellet was resuspended in 20 ml of TBF1 buffer (2/5 volume) and incubated on ice for 5 minutes. The cell suspension was centrifuged at 3000 rpm (RC-3B) and 4°C for 5 minutes, the supernatant was discarded and the bacterial pellet was resuspended in 2 ml TBF2 buffer (1/25 initial volume). After 15 minutes of incubation on ice 200 µl aliquots were frozen in liquid nitrogen and stored at -80°C.

Transformation of Chemocompetent Escherichia coli XL-10 Cells

100µl of competent E.coli XL-10 cells were added to DNA in an eppendorf tube and incubated on ice for 30 minutes, the mixture was then heat shocked for 90 seconds at 42°C, followed by a 3 minutes incubation on ice. 1 ml of LB medium was added and the cell suspension was incubated for 30 minutes at 37°C. The transformation was then plated on the appropriate selective media and incubated at 37°C overnight.

2.2. DNA - Genomic

Preparation of Genomic DNA

For genomic DNA preparation from mouse tails, each tail was incubated O/N in 800 µl of tail buffer at 56°C. The following day, 300 µl of 6M NaCl was added, mixed by shaking and centrifuged for 10 minutes at 14000 rpm (Eppendorf) at RT to pellet the tail debris. 800 µl of the supernatant were transferred into a new reaction tube, mixed with 500 µl isopropanol

and centrifuged at 14000 rpm for 5 minutes to pellet the genomic DNA. Supernatant was aspirated and the recovered DNA was washed in 70 % ethanol and dissolved in 100 µl TE.

For preparation of genomic DNA from ES cells, a confluent well from a 24 well plate was washed twice in PBS, incubated O/N at 37°C with 600 µl ES cell lysis buffer (10 mM Tris-HCl pH 8.0, 5 mM EDTA, 100 mM NaCl, 1 % SDS and 0.5 mg/ml Proteinase K). The next day DNA was precipitated in an eppendorf tube by addition of 500 µl isopropanol followed by mixing by shaking. The precipitated DNA was transferred to a new tube with a tip and washed O/N with 70% ethanol. After spinning down the DNA was resuspended in 100 µl TE O/N at 37°C

Polymerase Chain Reaction (PCR)

For mouse genotyping, the standard genomic PCR reaction was performed with a hot start protocol that employs the HotStarTaq DNA polymerase. One reaction contained:

16 µl H2 O

2 µl 10xPCR buffer (QIAGEN)

0.5 µl dNTP solution (10 mM each dATP, dCTP, dGTP, dTTP)

0.2 µl each primer 50 mmol)

0.5-1U HotStarTaq DNA polymerase

5 ng DNA in 1µl

The PCR reaction was initially denatured for 15 minutes at 94°C to activate the HotStarTaq DNA polymerase. PCR was typically performed with 36 cycles, each cycle consisting of 40 seconds denaturation at 95°C; 40 seconds annealing at the appropriate temperature for the oligonucleotides; 45 seconds of elongation at 72°C. The amplification products were assayed by gel electrophoresis.

For cloning purposes either the Pwo polymerase or the High Fidelity Expand was used, essentially as described above with the omission of the 15 minutes hot start.

Southern Blot Analysis

Detection and localization of specific sequences within genomic or plasmid DNA was performed using the Southern blot technique. Typically the DNA (20mg of mouse genomic DNA or 1mg of plasmid DNA) was treated with one or more restriction enzymes. The resulting fragments were separated according to size by agarose gel electrophoresis (0.6 % gel cast in 0.5xTBE containing 0.5 µg/ml Ethidium Bromide was used for genomic DNA

samples). After electrophoresis was completed, the gel was photographed with a fluorescent ruler for subsequent assessment of band sizes. The DNA was denatured by incubating the gel in 0.5 N NaOH, 1.5 M NaCl for 30 minutes with agitation. The gel was then briefly rinsed in deionized water before the DNA was transferred to a nylon membrane by capillarity under neutral conditions (20XSSC). The membrane was incubated for two hours at 80°C to immobilize the attached DNA, followed by a pre-hybridization (2 hours minimum) in Church hybridization solution at 65°C. The membrane was then probed O/N at 65°C with a radio-labelled probe in 30 ml of Church buffer. After hybridization, the membranes were washed in the following series of increasingly stringent washing buffers: 2xSSC, 1 % SDS; 0.5xSSC, 1 %SDS; 0.1xSSC, 1 % SDS for 30 minutes each. Autoradiography was used to locate the positions of the bands complementary to the probe.

Radiolabelling of DNA Probes for Southern Blot Analysis

Labelling of DNA probes for southern blot analysis was performed by PCR. Using a previously amplified sequence (in general with a size of 300 bp) as a template in the following reaction, radioactive dCTP-P^{32} was incorporated of into the probe:

8 µl H_2O

2 µl 10xPCR buffer

3 µl dNTP solution (0.5 mM each dATP, dGTP, dTTP)

5 µl P^{32}-dCTP (50 µCi)

0.5 µl each primer

2.5 U HotStarTaq DNA polymerase (QIAGEN)

5 ng Template DNA

The PCR reaction was initially denatured for 15 minutes at 94°C to activate the HotStarTaq DNA polymerase. PCR was usually performed with 30 cycles, each cycle consisting of 20 seconds denaturation at 95°C, 30 seconds annealing at 50°C, and 30 seconds of extension at 72°C. A final cycle of 5 minutes at 95°C was performed to denature the labelled probe. The probe was incubated on ice for 4-5 minutes and then added into the hybridization solution.

2.3. Cell Culture Methods

Culture Conditions

Mouse embryonic stem (ES) cells were cultured in ES medium: glucose-rich (4500 mg/liter) DMEM medium supplemented with 15 % Fetal Calf Serum (FCS), 1 % MEM nonessential amino acids, 1 % Penicillin-Streptomycin, 1 % Glutamine, 1 % Sodium Pyruvat and Leukaemia Inhibitory Factor (LIF). The ES cells were grown on a layer of mouse embryo fibroblasts (MEFs) at 37°C with 5 % CO_2. Selection of ES cell clones that had integrated the replacement vectors was done by adding the antibiotic G418 (250 µg/ml) to the medium. Ongoing cultures of MEFs and ES cells were typically grown in 10-cm tissue culture dishes with 10 ml of medium. All ES-cell medium was changed daily. Puromycin was added to 1µg/ml.

TS cells medium:

20% FCS (Invitrogen), 50 µg / ml Penicillin (Invitrogen), 50 µg / ml Streptomycin (Invitrogen), 1 mM Sodium pyruvate (Invirtogen), 100 µM ß-mercaptoethanol (SIGMA Aldrich), 2 mM L-glutamine (Invitrogen), in RPMI 1640 medium (Invitrogen); If required, Fgf-4 and heparin were added freshly before use: 10 µl of 1000x Fgf-4 (25 ng / ml / SIGMA) and 10 µl of 1000x heparin (1 µg / ml / SIGMA) into 10 ml of TS medium.

Collecting embryos and for handling them outside the incubator:

M2, a modified Krebs-Ringer solution with HEPES buffer substituted for some of the bicarbonate: 94.44 mM NaCl, 4.78 mM KCl, 1.71 mM $CaCl_2$• $2H_2O$, 1.19 mM KH_2PO_4, 1.19 mM $MgSO_4$ • $7H_2O$, 4.15 mM $NaHCO_3$, 20.85 mM HEPES, 23.28 mM Sodium lactate, 0.33 mM Sodium pyruvate, 5.56 mM Glucose, 4 g/l BSA, 0.060 g/l Penicillin G • Potassium salt (final conc. 100 units / ml), 0.050 g/l Streptomycin sulfate (final conc. 50 mg / ml), 0.010 g/l Phenol red, 2x glass-distilled H_2O. All reagents were bought from SIGMA Aldrich.

Overnight culture of blastocysts:

M16 is a modified Krebs-Ringer bicarbonate solution and is very similar to Whitten's medium 94.66 mM NaCl, 4.78 mM KCl, 1.71 mM $CaCl_2$• $2H_2O$, 1.19 mM KH_2PO_4, 1.19 mM $MgSO_4$ • $7H_2O$, 25.00 mM $NaHCO_3$, 23.28 mM Sodium lactate, 0.33 mM Sodium pyruvate, 5.56 Glucose, 4 g/l BSA, 0.060 g/l Penicillin G • Potassium salt (final conc. 100 units / ml),

0.050 g/l Streptomycin sulfate (final conc. 50 mg / ml), 0.010 g/l Phenol red, 2x glass-distilled H_2O. All reagents were bought from SIGMA Aldrich.

Trypsinization of Cells

ES cells and fibroblasts at sub-confluent density were passed into new tissue culture dishes at lower density. The medium was aspirated and the cells were washed once with PBS. The cells were subsequently covered with 0.5 % (w/v) trypsin and incubated at room temperature until they were visibly detached from the dish. The trypsinization reaction was stopped by addition of FCS containing medium. Cells were collected in a 50 ml tube and centrifuged (1500 rpm, 3 minutes) for large volumes. The pellet was resuspended in new medium. Small volumes were transferred onto the new plate directly. For Electroporation, the cells were counted using a Neubauer chamber and plated at appropriate numbers on new tissue culture dishes.

Mitomycin C Treatment of Mouse Embryo Fibroblasts

To mitotically inactivate feeder cells, confluent layers of MEFs in tissue culture dishes were treated with medium containing 10 mg/ml of Mitomycin C for 3 hours at 37°C with 5 % CO_2. The dishes were washed three time with PBS and the cells collected by trypsinization. After centrifugation, the cells were resuspended in feeder medium (glucose-rich (4500mg/liter) DMEM medium supplemented with 15 FCS, 1 % MEM nonessential amino acids, 1% penicillin-streptomycin, 1% glutamine), counted and the appropriate number was transferred to tissues culture dishes to ensure the production of a uniform confluent mono layer of cells.

Freezing and Thawing Cells

Cells were trypsinised, counted and centrifuged for 3 minutes at 1500 rpm. The supernatant was discarded and the pellet was resuspended in new medium at a density of $2.5x10^6$ cells/ml. Aliquots of $1x10^6$ cells were transferred into individual cryo vials and an equal volume of medium containing 20% DMSO was added to each vial (10% DMSO final concentration). The cell suspension was frozen, by storing the vials at -80°C. For long-term storage, vials were transferred into liquid nitrogen. For thawing, cells the vials were rapidly warmed at 37°C and the cell suspension was transferred into a centrifuge tube containing medium. The cells were centrifuged at 1500 rpm for 3minutes, resuspended in new medium

and transferred to tissue culture dishes. ES cells were transferred to tissue culture dishes containing a confluent layer of mitotically inactive MEFs.

Electroporation of ES Cells

An exponentially growing culture of ES cells at passage 4-6 was prepared for electroporation by changing medium approximately 3 hours before harvesting the cells. The cells were trypsinised and medium added to stop the trypsinization reaction. Cells were centrifuged for 4 minutes at 3000 rpm and RT, washed once and resuspended in PBS as a single cell suspension. The ES cell suspension was counted, and the cell suspension was centrifuged once more, washed and resuspended at a density of $1.5x10^7$ cells/ml in Ca^{2+} /Mg^{2+} free PBS. For each electroporation, 0.8 ml ($1.5x10^7$ cells) of cell suspension was placed in a sterile cuvette and 30 μg of linearized DNA were added and mixed well. This mixture was incubated on ice for 5 minutes and a single pulse of 240 V, 500 mF was applied to the cuvette. After electroporation, the cells were incubated on ice for 5 minutes and plated onto 10-cm gelatinized tissue culture dishes. Each 0.8 ml sample from the electroporation was divided and equally distributed on twelve tissue culture dishes. The cells were allowed to recover for 48 hours in non-selective medium. The G418 selection of ES clones that had integrated the replacement vectors started 48 hours after electroporation.

Isolation of Individual ES Cell Colonies

Selected individual ES cell colonies were isolated 9-10 days after electroporation. Sets of 96-well tissue culture dishes containing confluent mono-layers of MEFs in fresh medium were prepared in advance. Culture medium was replaced by PBS. A 20 μl Gilson pipette with the tip of a 10 μl micropipette was used to pick individual colonies under microscopic observation in 20 μl PBS. The colonies were placed into 96 well U-plates (NUNC) and incubated for 20 minutes at room temperature. Subsequently, 200 μl ESC medium was added, clones disrupted by repetitive pipetting and they were placed into the 96-well F dishes coated with MEFs and incubated at 37°C with 5 % CO_2. 48 hours after isolation, the individual clones were visible giving rise to an exponentially growing set of evenly distributed ES cell colonies. The growth of all ES cell clones was monitored daily and when considered appropriate (~50% confluency for the highest number of clones) all clones in one dish were trypsinized and divided in the following way: Half of the ES cells of each clone were transferred to a new 96-well dish containing mitomycine C treated MEFs and these cells were

subsequently frozen; the remaining cells from each clone were transferred to gelatine coated 24-well dishes for later preparation of genomic DNA. Homologous recombination was verified by southern blot.

De Novo Isolation of ES Cell Lines from Blastocysts

Embryos were flushed from the uterine horns with M2 and placed individually into 10 mm well tissue culture dishes containing a preformed feeder layer of MEF and 0.5 ml of ES cell culture medium.

Four to five days after explantation into culture, ICMs were dislodged from the underlying sheet of trophoblast cells using a finely drawn Pasteur pipette and transferred through two changes of Ca^{++}/Mg^{++}-free PBS.

Incubation in a microdrop of trypsin lead to disassociations at 37˚C for 3-4 minutes. After complete mechanical separation of cells, they were transferred into a fresh 10-mm feeder cell tissue culture well. Individual cultures were monitored for ES cell growth. After 2-5 days of culture primary colonies become visible. If clumps with ES cell morphology dominated the culture, clones were expanded.

Immunosurgery

Immunosurgery was performed according to Solter and Knowles, 1975.

The zona pellucida was removed prior to immunosurgery with Tyrode's Acid.

ES Cell Injection into Blastocysts and Chimera Production

Blastocyst injection was performed at the EMBL Transgenic Service. IB10 or R1 targeted ES cells with a 129/OLA genetic background were injected into e3.5 mouse blastocysts of the C57/Black6 strain produced by natural matings. Always approximately 12 cells are injected with an eppendorff microinjection apparatus in an inverted microscope under phase contrast. Two hours after injection, living blastocysts were implanted into pseudopregnant CD1 females (7-8 per oviduct arm) 2.5 days after mating to a sterile male to induce pseudopregnancy.

2.4. Tissue sectioning

Tissue Preparation and Fixation

Fixation of mouse tissues, organs and whole embryos was performed differently depending on whether the samples would be used for histology, Immunohistochemistry or *in situ* hybridization.

For histological analysis of early embryos using paraffin sections, implantation sites were isolated and fixed 5 hours at 4°C in 2 % PFA. The tissue was subsequently washed in PBS for 10 minutes at 4°C and stored indefinitely in 70% ethanol at 4°C.

For analysis of gene expression by *in situ* hybridization, organs from mice perfused in 10 % formalin and embryos were properly dissected and stored indefinitely in 10% formalin.

Cryosectioning

After fixation, embryos and organs were incubated O/N at 4°C in RNAse-free 20 % Sucrose/PBS (w/v) solution O/N or until the tissue sank. OCT medium was placed in embedding moulds and the tissue sample was introduced in the medium. The moulds were placed in dry ice-cooled ethanol to freeze the medium. Typically, 18 μm sections were cut on a Leica Cryotome with a chamber temperature of -20°C. Tissue sections were transferred onto glass slides and air dried for a minimum of 2 hours and stored at -20°C for later processing.

Paraffin Embedding, Mounting and Sectioning

Specimen, preserved in 70% ethanol, were dehydrated in a series of 85%, 96% and three times 100% ethanol at least for 1 hour for each step at room temperature. The samples were cleared in two changes of xylene for 30 minutes each and incubated O/N in xylene at room temperature. Tissues were carried through a series of four changes in paraffin 1hour each at 60°C and embedded in the proper orientation in embedding moulds. The moulds were stored at 4°C until sectioning. The microtome was set to cut 7 μm sections.

2.5. Histochemistry and Immunohistochemistry

PAS Staining

Paraffin sections of implantation sites were incubated twice for 10min in xylene followed by rehydration in a series of 100%, 96% and 70% ethanol for three minutes each.

After washing in distilled H2O, the sections were incubated in Periodic acid 0.5% for 5 min. Sections were washed through a series of 3 fast changes of distilled H2O and incubated in Schiff's Reagent (0.5% Basic fuchsin; 0.5% Sodium metabisulfite; 0.2% Activated charcoal, filtrated) for 20min.

Sections were rinsed 7min under running tap water, stained 1min in Hematoxylene GILL-3 and rinsed again 7min under running tap water. Dehydration in an ascending ethanol series was followed by xylene incubation for clearance and mounting.

Nuclear Fast Red Counterstaining

Subsequent to β-galactosidase staining or to DAB-substrate to visualize HRP-coupled antibodies, the slides were washed in PBS for 15min at room temperature, rinsed in distilled water and counterstained with Nuclear Fast Red (Vector) for 5min. Sections were washed in distilled H2O for 10min and dehydrated through a ethanol series (twice 1min each in 70%, 95% and 100%ethanol). Sections were cleared in xylene twice for 5 to 10 minutes and mounted in xylene based medium.

β–galactosidase Staining

Expression of the *lacZ* reporter gene present in the replacement vectors for the *Sall1, Sall2* and *Sall3* genes was visualized by staining whole embryos and frozen tissue sections for β-galactosidase activity. A β-galactosidase staining solution containing Bluo-Gal substrate (20mM K3; 20mM K4; 0.01% Na-deoxycholate; 2mM MgCl2; 0.02% (v/v) NP-40; 1mg/ml Bluo-Gal) was used. Staining was carried out at 37°C over night.

Thionine Staining

Staining solution was prepared by slow addition of 0.625g Thionin powder to a solution of 191ml water, 9ml NaOH-1M, 50ml acetic acid-1M at 60°C. Solution was boiled for 45 minutes, filtrated and stored at 4°C for a maximum of 4 months. Staining was done subsequent to incubation in water and takes 5 minutes. Sections were dehydrated in an ethanol series directly after and mounted in xylene based medium.

Immunofluorescence

Imunofluorescent stainings were performed on 18 μm cryostat sections prepared as described. Sections were rinsed for 10 minutes in PBS. Blastocysts were fixed 5 minutes in

2% paraformaldehyde in PBS, permeabilized in 0.2% Triton X100 for 5 minutes and directly placed on a coated glass slide and allowed to dry in. For blocking, the tissue sections were incubated for 2 hours with 5 % (v/v) 2nd antibody species serum in TBS/0.4 % Triton X-100 at room temperature. sections were washed twice with TBS/0.4% triton X-100 and incubated with the primary antibody at the appropriate dilution in TBS/0.4% triton X-100 O/N at 4°C. The slides were rinsed twice with TBS/0.4% triton X-100 and incubated with fluorescent-labelled secondary antibody in TBS/0.4 % Triton X-100 for 1 hour at room temperature in the dark. The slides were washed twice in 1xTBS/0.4% triton X-100, counterstained in 1xDAPI for 5 minutes and mounted with 50µl of slow fade reagent (Prolong Antifade kit, Molecular Probes) according to the manufacturer's indications. Slides were stored at 4°C in the dark.

Immunohistochemistry

Immunohistochemical stainings were performed on sections fixed in 4% paraformaldehyde in PBS for 2-12 hours. Paraffin sections were initially incubated in xylene for 3x5 minutes to dissolve the paraffin followed by rehydration in a series of 100%, 95% and 70% ethanol (3 minutes each). After washing in PBS for 5 minutes, antigen was unmasked by boiling 1 hour in unmasking solution (Vector) and endogenous peroxidases were inactivated by incubation for 10 minutes in 3 % H_2O_2 in PBS followed by 5 minutes washing in PBS with agitation if DAB stainings were performed.

The sections were incubated with blocking buffer (0.2 % Triton, 1 % Glycine, 3% BSA, 5 % normal serum in PBS) in a wet chamber for 2 hours at room temperature. After blocking, the sections were incubated with primary antibody in blocking buffer in a wet chamber ON at 4°C. The slides were washed for 3x20 minutes in wash buffer (0.4 % Triton, 1 % Glycine, 3 % BSA in 1xTBS) with agitation. The sections were incubated with biotinylated secondary antibody (vector) in blocking buffer in a wet chamber for 1 hour at room temperature and then washed as after primary antibody. The staining procedure using the ABC reagent and DAB substrate was performed according to the manufacturer's recommendations. The sections were counterstained with methyl green stain (1:8 dilution from 2 % stock 10 minutes) dehydrated 2x10 minutes in 1-Butanol, cleared in xylene and mounted with 60 µl Eukitt.

Embryo Staining for Cartilage and Bone

In order to reveal the skeletal structure and cartilage formation at e18.5, the following protocol was performed. Mice were sacrificed and placed in tap water for 24 hours. The skin was carefully peeled off with forceps by making a cut around the waist, neck, and base of the

limbs and pulling the skin back intact. Mice were eviscerated and fixed in 95%ethanol for 3-5 days. Staining for cartilage was performed by placing the mice in Alcian Blue solution (0.015% w/v) (Sigma). Staining was performed for 24 hours. The mice were then rinsed twice in 95% ethanol and then let sit again in 95%ethanol for 24 hours. The mice were then placed in 2% KOH for a period of time that allows them to be cleared. Counterstaining for bone with Alizarin Red solution (0.005% w/v) (Sigma) stain was performed for 1-3 hours until an optimal level of staining was observed. The mice were cleared by placing them in 2% KOH of decreasing strengths. Initially a 2% KOH solution was used until the mice were almost clear. The clearing process was completed in the following ratios of 2% KOH to glycerol; 80:20, 60:40, 40:60, 20:80. The mice were stored in 2% KOH, glycerol (20:80).

2.6. In situ hybridization

Generation of *in Situ* Probes by *in Vitro* Transcription

The radioactive labelling of RNA probes was performed using the Riboprobe® Combination System-SP6/T7. An *in vitro* transcription labelling reaction was set up and incubated for 1 hour at 37°C:

4 µl of 5xtranscription buffer
2 µl of 100 mM DTT
0.6 µl RNasin
4 µl of A/G/C mix (2.5mM each nucleotide)
1 µg of linearised template
Nuclease free H_2O to 14 µl
5 µl ^{35}S-UTP (100 µCi)
1 µl of RNA polymerase

The labelling reaction was followed by a treatment with DNaseI at 37°C for 15 minutes. This reaction was stopped by addition of 30 µl nuclease-free water and the labelled RNA was recovered using ProbeQuant™ G-50 Micro Columns according to the manufacturer's protocol. A probe concentration of $5x10^6$ cpm/ml was routinely used on the different tissue sections. The calculated amount of ^{35}S-labeled probe for 1 ml Hybridization solution was initially mixed with 50 µl tRNA, 10 µl 1M DTT and nuclease-free H_2O to 200 µl. This

mixture was incubated at 65°C for 5 minutes to denature annealed strands in the labelled probe and added to 800 µl Hybridization solution (50 ml l Hybridization solution: 25 ml formamide, 10 ml 50 % dextran sulphate, 3 ml 5 M NaCl, 1 ml 1xDenhardt's solution, 0.5 ml 1 M Tris-HCl pH 8.0, 0.1 ml 0.5 M EDTA pH 8.0, 0.4 ml nuclease free H_2O). The hybridization mixture was stored at -20°C until the hybridization step for a maximum of 1 week.

Pre-treatment and Prehybridization of Tissue Sections

18 µm frozen cryostat sections or 7µm paraffin sections were used for *in situ* hybridization. The tissue sections were initially treated with Proteinase K (10 mg/ml) in proteinase digestion buffer (0.1 M Tris-HCl pH 8.0, 50 mM EDTA pH 8.0) at 37°C for 30 minutes. Slides were rinsed briefly in dH_20 and incubated in 0.1 M triethanolamine (TEA) pH 8.0 for 3 minutes at room temperature, and then acetylated in acetylation buffer (625 µl acetic anhydrid, 250 ml 0.1 M TEA) for 10 minutes at room temperature. Slides were then incubated in 2xSSC for 5 minutes and dehydrated through an ethanol series of 50%, 70%, 90% and 2 times 100% ethanol for 3 minutes each step. The slides were drained for 5 minutes and dried under vacuum in an exciccator for a minimum of 2 hours until hybridization.

Hybridization

To hybridize the tissue sections with a ^{35}S-labelled probe, an aliquot of the hybridization mixture was placed at 65°C for 10 minutes and then centrifuged for 10 minutes at 3000 rpm to eliminate any precipitates that may have formed. 65-75 µl of hybridization mixture was applied to each slide (50x22 mm) and the tissue sections were then covered with a coverslip and sealed by the application of liquid DPX mounting medium around the edges of the cover slip. For hybridization, the slides were incubated O/N on a slide-warming tray at 60°C. After hybridization, the slides were cooled and incubated four times 15 minutes with 4xSSC. The slides were subsequently treated with RNaseA in RNase buffer (RNaseA 20 µg/ml, 25 ml 5 M NaCl, 2.5 ml 1 M Tris-HCl pH 8.0, 500 µl 0.5 M EDTA pH 8.0, dH_2O to 250 ml) for 37 minutes at 37°C. This was followed by a washing and gradual de-salting series of the following solutions: 2xSSC, 1 mM DTT for 5 minutes at RT; 1xSSC, 1 mM DTT for 5 minutes at RT; 0.5xSSC, 1mM DTT for 5 minutes at RT; 0.1xSSC, 1mM DTT for 37 minutes at 65°C; 0.1xSSC, 1mM DTT for 3 minutes at RT. Slides were then dehydrated through an ethanol series of 50 % ethanol (with 0.08xSSC, 1mM DTT), 70 % ethanol (with 0.08xSSC,

1mM DTT), 95 % ethanol and 100% ethanol each step for 3 min. The slides were finally drained for 5 minutes and dried under vacuum in an exciccator for a minimum of 2 hours.
Slides were dipped in liquid photographic emulsion in complete darkness, dried and stored at 4°C for 1-4 weeks. Emulsion was developed in standard developer at room temperature for 3 minutes, fixed, counterstained in Bisbenzimide, dehydrated, transferred into Xylene and mounted in Eukitt.

Slides were photographed in an Axiophot microscope in dark field to visualize silver grains that formed around radioactively labelled probe and represent expression. The Histology was subsequently visualized by the blue fluorescence. Pictures were overlayed in Photoshop 6.0.

2.7. Mouse Methods

Mice were housed in a specific pathogen free environment (SPF) and under controlled light, temperature (21°C) and humidity (50 % relative humidity) conditions. Water and food were provided as needed. Mice were routinely weaned from their parents at the age of 3 weeks with males and females being housed separately. Mice were tagged using six-digit number eartags and tail biopsies were taken for genotyping.
For experiments with embryos, breedings were set up with one male and one female. Vaginal plug-checks were performed daily at 8am, starting the morning following the set up of the breeding. The presence of a plug was considered day e0.5 of pregnancy. For genotyping of embryos at day e12.5 or younger the yolk sacs were separated from the placenta and used for DNA extraction. With embryos older than e12.5 tissue samples for genomic DNA were normally taken from the tail. The day of birth was considered postnatal day zero (P0).

BrdU Labelling of Cells Undergoing DNA Synthesis

BrdU (Sigma) was dissolved in PBS at 8mg/ml, sterile filtrated and stored at -80°C. If labelling was desired, aliquots were thawed and pregnant dams received 0.4 ml of BrdU solution into the tail vein. For Immunostaining, sectiones were pretreated with 2M HCl for 20 minutes at 37°C to denature DNA and expose the antigen.

Collection of Preimplantation Embryos

Superovulation was induced by an intraperitoneal injection of 5 units of PMSG 11am followed by 5 units of HCG 47 hours later. Mating was started immediately and plugs scored next morning. To collect preimplantation embryos, mothers were sacrificed e0.5-e3.5, oviducts and uteri dissected out and flushed in M2 medium. Blastocysts were collected using a fine drawn glass capillary, washed in M2 and processed further.

To reimplant embryos, foster mothers were bred to vasectomized, sterile males to induce pseudopregnancy and embryos were transferred dorsally into anesthetized females at e2.5 of pseudopregnancy.

2.8. Proteins

Cell-extracts

Extracts from cell-culture were done in RIPA buffer (50mM Tris-HCl pH 7.5, 150 mM NaCl, 1 % NP-40, 0.1 % SDS, 0.5 % Sodiumdeoxycholate, 0.05 % β-mercaptoethanol, protease and phosphatase inhibitors).

Cells were trypsinized, the enzyme inactivated by addition of FCS, spun at 1500 rpm for 3 minutes, the supernant removed and the pellet snap frozen in liquid nitrogen and lysed in RIPA buffer in an eppendorf tube by vortexing for two minutes. The extracts were sheared with a pipette tip and insoluble debris spun down at 14000 rpm.

Protein Concentration Measurements

Protein concentrations were determined by Bio-Rad Bradford protein assay according to the manufactures protocol. In all assays, BSA was used as a standard.

SDS-PAGE

All SDS-PAGE gels were run on the Mini Protean 3 system and the gels wereprepared as follows:

Separation gel: 0.375 M Tris-HCl pH 8.8, 8-18 % acrylamide, 0.1 % SDS, 0.05 % APS and TEMED (3 µl in 5 ml)

Stacking gel: 0.125 M Tris-HCl pH 6.8, 3.9 % acrylamide, 0.1 % SDS, 0.05 % APS and TEMED (5 µl in 5 ml)

Extracts were heated to 90°C for 4 min. in Laemmli buffer (5xLeammli buffer: 0.15 M Tris-HCl pH 6.8, 5 % SDS, 25 % glycerol, and 0.050 % Bromidephenolblue) with 3 % β-mercapto-ethanol added freshly just before heating. The heated samples were spun at 14000 rpm for 2 minutes before loading on the gels. Gels were run at 100-200V in running buffer (25 mM Tris, 192 mM Glycine)

Western Blotting

All blots were done using the Mini Trans-Blot Cell. In all cases nitrocellulose membranes were used. The mini-gels were blotted in a buffer containing 20 % Methanol (analytic grade), 0.1 % SDS, 8.3 mM Tris, and 64 mM Glycine at 80 volts for 60-90 minutes at 4°C.

Probing

Membranes were blocked in PBS, 0.075% Tween20 + 5 % milk for an hour at room temperature or overnight at 4°C. All primary antibodies were diluted in PBS, 0.075% + 5 % milk. In all cases incubations were for 2 hours at RT or overnight at 4°C. All washes were done in PBS, 0.075% Tween20. All secondary antibodies were used in 1:5000 dilution in PBS, 0.075% Tween20 (washing buffer) + 5 % milk for 1 hour at RT. Western blot were developed using ECL from Promega

Cited Literature

Al-Baradie, R., Yamada, K., St Hilaire, C., Chan, W. M., Andrews, C., McIntosh, N., Nakano, M., Martonyi, E. J., Raymond, W. R., Okumura, S., *et al.* (2002). Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. Am J Hum Genet *71*, 1195-1199.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Bard J. B. L., D. J. A., Karavanova I., Lehtonennext term E., Sariola H. and Vainio S. (1996). Kidney development: the inductive interactions. Seminars in Cell & Developmental Biology *7*, 195-202.

Barrio R, S. M., Carulli J, Lipkow K, Gaul U, Frommer G, Schuh R, Jäckle H, C. Kafatos FC (1996). The spalt-related gene of Drosophila melanogaster is a member of an ancient gene family, defined by the adjacent, region specific homeotic gene spalt. Dev Genes Evol *206*, 315-325.

Basson, M., and Horvitz, H. R. (1996). The Caenorhabditis elegans gene sem-4 controls neuronal and mesodermal cell development and encodes a zinc finger protein. Genes Dev *10*, 1953-1965.

Belluscio, L., Gold, G. H., Nemes, A., and Axel, R. (1998). Mice deficient in G(olf) are anosmic. Neuron *20*, 69-81.

Bestor, T. H. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. Embo J *11*, 2611-2617.

Blau, H. M., Brazelton, T. R., and Weimann, J. M. (2001). The evolving concept of a stem cell: entity or function? Cell *105*, 829-841.

Borozdin, W., Boehm, D., Leipoldt, M., Wilhelm, C., Reardon, W., Clayton-Smith, J., Becker, K., Muhlendyck, H., Winter, R., Giray, O., *et al.* (2004). SALL4 deletions are a common cause of Okihiro and acro-renal-ocular syndromes and confirm haploinsufficiency as the pathogenic mechanism. J Med Genet *41*, e113.

Botzenhart, E. M., Green, A., Ilyina, H., Konig, R., Lowry, R. B., Lo, I. F., Shohat, M., Burke, L., McGaughran, J., Chafai, R., *et al.* (2005). SALL1 mutation analysis in Townes-Brocks syndrome: twelve novel mutations and expansion of the phenotype. Hum Mutat *26*, 282.

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

Bradley, A., Evans, M., Kaufman, M. H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Brunet, L. J., Gold, G. H., and Ngai, J. (1996). General anosmia caused by a targeted disruption of the mouse olfactory cyclic nucleotide-gated cation channel. Neuron *17*, 681-693.

Bulfone, A., Wang, F., Hevner, R., Anderson, S., Cutforth, T., Chen, S., Meneses, J., Pedersen, R., Axel, R., and Rubenstein, J. L. (1998). An olfactory sensory map develops in the absence of normal projection neurons or GABAergic interneurons. Neuron *21*, 1273-1282.

Caviness, V. S., Jr., and Sidman, R. L. (1973). Time of origin or corresponding cell classes in the cerebral cortex of normal and reeler mutant mice: an autoradiographic analysis. J Comp Neurol *148*, 141-151.

Chambers, I. (2004). The molecular basis of pluripotency in mouse embryonic stem cells. Cloning Stem Cells *6*, 386-391.

Cheng, A. M., Saxton, T. M., Sakai, R., Kulkarni, S., Mbamalu, G., Vogel, W., Tortorice, C. G., Cardiff, R. D., Cross, J. C., Muller, W. J., and Pawson, T. (1998). Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. Cell *95*, 793-803.

Chenn, A., and Walsh, C. A. (2002). Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. Science *297*, 365-369.

Coskun, V., and Luskin, M. B. (2001). The expression pattern of the cell cycle inhibitor p19(INK4d) by progenitor cells of the rat embryonic telencephalon and neonatal anterior subventricular zone. J Neurosci *21*, 3092-3103.

de Celis, J. F., and Barrio, R. (2000). Function of the spalt/spalt-related gene complex in positioning the veins in the Drosophila wing. Mech Dev *91*, 31-41.

de Celis, J. F., Barrio, R., and Kafatos, F. C. (1999). Regulation of the spalt/spalt-related gene complex and its function during sensory organ development in the Drosophila thorax. Development *126*, 2653-2662.

Doetsch, F., Caille, I., Lim, D. A., Garcia-Verdugo, J. M., and Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. Cell *97*, 703-716.

Domingos, P. M., Brown, S., Barrio, R., Ratnakumar, K., Frankfort, B. J., Mardon, G., Steller, H., and Mollereau, B. (2004a). Regulation of R7 and R8 differentiation by the spalt genes. Dev Biol *273*, 121-133.

Domingos, P. M., Mlodzik, M., Mendes, C. S., Brown, S., Steller, H., and Mollereau, B. (2004b). Spalt transcription factors are required for R3/R4 specification and establishment of planar cell polarity in the Drosophila eye. Development *131*, 5695-5702.

Dong, P. D., Todi, S. V., Eberl, D. F., and Boekhoff-Falk, G. (2003). Drosophila spalt/spalt-related mutants exhibit Townes-Brocks' syndrome phenotypes. Proc Natl Acad Sci U S A *100*, 10293-10298.

Donohoe, M. E., Zhang, X., McGinnis, L., Biggers, J., Li, E., and Shi, Y. (1999). Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. Mol Cell Biol *19*, 7237-7244.

Estivill-Torrus, G., Pearson, H., van Heyningen, V., Price, D. J., and Rashbass, P. (2002). Pax6 is required to regulate the cell cycle and the rate of progression from symmetrical to asymmetrical division in mammalian cortical progenitors. Development *129*, 455-466.

Fleming, T. P., and Johnson, M. H. (1988). From egg to epithelium. Annu Rev Cell Biol *4*, 459-485.

Franch-Marro, X., and Casanova, J. (2002). spalt-induced specification of distinct dorsal and ventral domains is required for Drosophila tracheal patterning. Dev Biol *250*, 374-382.

Garcia, A. D., Doan, N. B., Imura, T., Bush, T. G., and Sofroniew, M. V. (2004). GFAP-expressing progenitors are the principal source of constitutive neurogenesis in adult mouse forebrain. Nat Neurosci *7*, 1233-1241.

Grant, K., Hanna-Rose, W., and Han, M. (2000). sem-4 promotes vulval cell-fate determination in Caenorhabditis elegans through regulation of lin-39 Hox. Dev Biol *224*, 496-506.

Gribnau, J., Luikenhuis, S., Hochedlinger, K., Monkhorst, K., and Jaenisch, R. (2005). X chromosome choice occurs independently of asynchronous replication timing. J Cell Biol *168*, 365-373.

Gritti, A., Bonfanti, L., Doetsch, F., Caille, I., Alvarez-Buylla, A., Lim, D. A., Galli, R., Verdugo, J. M., Herrera, D. G., and Vescovi, A. L. (2002). Multipotent neural stem cells reside into the rostral extension and olfactory bulb of adult rodents. J Neurosci *22*, 437-445.

Gupta, A., Sanada, K., Miyamoto, D. T., Rovelstad, S., Nadarajah, B., Pearlman, A. L., Brunstrom, J., and Tsai, L. H. (2003). Layering defect in p35 deficiency is linked to improper neuronal-glial interaction in radial migration. Nat Neurosci *6*, 1284-1291.

Handyside, A. H. (1978). Time of commitment of inside cells isolated from preimplantation mouse embryos. J Embryol Exp Morphol *45*, 37-53.

Hanna, L. A., Foreman, R. K., Tarasenko, I. A., Kessler, D. S., and Labosky, P. A. (2002). Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. Genes Dev *16*, 2650-2661.

Hayashi, S., Lewis, P., Pevny, L., and McMahon, A. P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. Mech Dev *119 Suppl 1*, S97-S101.

Hevner, R. F., Shi, L., Justice, N., Hsueh, Y., Sheng, M., Smiga, S., Bulfone, A., Goffinet, A. M., Campagnoni, A. T., and Rubenstein, J. L. (2001). Tbr1 regulates differentiation of the preplate and layer 6. Neuron *29*, 353-366.

Hogan, B., and Tilly, R. (1978). In vitro development of inner cell masses isolated immunosurgically from mouse blastocysts. I. Inner cell masses from 3.5-day p.c. blastocysts incubated for 24 h before immunosurgery. J Embryol Exp Morphol *45*, 93-105.

Hongo, T., Hakuba, A., Shiota, K., and Naruse, I. (2000). Suckling dysfunction caused by defects in the olfactory system in genetic arhinencephaly mice. Biol Neonate *78*, 293-299.

Hunter, K. E., and Hatten, M. E. (1995). Radial glial cell transformation to astrocytes is bidirectional: regulation by a diffusible factor in embryonic forebrain. Proc Natl Acad Sci U S A *92*, 2061-2065.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., and Jaenisch, R. (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet *27*, 31-39.

Jiang, Y., Jahagirdar, B. N., Reinhardt, R. L., Schwartz, R. E., Keene, C. D., Ortiz-Gonzalez, X. R., Reyes, M., Lenvik, T., Lund, T., Blackstad, M., *et al.* (2002a). Pluripotency of mesenchymal stem cells derived from adult marrow. Nature *418*, 41-49.

Jiang, Y., Vaessen, B., Lenvik, T., Blackstad, M., Reyes, M., and Verfaillie, C. M. (2002b). Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain. Exp Hematol *30*, 896-904.

Johnson, M. H., and McConnell, J. M. (2004). Lineage allocation and cell polarity during mouse embryogenesis. Semin Cell Dev Biol *15*, 583-597.

Johnson, M. H., and Ziomek, C. A. (1983). Cell interactions influence the fate of mouse blastomeres undergoing the transition from the 16- to the 32-cell stage. Dev Biol *95*, 211-218.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-191.

Kaplan, M. S., McNelly, N. A., and Hinds, J. W. (1985). Population dynamics of adult-formed granule neurons of the rat olfactory bulb. J Comp Neurol *239*, 117-125.

Kessaris, N., Pringle, N., and Richardson, W. D. (2001). Ventral neurogenesis and the neuron-glial switch. Neuron *31*, 677-680.

Kidder, G. M., and McLachlin, J. R. (1985). Timing of transcription and protein synthesis underlying morphogenesis in preimplantation mouse embryos. Dev Biol *112*, 265-275.

Kiefer, S. M., McDill, B. W., Yang, J., and Rauchman, M. (2002). Murine Sall1 represses transcription by recruiting a histone deacetylase complex. J Biol Chem *277*, 14869-14876.

Kiefer, S. M., Ohlemiller, K. K., Yang, J., McDill, B. W., Kohlhase, J., and Rauchman, M. (2003). Expression of a truncated Sall1 transcriptional repressor is responsible for Townes-Brocks syndrome birth defects. Hum Mol Genet *12*, 2221-2227.

Kohlhase, J., Altmann, M., Archangelo, L., Dixkens, C., and Engel, W. (2000). Genomic cloning, chromosomal mapping, and expression analysis of msal-2. Mamm Genome *11*, 64-68.

Kohlhase, J., Chitayat, D., Kotzot, D., Ceylaner, S., Froster, U. G., Fuchs, S., Montgomery, T., and Rosler, B. (2005). SALL4 mutations in Okihiro syndrome (Duane-radial ray syndrome), acro-renal-ocular syndrome, and related disorders. Hum Mutat *26*, 176-183.

Kohlhase, J., Heinrich, M., Liebers, M., Frohlich Archangelo, L., Reardon, W., and Kispert, A. (2002a). Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. Cytogenet Genome Res *98*, 274-277.

Kohlhase, J., Heinrich, M., Schubert, L., Liebers, M., Kispert, A., Laccone, F., Turnpenny, P., Winter, R. M., and Reardon, W. (2002b). Okihiro syndrome is caused by SALL4 mutations. Hum Mol Genet *11*, 2979-2987.

Kohlhase, J., Schubert, L., Liebers, M., Rauch, A., Becker, K., Mohammed, S. N., Newbury-Ecob, R., and Reardon, W. (2003). Mutations at the SALL4 locus on chromosome 20 result in a range of clinically overlapping phenotypes, including Okihiro syndrome, Holt-Oram syndrome, acro-renal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. J Med Genet *40*, 473-478.

Kohlhase, J., Taschner, P. E., Burfeind, P., Pasche, B., Newman, B., Blanck, C., Breuning, M. H., ten Kate, L. P., Maaswinkel-Mooy, P., Mitulla, B., *et al.* (1999). Molecular analysis of SALL1 mutations in Townes-Brocks syndrome. Am J Hum Genet *64*, 435-445.

Kohlhase, J., Wischermann, A., Reichenbach, H., Froster, U., and Engel, W. (1998). Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. Nat Genet *18*, 81-83.

Koster, R., Stick, R., Loosli, F., and Wittbrodt, J. (1997). Medaka spalt acts as a target gene of hedgehog signaling. Development *124*, 3147-3156.

Kuhnlein, R. P., Frommer, G., Friedrich, M., Gonzalez-Gaitan, M., Weber, A., Wagner-Bernholz, J. F., Gehring, W. J., Jackle, H., and Schuh, R. (1994). spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. Embo J *13*, 168-179.

Kunath, T., Arnaud, D., Uy, G. D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R. L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649-1661.

Lan, Z. J., Xu, X., and Cooney, A. J. (2004). Differential oocyte-specific expression of Cre recombinase activity in GDF-9-iCre, Zp3cre, and Msx2Cre transgenic mice. Biol Reprod *71*, 1469-1474.

Landecker, H. L., Sinclair, D. A., and Brock, H. W. (1994). Screen for enhancers of Polycomb and Polycomblike in Drosophila melanogaster. Dev Genet *15*, 425-434.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet *3*, 662-673.

Long, J. E., Garel, S., Depew, M. J., Tobet, S., and Rubenstein, J. L. (2003). DLX5 regulates development of peripheral and central components of the olfactory system. J Neurosci *23*, 568-578.

Ma, Y., Chai, L., Cortez, S. C., Stopa, E. G., Steinhoff, M. M., Ford, D., Morgan, J., and Maizel, A. L. (2002). SALL1 expression in the human pituitary-adrenal/gonadal axis. J Endocrinol *173*, 437-448.

Mason, I. J., Murphy, D., Munke, M., Francke, U., Elliott, R. W., and Hogan, B. L. (1986). Developmental and transformation-sensitive expression of the Sparc gene on mouse chromosome 11. Embo J *5*, 1831-1837.

Maunoury, R., Robine, S., Pringault, E., Huet, C., Guenet, J. L., Gaillard, J. A., and Louvard, D. (1988). Villin expression in the visceral endoderm and in the gut anlage during early mouse embryogenesis. Embo J *7*, 3321-3329.

McEvilly, R. J., de Diaz, M. O., Schonemann, M. D., Hooshmand, F., and Rosenfeld, M. G. (2002). Transcriptional regulation of cortical neuron migration by POU domain factors. Science *295*, 1528-1532.

Metzger, D., Li, M., and Chambon, P. (2005). Targeted somatic mutagenesis in the mouse epidermis. Methods Mol Biol *289*, 329-340.

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

Mollereau, B., Dominguez, M., Webel, R., Colley, N. J., Keung, B., de Celis, J. F., and Desplan, C. (2001). Two-step process for photoreceptor formation in Drosophila. Nature *412*, 911-913.

Mori, K., Nagao, H., and Yoshihara, Y. (1999). The olfactory bulb: coding and processing of odor molecule information. Science *286*, 711-715.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Neff, A. W., King, M. W., Harty, M. W., Nguyen, T., Calley, J., Smith, R. C., and Mescher, A. L. (2005). Expression of Xenopus XlSALL4 during limb development and regeneration. Dev Dyn *233*, 356-367.

Netzer, C., Rieger, L., Brero, A., Zhang, C. D., Hinzke, M., Kohlhase, J., and Bohlander, S. K. (2001). SALL1, the gene mutated in Townes-Brocks syndrome, encodes a transcriptional repressor which interacts with TRF1/PIN2 and localizes to pericentromeric heterochromatin. Hum Mol Genet *10*, 3017-3024.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Nishinakamura, R., Matsumoto, Y., Nakao, K., Nakamura, K., Sato, A., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Scully, S., Lacey, D. L., *et al.* (2001). Murine homolog of SALL1 is essential for ureteric bud invasion in kidney development. Development *128*, 3105-3115.

Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Novak, A., Guo, C., Yang, W., Nagy, A., and Lobe, C. G. (2000). Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. Genesis *28*, 147-155.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol *21*, 4330-4336.

Onai, T., Sasai, N., Matsui, M., and Sasai, Y. (2004). Xenopus XsalF: anterior neuroectodermal specification by attenuating cellular responsiveness to Wnt signaling. Dev Cell *7*, 95-106.

Ott, T., Parrish, M., Bond, K., Schwaeger-Nickolenko, A., and Monaghan, A. P. (2001). A new member of the spalt like zinc finger protein family, Msal-3, is expressed in the CNS and sites of epithelial/mesenchymal interaction. Mech Dev *101*, 203-207.

Parrish, M., Ott, T., Lance-Jones, C., Schuetz, G., Schwaeger-Nickolenko, A., and Monaghan, A. P. (2004). Loss of the Sall3 gene leads to palate deficiency, abnormalities in cranial nerves, and perinatal lethality. Mol Cell Biol *24*, 7102-7112.

Pavletich, N. P., and Pabo, C. O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809-817.

Qi, Y., Stapp, D., and Qiu, M. (2002). Origin and molecular specification of oligodendrocytes in the telencephalon. Trends Neurosci *25*, 223-225.

Raff, M. (2003). Adult stem cell plasticity: fact or artifact? Annu Rev Cell Dev Biol *19*, 1-22.

Rastan, S., and Robertson, E. J. (1985). X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J Embryol Exp Morphol *90*, 379-388.

Ribeiro, C., Neumann, M., and Affolter, M. (2004). Genetic control of cell intercalation during tracheal morphogenesis in Drosophila. Curr Biol *14*, 2197-2207.

Rice, D. S., and Curran, T. (2001). Role of the reelin signaling pathway in central nervous system development. Annu Rev Neurosci *24*, 1005-1039.

Rideout, W. M., 3rd, Eggan, K., and Jaenisch, R. (2001). Nuclear cloning and epigenetic reprogramming of the genome. Science *293*, 1093-1098.

Risser, J. M., and Slotnick, B. M. (1987). Nipple attachment and survival in neonatal olfactory bulbectomized rats. Physiol Behav *40*, 545-549.

Robson, P. (2005). Genome wide mapping of Oct4-DNA interactions in embryonic stem cells. mouse molecular genetics meeting, EMBL Heidelberg.

Rodriguez, C. I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A. F., and Dymecki, S. M. (2000). High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet *25*, 139-140.

Rodriguez, T. A., Sparrow, D. B., Scott, A. N., Withington, S. L., Preis, J. I., Michalicek, J., Clements, M., Tsang, T. E., Shioda, T., Beddington, R. S., and Dunwoodie, S. L. (2004).

Cited1 is required in trophoblasts for placental development and for embryo growth and survival. Mol Cell Biol *24*, 228-244.

Rossant, J. (2004). Lineage development and polar asymmetries in the peri-implantation mouse blastocyst. Semin Cell Dev Biol *15*, 573-581.

Rossant, J., and Lis, W. T. (1979). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. Dev Biol *70*, 255-261.

Royet, J. P., Distel, H., Hudson, R., and Gervais, R. (1998). A re-estimation of the number of glomeruli and mitral cells in the olfactory bulb of rabbit. Brain Res *788*, 35-42.

Santos, F., and Dean, W. (2004). Epigenetic reprogramming during early development in mammals. Reproduction *127*, 643-651.

Sato, A., Kishida, S., Tanaka, T., Kikuchi, A., Kodama, T., Asashima, M., and Nishinakamura, R. (2004). Sall1, a causative gene for Townes-Brocks syndrome, enhances the canonical Wnt signaling by localizing to heterochromatin. Biochem Biophys Res Commun *319*, 103-113.

Sato, A., Matsumoto, Y., Koide, U., Kataoka, Y., Yoshida, N., Yokota, T., Asashima, M., and Nishinakamura, R. (2003). Zinc finger protein sall2 is not essential for embryonic and kidney development. Mol Cell Biol *23*, 62-69.

Schneider, A., Zhang, Y., Guan, Y., Davis, L. S., and Breyer, M. D. (2003). Differential, inducible gene targeting in renal epithelia, vascular endothelium, and viscera of Mx1Cre mice. Am J Physiol Renal Physiol *284*, F411-417.

Schwenk, F., Baron, U., and Rajewsky, K. (1995). A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. Nucleic Acids Res *23*, 5080-5081.

Smith, A. G. (2001). Embryo-derived stem cells: of mice and men. Annu Rev Cell Dev Biol *17*, 435-462.

Solter, D., and Knowles, B. B. (1975). Immunosurgery of mouse blastocyst. Proc Natl Acad Sci U S A *72*, 5099-5102.

Song, H., Stevens, C. F., and Gage, F. H. (2002a). Astroglia induce neurogenesis from adult neural stem cells. Nature *417*, 39-44.

Song, H. J., Stevens, C. F., and Gage, F. H. (2002b). Neural stem cells from adult hippocampus develop essential properties of functional CNS neurons. Nat Neurosci *5*, 438-445.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet *21*, 70-71.

Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Sunaga, S., Maki, K., Komagata, Y., Ikuta, K., and Miyazaki, J. I. (1997). Efficient removal of loxP-flanked DNA sequences in a gene-targeted locus by transient expression of Cre recombinase in fertilized eggs. Mol Reprod Dev *46*, 109-113.

Sweetman, D., Smith, T., Farrell, E. R., Chantry, A., and Munsterberg, A. (2003). The conserved glutamine-rich region of chick csal1 and csal3 mediates protein interactions with other spalt family members. Implications for Townes-Brocks syndrome. J Biol Chem *278*, 6560-6566.

Tada, T., and Tada, M. (2001). Toti-/pluripotential stem cells and epigenetic modifications. Cell Struct Funct *26*, 149-160.

Takahashi, T., Goto, T., Miyama, S., Nowakowski, R. S., and Caviness, V. S., Jr. (1999). Sequence of neuron origin and neocortical laminar fate: relation to cell cycle of origin in the developing murine cerebral wall. J Neurosci *19*, 10357-10371.

Takasato, M., Osafune, K., Matsumoto, Y., Kataoka, Y., Yoshida, N., Meguro, H., Aburatani, H., Asashima, M., and Nishinakamura, R. (2004). Identification of kidney mesenchymal genes by a combination of microarray analysis and Sall1-GFP knockin mice. Mech Dev *121*, 547-557.

Tam, P. P., and Rossant, J. (2003). Mouse embryonic chimeras: tools for studying mammalian development. Development *130*, 6155-6163.

Tamamaki, N., Nakamura, K., Okamoto, K., and Kaneko, T. (2001). Radial glia is a progenitor of neocortical neurons in the developing cerebral cortex. Neurosci Res *41*, 51-60.

Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.

Taraviras, S., Monaghan, A. P., Schutz, G., and Kelsey, G. (1994). Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. Mech Dev *48*, 67-79.

Toker, A. S., Teng, Y., Ferreira, H. B., Emmons, S. W., and Chalfie, M. (2003). The Caenorhabditis elegans spalt-like gene sem-4 restricts touch cell fate by repressing the selector Hox gene egl-5 and the effector gene mec-3. Development *130*, 3831-3840.

Townes , P. L., Brooks,E. R. (1972). Hereditary Syndrome of imperforate Anus with Hand, Foot and Ear Anomalies. Journal of Pediatry *8*, 321-326.

Tsonis, P. A., and Goetinck, P. F. (1988). The Drosophila homoeotic gene spalt is structurally related to collagen alpha 1(IV) chain. Coll Relat Res *8*, 451-452.

Wagner-Bernholz, J. T., Wilson, C., Gibson, G., Schuh, R., and Gehring, W. J. (1991). Identification of target genes of the homeotic gene Antennapedia by enhancer detection. Genes Dev *5*, 2467-2480.

Weissman, I. L., Anderson, D. J., and Gage, F. (2001). Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. Annu Rev Cell Dev Biol *17*, 387-403.

Wiley, L. M., and Pedersen, R. A. (1977). Morphology of mouse egg cylinder development in vitro: a light and electron microscopic study. J Exp Zool *200*, 389-402.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Wolpert, L. (2002). Limb patterning: reports of model's death exaggerated. Curr Biol *12*, R628-630.

Wolpert, L., Tickle, C., and Sampford, M. (1979). The effect of cell killing by x-irradiation on pattern formation in the chick limb. J Embryol Exp Morphol *50*, 175-193.

Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2005). Nanog expression in mouse germ cell development. Gene Expr Patterns *5*, 639-646.

Yokoo, H., Nobusawa, S., Takebayashi, H., Ikenaka, K., Isoda, K., Kamiya, M., Sasaki, A., Hirato, J., and Nakazato, Y. (2004). Anti-human Olig2 antibody as a useful immunohistochemical marker of normal oligodendrocytes and gliomas. Am J Pathol *164*, 1717-1725.

Zechner, D., Fujita, Y., Hulsken, J., Muller, T., Walther, I., Taketo, M. M., Crenshaw, E. B., 3rd, Birchmeier, W., and Birchmeier, C. (2003). beta-Catenin signals regulate cell growth and the balance between progenitor cell expansion and differentiation in the nervous system. Dev Biol *258*, 406-418.