

Functional analysis of murine *Sall4*

Dissertation

**zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen**



vorgelegt von

Lina Malinouskaya

aus Minsk, Republic of Belarus

Göttingen 2005

D7

Referent: Prof. Dr. W. Engel

Korreferentin: PD Dr. S. Hoyer-Fender

Tag der mündlichen Prüfungen: 18.01.06

I dedicate this work to my parents

CONTENTS

ABBREVIATIONS	V
1. INTRODUCTION	1
1.1 SALL family of proteins.....	1
1.2 <i>SALL</i> genes and human syndromes	5
1.3 Mice as model organisms for human diseases	8
1.4 Gene trap approach.....	9
1.5 Mouse models for Sall gene function.....	10
1.6 Aims of the study.....	11
2. MATERIALS AND METHODS	12
2.1 Materials	12
2.1.1 Chemicals	12
2.1.2 Solutions, buffers and media	15
2.1.3 Laboratory materials.....	17
2.1.4 Sterilisation of solutions and equipments.....	18
2.1.5 Media, antibiotics and agar-plates.....	18
2.1.5.1 Media for bacteria.....	18
2.1.5.2 Media for cell culture	19
2.1.5.3 Antibiotics	19
2.1.5.4 IPTG/X-Gal LB agar	20
2.1.6 Bacterial strains	20
2.1.7 Plasmids.....	20
2.1.8 Synthetic oligonucleotide primers.....	20
2.1.9 Eukaryotic cell lines	23
2.1.10 Mouse strains.....	23
2.1.11 Antibodies.....	23
2.1.12 Kits	24
2.1.13 Instruments	24
2.2 Methods	25
2.2.1 Isolation of nucleic acids.....	25
2.2.1.1 Small-scale isolation of plasmid DNA (mini-prep).....	25
2.2.1.2 Large-scale preparation of plasmid DNA (midi-prep)	26
2.2.1.3 Isolation of genomic DNA from tissue samples.....	27
2.2.1.4 Isolation of genomic DNA from mouse embryos	27
2.2.1.5 Isolation of genomic DNA from preimplantation embryos.....	28
2.2.1.6 Isolation of genomic DNA from cultured blastocysts.....	28
2.2.1.7 Isolation of total RNA from tissue samples and cultured cells	28

2.2.1.8 Determination of the nucleic acid concentration.....	29
2.2.2 Nucleic acid electrophoresis.....	29
2.2.3 Isolation of DNA fragments from agarose gel.....	30
2.2.4 Enzymatic modifications of DNA.....	31
2.2.4.1 Digestion of DNA using restriction enzymes.....	31
2.2.4.2 Dephosphorylation of 5' ends of DNA.....	31
2.2.4.3 Ligation of DNA fragments.....	31
2.2.4.4 TA - cloning.....	32
2.2.4.5 Non-radioactive dye terminator cycle sequencing.....	32
2.2.5 Transformation of competent bacteria.....	33
2.2.6 Polymerase Chain Reaction (PCR).....	33
2.2.6.1 PCR amplification of DNA fragments.....	34
2.2.6.2 Genotyping PCR.....	34
2.2.6.3 Reverse transcription PCR (RT-PCR).....	36
2.2.7 Real time RT-PCR using SYBR Green I dye.....	36
2.2.8 Protein biochemical methods.....	38
2.2.8.1 Isolation of total protein.....	38
2.2.8.2 Electrophoresis of proteins.....	38
2.2.8.3 Transfer of protein onto nitrocellulose membranes.....	39
2.2.8.4 Staining of polyacrylamide gels.....	39
2.2.8.5 Western blotting.....	39
2.2.8.6 Generation of polyclonal antibody.....	40
2.2.8.6.1 Sequence analysis of <i>Sall4</i> protein.....	40
2.2.8.6.2 Amplification and cloning of <i>Sall4</i> fusion construct.....	41
2.2.8.6.3 Expression of recombinant proteins in the pET vector.....	41
2.2.8.6.4 Purification of GST fusion proteins.....	42
2.2.8.6.5 Immunisation of rabbits.....	42
2.2.8.6.6 Affinity purification of polyclonal antibody.....	42
2.2.8.6.6.1 Ligand coupling.....	43
2.2.8.6.6.2 Washing and deactivation.....	43
2.2.8.6.6.3 Affinity purification of antiserum.....	43
2.2.8.6.7 Protein G purification of antibody.....	44
2.2.9 Eukaryotic cell culture methods.....	44
2.2.9.1 Cell culture conditions.....	44
2.2.9.2 Trypsinisation of eukaryotic cells.....	44
2.2.9.3 Cryopreservation and thawing of eukaryotic cells.....	44
2.2.9.4 Transient transfection of the cells.....	45
2.2.9.5 Immunocytochemical analysis.....	45
2.2.10 Histological techniques.....	46
2.2.10.1 Tissue preparation for paraffin-embedding.....	46
2.2.10.2 Sections of the paraffin block.....	46
2.2.10.3 Staining of the histological sections (Hematoxylin-Eosin staining).....	46
2.2.10.4 Immunohistochemistry.....	47
2.2.11 Techniques for recovery and culture of preimplantation embryos.....	47
2.2.11.1 Superovulation.....	47
2.2.11.2 Preparation of blastocyst stage embryos.....	47
2.2.11.3 In vitro culture of blastocyst stage embryos.....	48
2.2.12 Whole mount <i>in situ</i> hybridisation.....	48

Contents

2.2.12.1 Embryo dissection	48
2.2.12.2 Pre-hybridisation treatment and hybridisation of the embryos	50
2.2.12.3 Detection of hybridisation signals	50
2.2.12.4 Colour development and storage of the embryos	51
2.2.13 X-gal staining	51
2.2.13.1 LacZ staining of mouse embryos.....	51
2.2.13.2 X-gal staining of blastocyst stage embryos	52
2.2.14 BrdU staining.....	52
2.2.15 Computer analysis	53
3. RESULTS.....	54
3.1 Gene trap cell line.....	54
3.1.1 Identification of <i>Sall4</i> gene trap cell line	54
3.1.2 Localisation of the gene trap vector integration site in <i>Sall4</i> gene.....	55
3.2 Characterisation of <i>Sall4</i> splicing forms	57
3.3 Expression analysis of <i>Sall4</i>	59
3.3.1 Expression of <i>Sall4</i> gene in preimplantation embryos	59
3.3.2 Expression of <i>Sall4</i> gene in postimplantation embryos.....	60
3.3.3 RT-PCR analysis of <i>Sall4</i> expression during embryonic development	62
3.3.4 Expression of splicing variants in <i>Sall4</i> mutant mice	63
3.4 Phenotype of <i>Sall4</i> mutant mice.....	65
3.4.1 Early embryonic lethality of <i>Sall4</i> mice.....	65
3.4.2 Phenotype of <i>Sall4</i> homozygous embryos	67
3.4.3 Histological analysis of <i>Sall4</i> embryos	67
3.4.4 Analysis of cellular proliferation in <i>Sall4</i> embryos.....	69
3.5 <i>Sall4</i> regulates the expression of <i>Pax3</i> and T-box genes	70
3.6 <i>Sall4</i> antibody generation	72
3.7 <i>Sall4</i> and <i>Sall1</i> interaction studies.....	74
3.7.1 Interaction on the genetic level.....	74
3.7.2 Co-localization of <i>Sall4</i> with truncated <i>Sall1</i> protein in F9 cells.....	76
4. DISCUSSION.....	78
4.1 <i>Sall4</i> expression analysis.....	79
4.2 <i>Sall4</i> is required for early heart development.....	81
4.3 Hypomorphic phenotype of <i>Sall4</i> gene trap mice	82
4.4 <i>Sall4</i> homozygous mutant mice die during embryonic development.....	84
4.5 <i>Sall4</i> regulation during embryonic development.....	85

Contents

4.6 Future perspectives.....	89
5. SUMMARY	92
6. REFERENCES	94
ACKNOWLEDGMENTS	103
CURRICULUM VITAE	105

ABBREVIATIONS

ABI	Applied Biosystem Instrument
AP	Alkaline Phosphatase
APS	Ammonium peroxydisulfate
ATP	Adenosintriphosphate
BCIP	5-Brom-4-Chlor-3-Indolyl-Phosphate
bp	base pair
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
°C	degree Celsius
cDNA	complementary DNA
Cy3	indocarbocyanine
DAPI	Diamidino-2-phenylindole dihydrochloride
dATP	Desoxyriboadenosintriphosphate
dCTP	Desoxyribocytosintriphosphate
DEPC	Diethylpyrocarbonate
dH ₂ O	distilled water
DIG	Digoxygenin
DMEM	DULBECCO's Modified Eagles Media
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dnase	Deoxyribonuclease
dNTP	Deoxynucleotidetriphosphate
DRRS	Duane-radial ray syndrome
dT	Deoxythymidinate
DTT	Dithiothreitol
E	embryonic day
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylenglycol-bis(β-Aminoethylether)-Tetraacetat
ES	Embryonic stem

Abbreviations

FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
g	gram
g	gravity
GGTC	German Gene Trap Consortium
GSDB	Goat serum dilution buffer
GST	Glutathione S-transferase
h (rs)	hour(s)
HEPES	N-(-hydroxymethyl)piperazin,N'-3-propansulfoneacid
HOS	Holt-Oram syndrome
IPTG	Isopropyl- β -thiogalactopyranoside
kb	kilobase
LB	Luria-Bertrani
LIF	Recombinant leukaemia inhibitory factor
M	molarity
MAB	Maleic acid buffer
Mb	Mega base pair
mg	milligram
min	minute
ml	milliliter
μ l	microliter
μ m	micrometer
MOPS	3-[N-Morpholino]-Propanesulfate
mRNA	messenger ribonucleic acid
NaAc	Sodium acetate
NBT	Nitro-blue tetrazolium
NCBI	National Center for Biotechnology Information
<i>Neo</i>	<i>Neomycin</i>
ng	nanogram
nm	nanometer
NTMT	Natriumchlorid-Tris-Magnesiumchlorid solution with Tween-20
NTP	nucleotidetriphospate

Abbreviations

OD	Optical density
ORF	Open Reading Frame
Pa	Pascal
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate buffer saline
PBT	Phosphate buffersaline + Tween
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
pH	prepondirance of hydrogen ions
PMSF	Phenylmethylsulfonyl fluoride
RACE	Rapid Amplification of cDNA Ends
RNA	Ribonucleic acid
RNAi	RNA interference
Rnase	Ribonuclease
shRNA	short hairpin RNA
rpm	revolution per minute
RT-PCR	reverse transcriptase-PCR
SDS	Sodium dodecylsulfate
sec	second
SSC	Standard Saline Citrat buffer
SV 40	Simian Virus 40
<i>Taq</i>	<i>Thermus aquaticus</i>
TBE	Tris-Borate-EDTA-Electrophoresis buffer
TBS	Townes-Brocks-Syndrom
TE	Tris-EDTA buffer
TEMED	Tetramethylethylene diamine
Tris	Trihydroxymethylaminomethane
Tween-20	Polyoxyethylensorbitan-Monolaureate
U	Unit
UV	Ultra violet
V	Voltage

Abbreviations

w/v	weight/volume
X-Gal	5-bromo-4-chloro-3-indolyl- β -galactosidase

Symbols of amino acids

A	Ala	Alanine	M	Met	Methionine
B	Asx	Asparagine	N	Asn	Asparagine
C	Cys	Cysteine	P	Pro	Proline
D	Asp	Aspartic acid	Q	Gln	Glutamine
E	Glu	Glutamic acid	R	Arg	Arginine
F	Phe	Phenylalanine	S	Ser	Serine
G	Gly	Glycine	T	Thr	Threonine
H	His	Histidine	V	Val	Valine
I	Ile	Isoleucine	W	Trp	Tryptophan
K	Lys	Lysine	Y	Tyr	Tyrosine
L	Leu	Leucine	Z	Glx	Glutamine

Symbols of nucleic acid

A	Adenosine
C	Cytidine
G	Guanosine
T	Thymidine
U	Uridine

1. INTRODUCTION

1.1 SALL family of proteins

Members of the conserved *spalt* (*sal*) gene family encode zinc finger proteins and have been implicated in multiple developmental processes. The region-specific homeotic gene *spalt* (*sal*) plays an important role in specifying the head and tail regions of the *Drosophila melanogaster* embryo (Jürgens, 1988). SAL function also contributes to the development of the wing imaginal discs and wing patterning (de Celis, 1998), participates in the development of sensory organs in the thorax (de Celis, 1999), required for the specification and establishment of planar cell polarity in the *D. melanogaster* eye (Domingos et al., 2004) and suppresses the molecular pathway that establishes tracheal development in the *D. melanogaster* embryo (Kühnlein et al., 1996). The SAL protein is characterised by three C₂H₂ double zinc finger motifs (Fig. 1.1).



Figure 1.1: Schematic representation of SAL protein structure. A typical SAL protein contains at least three widely spaced C₂H₂ double zinc finger (DZF) motifs (yellow) with single C₂H₂ zinc finger attached to the second domain (blue). In addition, all vertebrate SAL proteins have an additional N-terminal C₂HC zinc finger (red). Glutamine rich region marked in blue. Zinc finger domains represented as ovals.

An additional C₂H₂ zinc finger motif is attached to the second double zinc finger domain (Kühnlein et al., 1994), and all vertebrate SAL proteins have an additional N-terminal C₂HC zinc finger. Each C-terminal zinc finger of the double zinc finger contains a conserved sequence called “SAL-box” which consists of eight conserved amino acids (FTTKGNLK). The SAL protein contains also some regions rich in glutamines, alanines, prolines and serines. C₂H₂ zinc finger domains are composed of 21 amino-acid residues including two conserved Cys and two conserved His residues which are coordinated by a zinc ion (Fig. 1.2). Zinc fingers have been found to bind to both RNA and DNA motives.

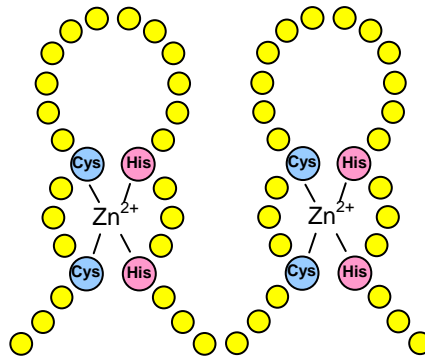


Figure 1.2: Schematic representation of series of two zinc fingers #-X-C-X(1-5)-C-X3-#-X5-#-X2-H-X(3-6)-[H/C]. Where X can be any amino acid, and numbers in brackets indicate the number of residues. The final position can be either His or Cys.

The double zinc finger motifs are highly conserved among the *D. melanogaster* spalt and spalt-related proteins and vertebrate SALL (SAL-like) proteins. The strong conservation among SAL family members has allowed to identify the homologous genes in various species. One of the medaka (*Oryzias latipes*) homologues of spalt was identified by Köster et al. (1997). Medaka *sal* is expressed in the eye, CNS, excretory system, the somites and the developing fins during embryogenesis and is controlled by the sonic hedgehog pathway (Köster et al., 1997). Three homologues of *sal* family were isolated in zebra fish (*Danio rerio*). Zebra fish *sall1a* and *sall1b* are orthologous to other vertebrate *sall1* genes and zebra fish *sall3* is orthologous to other *sall3* vertebrate genes, except *Xenopus sall3* (Camp et al., 2003). *sall1a* and *sall3* are expressed in various regions of the CNS, including in primary motor neurons. Outside of the CNS, *sall1a* expression was observed in the otic vesicle (ear), heart and in a discrete region of the pronephric ducts. Like other vertebrate *sal* genes, zebrafish *sal* genes have a role in neural development (Camp et al., 2003). Zebrafish *sall4* homologue is listed in the database, but no data are available on the expression and function. *Xenopus laevis Xsal-1* encodes a protein with the structure similar to the *Drosophila spalt*, but contains additional double zinc finger domain (Holleman et al., 1996). During early development *Xsal-1* expression is confined the midbrain, hindbrain, and spinal cord. *Xsal-1* expression was also found in the developing limb buds (Holleman et al., 1996). The pattern of expression suggests that *Xsal-1* takes part in neuronal cell specification (Holleman et al., 1996). Alternatively spliced *Xsal-3* transcripts give rise to RNAs coding either two or three double zinc fingers, and the longer form is expressed maternally (Onuma et al., 1999). *Xsal-3* is expressed in the neural tube, the mandibular, hyoid, and branchial arch and the pronephric

duct (Onuma et al., 1999). *XsalF* gene plays an essential role for the forebrain/midbrain determination and functions by antagonising canonical Wnt signalling pathway in *Xenopus* (Onai et al., 2004). A new member of the *Xenopus sal* family was recently identified (*XISALL4*) by a competitive subtractive hybridization screen of genes expressed in *Xenopus* hindlimb regeneration blastemas (Neff et al., 2005). The expression pattern of *XISALL4* transcripts was similar to *Xsal-3* during embryonic development. *XISALL4* may regulate normal forelimb and hindlimb development and is necessary for the epimorphic limb regeneration (Neff et al., 2005).

In the chick three members of the spalt family have been described so far, *csal1*, *csal3* and *csal4* (Sweetman et al., 2005, Farrell et al., 2001, Barembaum and Bronner-Fraser, 2004). During early embryonic development *csal1* expression was found in ectoderm, involuting mesoderm and presegmented mesoderm. *csal1* expression was also observed in the heart, the neural plate and the pharynx (Sweetman et al., 2005). At later developmental stages *csal1* transcripts were detected in the CNS, tail bud and developing limb buds and the *csal1* expression is regulated by members of the FGF and Wnt families (Farrell and Munsterberg, 2000). *csal3* is expressed in the nervous system, developing kidney, cloaca and limb bud (Farrell et al., 2001). *csal4* is expressed in the neural tube, migrating neural crest and branchial arches and might be involved in the neural crest cell determination (Barembaum and Bronner-Fraser, 2004).

The mouse *Sall* as well as the human *SALL* families contain four members each. The murine homologue of the human *SALL1* gene was isolated by Buck et al., 2000. *Sall1* consists of three exons and *SALL1* protein shares a high (89%) sequence homology with human *SALL1*. Northern blot analysis of *Sall1* showed strong expression of *Sall1* transcript in kidney. A lower amount of the transcript was found in brain, liver, testis, and heart tissue (Buck et al., 2000). During embryonic development prominent *Sall1* expression was found in the developing brain and the limbs. Other sites of expression include the meso- and metanephros, lens, olfactory bulbs, heart, primitive streak and the genital tubercle (Buck et al., 2001). *Sall2* consists of three exons (exons 1, 1a and 2) with 1 and 1a encoding alternative translational start sites. *Sall2* encodes only three double zinc finger domains, the most carboxyterminal of which only distantly resembles spalt-like zinc fingers (Kohlhase et al., 2000b). *Sall2* is expressed throughout embryonic development but also in adult tissues, predominantly in brain (Kohlhase et al., 2000b). *Sall3* is expressed in the developing neuroectoderm of the brain, the

inner ear, the spinal cord, the testis, ovaries and kidneys. A weaker and transient expression was also observed in the branchial arches, the notochord, the limb buds and the heart (Ott et al., 1996). The mouse *Sall3* locus is an epigenetic hotspot of aberrant DNA methylation associated with placentomegaly of cloned mice (Ohgane et al., 2004).

SAL family members represent a group of evolutionary conserved zinc finger transcription factors which have been identified in many species and appear to be critical developmental regulators. Despite strong conservation among the family and partial overlapping expression pattern members of SALL family have some common and some unique functions during development. The summary of the known and predicted protein sequences of SALL family is presented on the figure 1.3.

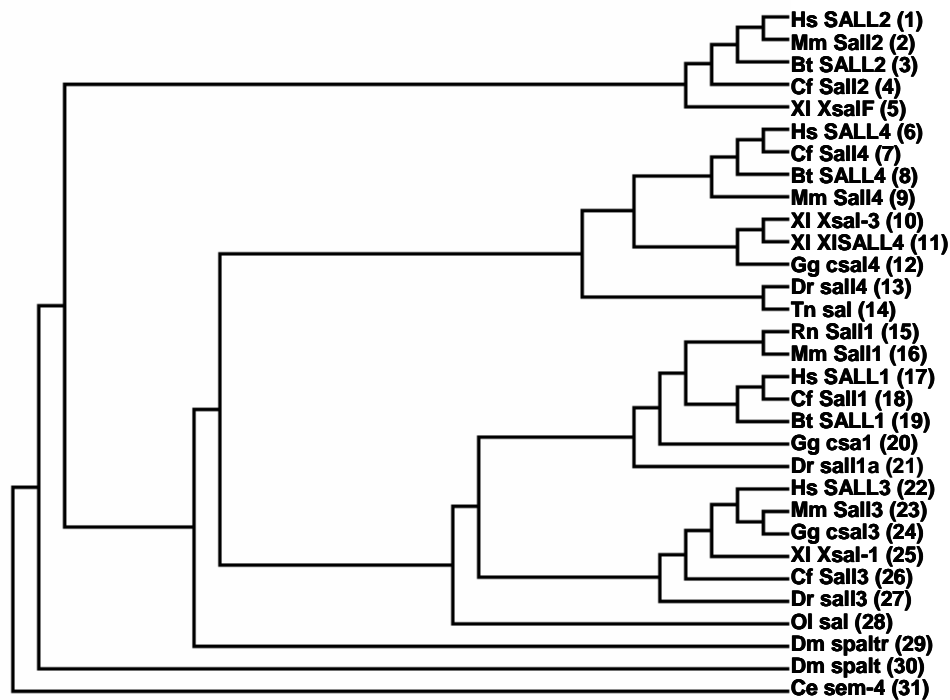


Figure 1.3: SAL protein family maximum parsimony phylogenetic tree. The phylogeny analysis was done by the ClustalW and TreeTop software. Species abbreviations: Bt, *Bos taurus*, Ce, *Caenorhabditis elegans*, Cf, *Canis familiaris*, Dm, *Drosophila melanogaster*, Dr, *Danio rerio*, Gg, *Gallus gallus*, Hm, *Homo sapiens*, Mm, *Mus musculus*, Ol, *Oryzalis latipes*, Rn, *Rattus norvegicus*, Tn, *Tetraodon nigroviridis*, Xl, *Xenopus laevis*. Accession numbers of the proteins: 1-NP_005398, 2-NP_056587, 3-XP_605967.2, 4-XP_539685.2, 5-AAS79482, 6-NP_065169, 7-XP_543055, 8-XP_614880.1, 9-NP_780512, 10-BAA85900, 11-AAP97930, 12-AAR01968, 13-XP_701344, 14-CAG03873, 15-XP_226329, 16-NP_067365, 17-NP_002959, 18-XP_544410, 19-XP_592327, 20-NP_990038, 21-XP_683406, 22-NP_741996, 23-NP_840064, 24-NP_989978, 25-AAC42233, 26-XP_848821, 27-XP_696441, 28-AAB51127, 29-NP_523548, 30-P39770, 31-NP_491997.

1.2 *SALL* genes and human syndromes

The human *SALL* protein family includes four members. The gene *SALL1* is localised on chromosome 16q12.1 and encodes a typical SAL-like protein with four double zinc finger domains of the SAL-type and an additional N-terminal C₂HC zinc finger motive (Kohlhase et al., 1996). Strong expression of *SALL1* was detected in adult kidney, brain and liver (Kohlhase et al., 1996). *SALL1* localises to heterochromatic foci within the nucleus and acts as a strong transcriptional repressor in mammalian cells (Netzer et al., 2001). Mutations in *SALL1* gene cause Townes-Brocks syndrome, a rare autosomal dominantly inherited malformation syndrome characterized by imperforate anus, dysplastic ears, and preaxial polydactyly and/or triphalangeal thumbs (Kohlhase et al., 1998). Less commonly TBS patients also show malformations of the feet, the kidneys, and the heart as well as hearing loss, impaired renal function, and mental retardation (Kohlhase et al., 1999a, Powell and Michaelis, 1999). The finding of disease-causing mutations of *SALL1* in TBS suggested an important developmental regulatory function for *SALL1* (Kohlhase et al., 1998). Most mutations causing TBS are clustered in the N-terminal third of the *SALL1* coding region and result in the production of truncated proteins containing only one or none of the C₂H₂ domains and the N-terminal transcriptional repressor domain of *SALL1* (Botzenhart et al., 2005). Two other members of SAL-family, *SALL2* (Kohlhase et al., 1996) and *SALL3* (Kohlhase et al., 1999b), have not been found yet to cause disease phenotype, but an involvement of *SALL2* has been discussed in ovarian carcinoma (Li et al., 2001) and *SALL3* haploinsufficiency was thought to contribute to the phenotype of patients with terminal deletions of chromosome 18q (Kohlhase et al., 1999b).

Mutations in the *SALL4* gene result in Okinhiro/Duane-radial ray syndrome (DRRS) syndrome (Kohlhase et al., 2002b), an autosomal dominant condition characterised by radial ray defects and Duane anomaly (Fig 1.4). Other abnormalities reported in this condition are anal, renal, cardiac, ear, foot malformations and hearing loss (Kohlhase et al., 2002b, Al-Baradie et al., 2002, Kohlhase et al., 2003, Borozdin et al., 2004a). *SALL4* mutations may also cause acrorenal-ocular syndrome (AROS), which differs from DRRS by the presence of structural eye anomalies, and phenotypes similar to thalidomide embryopathy and Holt-Oram syndrome (HOS) (Borozdin et al., 2004b).

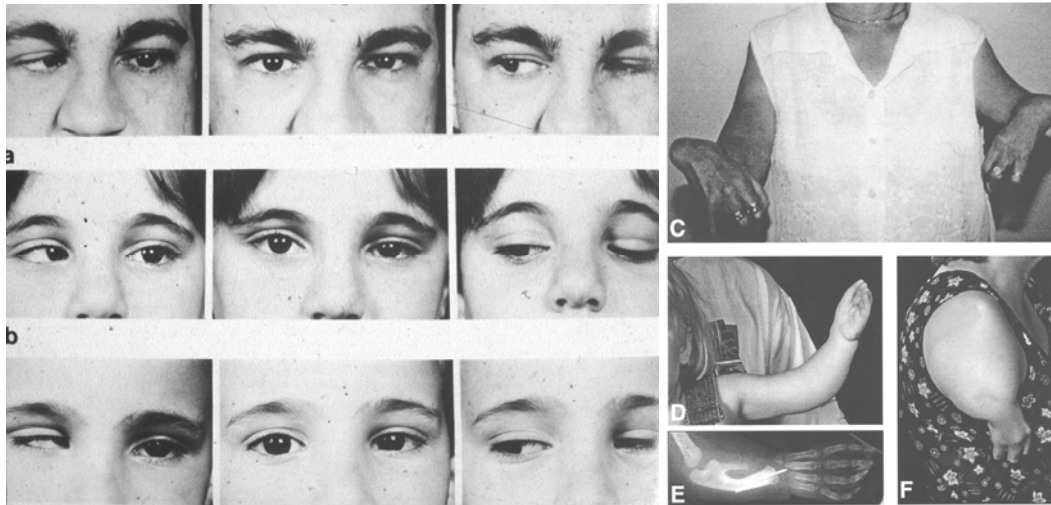


Figure 1.4: Clinical features of Okihiro syndrome patients. (A, B) Family demonstrating Duane retraction anomaly. The patients are unable to abduct the left eye unilaterally and on adducting the same eye there is retraction of the globe and narrowing of the palpebral fissure. (C–F) Limb malformations in three affected individuals. (C) Patient has bilateral radial club hands with absence of the thumbs. (D) Patient has a normal humerus but short forearm and absence of the thumb. (E) There is radiological fusion of the ulna and radius, which are foreshortened. The first metacarpals are absent. (F) Patient shows a very rudimentary upper limb, with absence of the thumb and forearm and a short humerus (pictures from: Kohlhase et al. (2002a). Okihiro syndrome is caused by *Sall4* mutations. *Hum Mol Genet* (11) published by Oxford University Press).

The human *SALL4* gene is localised on chromosome 20q13.13-q13.2 (Kohlhase et al., 2002a, Al-Baradie et al., 2002). *SALL4* consists of 4 exons with an open reading frame of 3159 bp. It encodes a protein with three C₂H₂ double zinc finger domains second of which has a single C₂H₂ zinc finger attached at its carboxy-terminal end, as well as an N-terminal C₂HC zinc finger motif typical for vertebrate SAL-like proteins (Fig 1.5). It expressed in several human tissues including testes, ovary and spleen as determined by Kohlhase et al. (2002a). It also expressed in the embryonal carcinoma cell lines H12.1 and 2102EP as shown by Kohlhase et al. (2002a). 17 of the 22 *SALL4* mutations known to date are located in exon 2, and five are located in exon 3 (Fig 1.5). These are nonsense mutations, short duplications, and short deletions. All of the mutations lead to preterminal stop codons and are thought to cause the phenotype via haploinsufficiency (Kohlhase et al., 2005).

Introduction

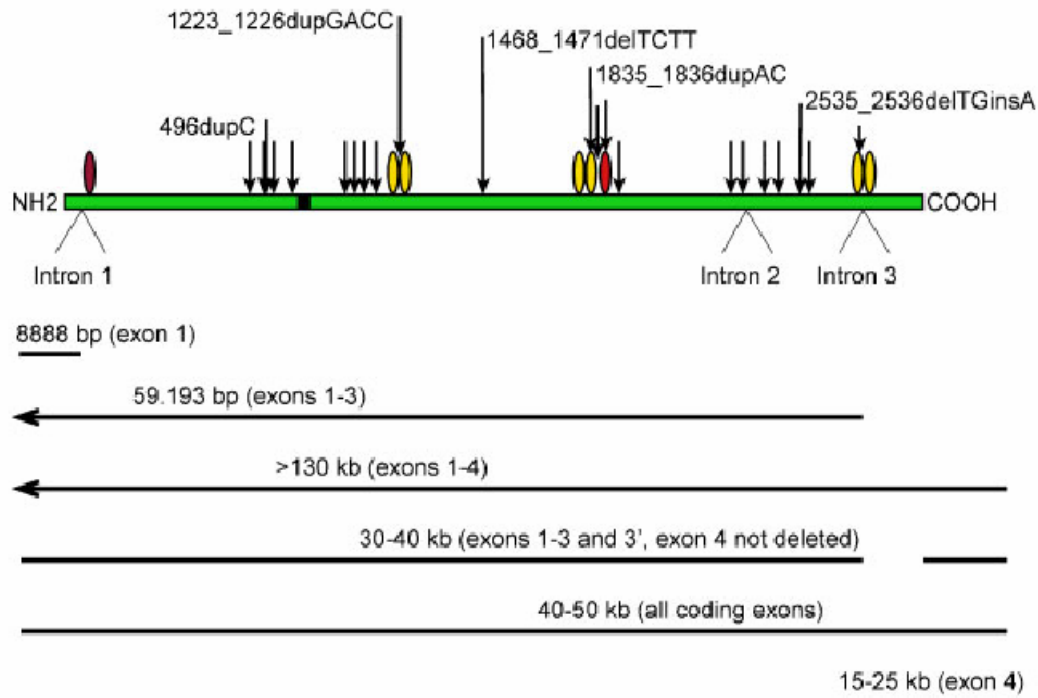


Figure 1.5: Schematic representation of the SALL4 protein and localization of the mutations and deletions identified to date. Zinc fingers are indicated as oval symbols. SALL4 encodes three C₂H₂ double zinc finger domains distributed over the protein. A single C₂H₂ domain is attached to the second double zinc finger. At the aminotermminus, a single C₂HC domain is found. Horizontal bars indicate the positions of bigger deletions with respect to the coding exons. The interrupted bar indicates that this deletion spares exon 4 but continues farther 30. Note that the mutations c.496 dupC and c.2593C4T have each been found in two unrelated families. All other mutations have been found only once. The newly described mutations are labelled, and the unlabeled arrows represent previously described mutations. Positions of the introns are indicated. Numbers refer to the amino acid sequence (1053 aa). (pictures from: Kohlhase et al. (2005). SALL4 mutations in Okihiro syndrome (Duane-radial ray syndrome), acro-renal-ocular syndrome, and related disorders. *Hum Mutat* (26) published by Wiley-Liss, Inc.).

The only missense point mutation of the *SALL4* gene known to date, (c.2663A→G, p.H888R) positioned within the first C₂H₂ zinc finger of the C-terminal double zinc finger motif in the *SALL4* gene, was investigated by a molecular modelling study on the wild type and mutated zinc finger domains (Freder et al., 2004). The modelled wild type and mutated forms of zinc finger motif of the *SALL4* did not display significant structural differences caused by steric strain or charge of R888. However, more significant differences were predicted in their binding affinities to DNA. Calculated higher DNA binding affinity of the mutated zinc finger

could alter activator/repressor function or the cause for erroneous targeting of a different DNA sequence (Freecer et al., 2004).

Up to now four murine homologues of human SALL family have been reported (Buck et al., 2000, Kohlhase et al., 2000b, Ott et al., 1996, Kohlhase et al., 2002b). The *Sall4* gene was isolated based on the high sequence homology (72%) to human *SALL4*. During embryonic development *Sall4* expression was found in the brain, branchial arches and limbs suggesting an important role of *Sall4* during development (Kohlhase et al., 2002b).

1.3 Mice as model organisms for human diseases

Among the model organisms, the mouse offers particular advantages for the study of human biology and disease.

- The mouse is a mammal, and its development, body plan, physiology, behaviour and diseases have much in common with those of humans
- Almost all (99%) mouse genes have homologues in humans
- The mouse genome supports targeted mutagenesis in specific genes by homologous recombination in embryonic stem (ES) cells, allowing genes to be altered efficiently and precisely (Austin et al., 2004)

Many models for genetic diseases were established through spontaneous mutations. Another way to produce mouse lines for the different genetic diseases is by targeted disruption of a specific gene which is responsible for the human syndrome. Various methods can be used to create mutated alleles, including gene targeting, gene trapping and RNA interference.

- Knockout technology is based on the homologous recombination of the targeting vector with genomic sequence in mouse ES cell (Smithies et al., 1985). Advantages of this method include flexibility in design of alleles, possibility to produce reporter knock-ins and conditional alleles and ability to target splice variants and alternative promoters. Some important limitations exist, like the gene may serve a different function in human than in mouse, about 15% of knockouts are developmentally lethal and the procedure is laborious and costly. The total number of knockout mice described in the literature is around 2500 unique genes (Austin et al., 2004)

Recent advances in RNA interference (RNAi) technology have provided a rapid loss-of-function method for assessing gene function in a number of organisms (Prawitt et al., 2004). Plasmid-based systems using H1 or U6 RNA polymerase III promoters driving expression of short hairpin RNA (shRNA) molecules were stably integrated into the genome of mice by microinjecting DNA into one cell mouse embryos or transecting ES cells (Hasuwa et al., 2002, Kunath et al., 2003). The desired knock-down phenotype was apparent and transmitted to the germline. Drawbacks of using this technology are cross-hybridisation of siRNA with similar sequences which lead to multiply gene phenotype, constitutive RNAi expression in knock-down mice which might provoke an embryonic lethal phenotype.

1.4 Gene trap approach

Gene trapping is a high-throughput approach that can be used to introduce insertional mutations across the genome in mouse embryonic stem (ES) cells (Hansen et al., 2003, Skarnes et al., 2004). Gene trap mutagenesis of mouse ES cells generates random loss-of-function mutations, which can be identified by a sequence tag and can often report the endogenous expression of the mutated gene (Wurst et al., 1995). 5' RACE is used to analyse genes trapped by all types of vectors. The gene trap vector pT1 β geo used in this study contains the *engrailed-2* splice acceptor site fused to the promoterless LacZ reporter gene (Fig. 1.6). Integration of pT1 β geo into the intron of an active gene can generate a fusion transcript between LacZ and an endogenous trapped gene. The bacterial resistance is supported by neomycin cassette driven by the PGK-1 promoter and containing an SV-40 polyadenylation signal.



Figure 1.6: Schematic representation of the pT1 β geo gene trap vector used in this study. En2in- *engrailed-2* intron; En2ex- *engrailed-2* exon; β geo, a fusion gene between the β -galactosidase (β gal) and neomycin phosphotransferase (neo) genes; pA, SV40 polyadenylation signal.

Gene trap ES cell lines are available at the different sources like German Gene Trap Consortium (<http://www.genetrapped.de>), the Mammalian Functional Genomics Centre (<http://www.escells.ca>), the Sanger Institute Gene Trap Resource (<http://www.sanger.ac.uk/genetrapped>) and others. Mutated ES cell clones could serve as a resource to accelerate genome annotation and the in vivo modeling of human disease.

1.5 Mouse models for Sall gene function

Recent studies have provided new insights into a putative role of SALL proteins (Nishinakamura et al., 2001, Sato et al., 2004). Generation of mice lacking Sall genes is necessary to address developmental role of Sall family. Several mouse models were created to study the function of *sal*-like genes in mouse and human, including *Sall1* (Nishinakamura et al., 2001, Kiefer et al., 2003), *Sall2* (Sato et al., 2003), *Sall3* (Parrish et al., 2004). *Sall1* is expressed in the metanephric mesenchyme surrounding uterine bud, and homozygous deletion of *Sall1* results in an incomplete uterine bud outgrowth and failure of tubule formation in the mesenchyme (Nishinakamura et al., 2001). Loss of the *Sall1* gene cause perinatal death in mice because of kidney agenesis or severe dysgenesis (Nishinakamura et al., 2001). No other abnormalities in *Sall1*-deficient mice were found. Another mouse model which mimics the phenotype of TBS patients was described by Kiefer et al. (2003). Mice, expressing *Sall1* transcriptional repression domain and lacking the DNA binding domain recapitulate the abnormalities found in human TBS. *Sall1* heterozygous mice display hearing loss, renal cystic hypoplasia and wrist abnormalities. Homozygous mutant mice have renal agenesis, exencephaly, limb and anal deformities (Kiefer et al., 2003). Murine *Sall2* has an overlapping expression pattern with *Sall1*, but *Sall2* knockout mice show no apparent abnormal phenotype (Sato et al., 2003). *Sall3* homozygous mutant mice die on the first postnatal day and fail to feed because of palate and epiglottis anomalies and glossopharyngeal nerve (IX) deficiency (Parrish et al., 2004).

1.6 Aims of the study

Preceding the presented study a *Sall4* mouse gene trap line was identified by comparing *Sall4* sequences to the database of the German Gene Trap Consortium. The present study aimed at characterizing this mutant in detail. Specifics were:

- 1) to identify and characterize potential *Sall4* isoforms
- 2) to characterize the mutation and the phenotype of the *Sall4* gene trap line
- 3) to perform detailed expression studies of *Sall4* during embryonic development
- 4) to analyze *Sall4* regulation during embryonic development
- 5) to generate a specific antibody against Sall4
- 6) to perform protein interaction studies of Sall4 and Sall1

2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

1 kb DNA Ladder	Invitrogen, Karlsruhe
Acetic acid	Merck, Darmstadt
Acrylamide/Bisacrylamide	Roth, Karlsruhe
Agar	Roth, Karlsruhe
Agarose	Invitrogen, Karlsruhe
Alkaline phosphatase	Invitrogen, Karlsruhe
Ammonium acetate	Fluka, Neu Ulm
Ampicillin	Sigma, Deisenhofen
Ampuwa	Fresenius, Bad Homburg
Aprotinin	Sigma, Deisenhofen
BCIP	Applichem, Darmstadt
Boric acid	ICN Biomedicals, Eschwege
Bromophenolblue	Sigma, Deisenhofen
BSA	Biomol, Hamburg
Cell culture media (DMEM)	PAN-Systems, Nürnberg
Chloroform	Merck, Darmstadt
Coomasie G-250	Sigma, Deisenhofen
DEPC	Sigma, Deisenhofen
Dithiothreitol	Sigma, Deisenhofen
DMSO	Merck, Darmstadt
DNA Marker	Invitrogen, Karlsruhe
dNTPs (100 mM)	Invitrogen, Karlsruhe
EDTA	ICN Biomedicals, Eschwege
EGTA	Applichem, Darmstadt
Ethanol	Merck, Darmstadt

Materials and Methods

Ethidium bromide	Eurobio, Les Ulis Cedex
FBS	PAN-Systems, Nürnberg
Ficoll 400	Applichem, Darmstadt
Formaldehyde	Merck, Darmstadt
Formamide	Sigma, Deisenhofen
Glutaraldehyde	Sigma, Deisenhofen
Glycerol	Merck, Darmstadt
Glycine	Merck, Darmstadt
Goat serum	PAN-Systems, Nürnberg
H ₂ O ₂	Merck, Darmstadt
HCl	Merck, Darmstadt
HEPES	Sigma, Deisenhofen
IPTG	Applichem, Darmstadt
Isopropanol	Merck, Darmstadt
K ₃ [Fe(CN) ₆]	Sigma, Deisenhofen
K ₄ [Fe(CN) ₆]	Sigma, Deisenhofen
Kanamycin	Sigma, Deisenhofen
KCl	Merck, Darmstadt
KH ₂ PO ₄	Merck, Darmstadt
Leupetin	Sigma, Deisenhofen
Lipofectamine 2000	Invitrogen, Karlsruhe
M16 medium	Sigma, Deisenhofen
Methanol	Merck, Darmstadt
2-Mercaptoethanol	Serva, Heidelberg
MgCl ₂	Merck, Darmstadt
Milk powder	Roth, Karlsruhe
Mineral oil	Sigma, Deisenhofen
MOPS	Applichem, Darmstadt
Na ₂ HPO ₄	Merck, Darmstadt
NaCl	Roth, Karlsruhe
NaH ₂ PO ₄	Merck, Darmstadt
NaN ₃	Merck, Darmstadt

Materials and Methods

NaOH	Merck, Darmstadt
NBT	Applichem, Darmstadt
Nonidet P40	Fluka, Neu Ulm
NuPAGE MES SDS running buffer	Invitrogen, Karlsruhe
NuPAGE SDS sample buffer	Invitrogen, Karlsruhe
Orange G	Sigma, Deisenhofen
Paraformaldehyde	Merck, Darmstadt
PBS	PAN-Systems, Nürnberg
Penicillin/Streptomycin	PAN-Systems, Nürnberg
Peptone	Roth, Karlsruhe
PfuI polymerase	Promega, Mannheim
Phenol	Biomol, Hamburg
Picric acid	Fluka, Neu Ulm
Platinum Taq polymerase	Invitrogen, Karlsruhe
Proteinase K	Roth, Karlsruhe
RNase A	Qiagen, Hilden
RNase away	Applichem, Darmstadt
RNase Inhibitor	Boehringer, Mannheim
S.O.C Medium	Invitrogen, Karlsruhe
Salmon sperm DNA	Sigma, Deisenhofen
SDS	Serva, Heidelberg
SeeBlue Plus2 protein marker	Invitrogen, Karlsruhe
Sodium acetate	Merck, Darmstadt
Sodium citrate	Merck, Darmstadt
Sodium deoxycholate	Sigma, Deisenhofen
Sucrose	Fluka, Neu Ulm
SuperScript II	Invitrogen, Karlsruhe
T4 DNA ligase	Promega, Mannheim
TEMED	Merck, Darmstadt
TRIsreagent	Sigma, Deisenhofen
Tris	Roth, Karlsruhe
Triton X-100	Serva, Heidelberg

Materials and Methods

Trypsin	PAN-Systems, Nürnberg
Tween-20	Promega, Mannheim
Vectashield (DAPI)	Vector, Burlingame
X-Gal	Applichem, Darmstadt
Xylene	Merck, Darmstadt
Yeast extract	Roth, Karlsruhe

All chemicals not mentioned above were ordered from Merck, Darmstadt, or Roth, Karlsruhe.

2.1.2 Solutions, buffers and media

All standard buffers and solutions were prepared according to Sambrook *et al.* (1989).

5x TBE buffer	450 mM Tris 450 mM Boric acid 20 mM EDTA (pH 8.0)
Stop-mix	15% Ficol 400 200 mM EDTA (pH 8.0) 0.025% Orange G 1% Glycerol
APS solution	10% Ammoniumpersulfate in H ₂ O
Stacking gel buffer (4x)	0.5 M Tris/HCl (pH 6.8) 0.4% SDS
Separating gel buffer (4x)	1.5 M Tris/HCl (pH 8.3) 0.4% SDS
Coomassie Blue (staining solution)	0.25% Coomassie Brilliant Blue R-250 45% Methanol 10% Acetic acid

Materials and Methods

Coomassie Blue (destaining solution)	45% Methanol 10% Acetic acid
NBT solution	75 mg/ml NBT in 70% DMF
BCIP solution	5% w/v in 100% DMF
Denhardt's solution (50x)	1% BSA 1% Polyvinylpyrrolidone 1% Ficoll 400
dNTP-Mix (2 mM)	2 mM dATP 2 mM dGTP 2 mM dCTP 2 mM dTTP
Sodium phosphate buffer (NaP)	0.5 M NaH ₂ PO ₄ (pH 7.4)
GSDB (Goat serum dilution buffer) (blocking solution)	30% goat serum 20 mM NaP buffer 450 mM NaCl 0.3% Triton X-100
GSDB (antibody solution)	3% goat serum 20 mM NaP buffer 450 mM NaCl 0.1% Triton X-100
Lysis buffer (pH 8.0) (bacterial cells)	100 mM Tris 500 mM NaCl

Materials and Methods

	0.5 mM EDTA
	0.1% Triton X-100
	0.1% Tween 20
	250 mM Urea
	5 mM β -Mercaptoethanol
	0.03% DTT
	5% Glycerol
	1 mM PMSF
Tissue lysis buffer	10 mM Tris/HCl (pH 8.0)
	1 mM EDTA
	2.5% SDS
	1 mM PMSF
PBS buffer	140 mM NaCl
	8 mM Na ₂ HPO ₄
	2 mM KH ₂ HPO ₄
	2.7 mM KCl
PBT buffer	0.1% Tween 20
	in PBS
20x SSC	3 M NaCl
	0.3 M Sodium citrate (pH 7.0)

2.1.3 Laboratory materials

The laboratory materials, which are not listed here, were bought from Schütt and Krannich, Göttingen.

Cell culture flask	Greiner, Nürtingen
Culture slides	BD Falcon, Heidelberg

Disposable filter Minisart	NMI Sartorius, Göttingen
Filter paper 0858	Schleicher and Schüll, Dassel
Hybond – C	Amersham, Braunschweig
HPTLC Aluminum folio	Merck, Darmstadt
Microcentrifuge tubes	Eppendorf, Hamburg
Petri dishes	Greiner, Nürtingen
Pipette tips	Eppendorf, Hamburg
RotiPlast paraffin	Roth, Karlsruhe
Superfrost slides	Menze, Gläser
Whatman blotting paper	Schleicher and Schüll, Dassel
X-ray films	Amersham, Braunschweig

2.1.4 Sterilisation of solutions and equipments

All solutions that are not heat sensitive were sterilised at 121°C, 10^5 Pa for 60 min in an autoclave (Webeco, Bad Schwartau). Heat sensitive solutions were filtered through a disposable sterile filter (0.2 to 0.45 μm pore size). Plastic wares were autoclaved as above. Glassware were sterilised overnight in an oven at 220°C.

2.1.5 Media, antibiotics and agar-plates

2.1.5.1 Media for bacteria

LB Medium (pH 7.5):

- 1% Peptone
- 0.5% Yeast extract
- 1% NaCl

LB-Agar:

- 1% Peptone
- 0.5% Yeast extract
- 1% NaCl
- 1.5% Agar

The LB medium was prepared with distilled water, autoclaved and stored at 4°C.

2.1.5.2 Media for cell culture

ES-cell medium:

DULBECCO's Modified Eagles Media (DMEM) with	0.1 mM Non essential amino acids
	1 mM Sodium pyruvate
	10 μ M β -Mercaptoethanol
	2 mM L-Glutamine
	20% Fetal calf serum (FCS)
	1000 U/ml Recombinant leukaemia inhibitory factor

Cell medium:

DULBECCO's Modified Eagles Media (DMEM) with	2 mM L-Glutamine
	1% penicillin/streptomycin
	10% Fetal calf serum

Freezing medium:

20% Fetal calf serum
10% DMSO
in DMEM

2.1.5.3 Antibiotics

The stock solutions were filtered through sterile disposable filters and stored at -20°C . When antibiotics were needed, in each case it was added after the autoclaved medium has cooled down to a temperature lower than 55°C .

	Stock solution	Working solution
Ampicillin	50 mg/ml	50 μ g/ml
Kanamycin	25 mg/ml	50 μ g/ml

Materials and Methods

actinF 5' CTTTGCAGCTCCTTCGTTGC 3'
actinR 5' ACGATGGAGGGGAATACAGC 3'
dom1F1 5' ACGGAATTCTGTTCTCAGGATGTTCCCAGTG 3'
dom1R1 5' GGACTCGAGGGAAGTGAGGGAAGTGGTGT 3'
S4E1aF3 5' AGTTTTCTGCGGATCACTCG 3'
S4E3F 5' GGAGAGAAGCCTTTCGTGTGT 3'
S4E4R 5' GTTCTCTATGGCCAGCTTCCT 3'
pTKneoF264 5' GGCTGACCGCTTCCTCGTG 3'
Ex2F164 5' TGATTACACGACATTTCCACTG 3'
SAR364 5' CTCTTCTTCCTGGCTAACTCG 3'
S4E2RTF1 5' TCTCCCCAACTTCGTGTCTC 3'
S4E3RTR1 5' GGCTGTGCTCGGATAAATGT 3'
Sp6 5' TTTAGGTGACACTATAGAA 3'
T7 5' AATACGACTCACTATAGGG 3'
S1F1new 5' AGTAAGCTTCGGAAGAGGGAGTACAGGGC 3'
S1pcDNAtrR 5' TATTCTAGAAGTGACATTTGGTGGGTTGC 3'
brachyuryF 5' CTGCGCTTCAAGGAGCTAAC 3'
brachyuryR 5' CCCC GTT CACATATTTCCAG 3'
Tbx1F 5' CGACAAGCTGAAACTGACCA 3'
Tbx1R 5' GTGACTGCAGTGAAGCGTGT 3'
Tbx5F 5' ATATTGTTCCCGCAGACGAC 3'
Tbx5R 5' TTGAGCTTCTGGAAGGAGACA 3'
Sall1seq 5' TGGGCATCCTTGCTCTTAGT 3'
S4ME2R3 5' CAGTAATCAATTCCCCATCAGC 3'
S4ME2R2 5' GAAAGGAGAGAAAAAGGTGAGC 3'
S4ME2R1 5' AGGGGAGTTCCTGAGCAC 3'
S4ME1F1 5' GGCTAAAATTTCCCAACTCCAG 3'
S4ME1F2 5' CAGGAATTGGTGGCGGAGAG 3'
S4ME1F3 5' GCACCATGTCGAGGCGCAAG 3'
pGeoR1 5' AAGGCGATTAAGTTGGGTAACG 3'
pGeoR2 5' ATGGAGGAGAAAGGGCAGAGG 3'
pGeoF1 5' ATCAGCCGCTACATGCAACAG 3'

Materials and Methods

pGeoF2 5' ATTGGTGGCGACGACTCCTG 3'
S4MI1R1 5' GTGAAGGAAGCCCAAGTTACC 3'
S4MI1R2 5' CCGACCTCTCCCCTTTCACC 3'
S4MI1F4 5' TGCCTCTCCCTCCAACCTTCTC 3'
S4MI1F5 5' ATGTGAGTTTTGGGGCGTTGG 3'
S4MI1F6 5' AAAAAATCCGCAGGTGGTCAAG 3'
S4MI1F7 5' TGGGTTTCTTTATTGGGGTTAC 3'
S4MI1F8 5' ATAAGTAGATGCTGAAGAGATACC 3'
S4MI1F9 5' ACTAGGTAGGCATCAGGTAAAAC 3'
S4MI1F10 5' GTTTCTCTTAGTATAGCCCCAG 3'
S4MI1F11 5' ATCCCCTAGACTTACTGTTGAG 3'
S4MI1F12 5' ATTCTGCACTTGGGAGGATGAG 3'
S4MI1F13 5' TTTCCATTTTCGCCTGTGTCGTG 3'
S4MI1F14 5' AATTCTTCTTCTTGGAGTGATGC 3'
pGeoR3 5' GTCACTCCAACGCAGCACCATCACC 3'
pGeoR4 5' TGCCAGTTTGAGGGGACGACGACAG 3'
pGeoF3 5' GAAAACTATCCCGACCGCCTTACTGC 3'
pGeoF4 5' CATATGGGGATTGGTGGCGACGACTC 3'
S4ME1F4 5' TAGTGGCCGCAAGGATGATTC 3'
S4MI1F1 5' TGGGGATTCCGGACTCTGTTC 3'
S4MI1F2 5' CAGGAAAGGAGAGATGCGGAG 3'
S4I1R1a 5' TTTAAAAGCGGCGCCACTAGA 3'
S4E1aRTF1 5' AAGCGAGTTTTCTGCGGATCAC 3'
S4E1aRTF2 5' CGGACTGCGACACACACCTG 3'
S4MRTR3 5' AGGGCACTGGGAAGGCGAGAC 3'
pGeoR7 5' CCACAACCAACGCACCCAAGC 3'
pGeoR6 5' GAGATGGATTGGCAGATGTAGC 3'
pGeoR5 5' TGCATTTGTTTTTCTCCCCTTATC 3'
S4E1aF3 5' AGTTTTCTGCGGATCACTCG 3'
Sall4ORF1f 5' GGGGCTAAAATTTCCCAACT 3'
Sall4ORF1r 5' GGGAAGCTCTGACAGCTGAGG 3'

Sall4ORF2f 5' GGCTACCTCAAAAGCTCACG 3'

Sall4ORF2r 5' ATCTCCCTGCAGAAGACTCC 3'

2.1.9 Eukaryotic cell lines

NIH3T3 Mouse embryonic fibroblast cell line, ATCC, Rockville, USA

F9 Mouse teratocarcinoma cell line, ATCC, Rockville, USA.

2.1.10 Mouse strains

Mouse strains C57BL/6J, 129X1/SvJ, CD-1 and NMRI were initially ordered from Charles River Laboratories, Wilmington, USA, and further bred in the animal facility of the Institute of Human Genetics, Göttingen.

2.1.11 Antibodies

Goat anti-rabbit IgG alkaline phosphatase conjugate	Sigma, Deisenhofen
Goat anti-mouse IgG alkaline phosphatase conjugate	Sigma, Deisenhofen
Goat anti-rabbit IgG Cy3 conjugate	Sigma, Deisenhofen
Rabbit anti-mouse IgG FITC conjugate	Sigma, Deisenhofen
Goat anti-rabbit IgG FITC conjugate	Sigma, Deisenhofen
Mouse monoclonal anti-E2 (977)	Abcam, Cambridge
Rabbit polyclonal anti-c-Myc	Sigma, Deisenhofen
Mouse monoclonal anti-c-Myc	Roche, Mannheim
Mouse monoclonal Penta-His	Qiagen, Hilden
Rabbit polyclonal anti-Sall4	generated in the present study
Rabbit polyclonal anti-GST	Novagen, San Diego
Mouse monoclonal anti- α -tubulin	Sigma, Deisenhofen
Sheep anti-Digoxigenin-AP, Fab fragments	Roche, Mannheim

2.1.12 Kits

Lipofectamin2000	Invitrogen, Karlsruhe
DYEnamic ET-Terminator mix	Amersham, Freiburg
Midi Plasmid Kit	Qiagen, Hilden
QIAquick Gel Extraction Kit	Qiagen, Hilden
SulfoLink Kit	Pierce, Bonn
Wizard® SV Gel and PCR Clean-Up System	Promega, Mannheim
Montage PCR	Millipore, Schwalbach
QuantiTest SYBR Green PCR Master Mix	Qiagen, Hilden
pGEM T-easy vector system	Promega, Mannheim
SuperScript II	Invitrogen, Karlsruhe

2.1.13 Instruments

ABI Prism Real-Time detection system 7900	Applied Biosystems, Foster City
Autoclave	Webeco, Bad Schwartau
Biofuge 13	Heraeus, Hanau
Centrifuge 5415 D	Eppendorf, Hamburg
Centrifuge 5417 R	Eppendorf, Hamburg
GeneAmp PCR System 9700	Perkin Elmer, Boston
Histocentre 2 embedding machine	Thermo, Dreieich
MegaBace Sequencer	Amersham, Freiburg
Megafuge 1.0 R	Heraeus, Hanau
Microscope BX60	Olympus, Hamburg
Microscope BZX12	Olympus, Hamburg
Microtom Hn 40 Ing.	Leica, Solms
Neubauer cell chamber	Schütt Labortechnik, Göttingen
Power supply	Gibco BRL, Karlsruhe
Refrigerated Superspeed Centrifuge RC-5B	Sorvall, Langenselbold
Semi-Dry-Blot	Biometra, Göttingen
Spectrophotometer Ultraspec 3000	Amersham, Freiburg

SpeedVac concentrator SVC 100H	Schütt Labortechnik, Göttingen
Thermocycler PTC-100	MJ Reseach, Waltham
Thermomixer 5436	Eppendorf, Hamburg
UV Stratalinker™ 1800	Leica, Solms

2.2 Methods

2.2.1 Isolation of nucleic acids

2.2.1.1 Small-scale isolation of plasmid DNA (mini-prep)

(Sambrook et al., 1989)

A single *E.coli* colony was inoculated in 5 ml of LB medium with the appropriate antibiotic and incubated in shaker for 16 hrs at 37°C at speed of 150 rpm. Seven hundred microliters of the bacterial culture were used for preparation of glycerol stocks (0.7 ml of culture and 0.3 ml of glycerol). A 1.5 ml aliquot of the bacterial culture was centrifuged at 5000xg for 5 min. The pellet was resuspended in 150 µl of P1 solution. The bacterial cells were then lysed with 300 µl of P2 solution and incubated for 5 min with 200 µl of P3 solution. The precipitated solution was centrifuged at 12000xg for 15 min at 4°C. The supernatant was transferred into a new tube and centrifugation was repeated. The supernatant was transferred into a new tube and 0.5 ml of isopropanol was added to precipitate the DNA. It was then stored on ice for 15 min, centrifuged at full speed for 20 min, and finally the pellet was washed with 70% ethanol and after air-drying, the pellet was dissolved in 30 µl of sterile water.

P1: 50 mM Tris/HCl (pH 8.0)
 10 mM EDTA
 100 µg/ml RNase A

P2: 200 mM NaOH,
 1% SDS

P3: 3.0 M potassium acetate (pH 5.5)

2.2.1.2 Large-scale preparation of plasmid DNA (midi-prep)

A single clone was inoculated in 2 ml LB medium with appropriate antibiotic as a pre-culture for 8 hrs in 37°C shaker. This pre-culture was further inoculated at dilution of 1:100 in 100 ml LB medium with appropriate antibiotic and was cultured overnight at 37°C. The bacterial culture was then centrifuged at 6000xg for 15 min. The pellet was resuspended in 4 ml of P1 solution and cells were then lysed by adding 4 ml of P2 and further incubating for 5 min at RT. The lysed culture was precipitated by adding 4 ml of P3 solution and incubating on ice for 15 min. The precipitated solution was centrifuged at 12000xg for 30 min at 4°C. During this time the column (Qiagen-tip), which was provided with the midi preparation kit, was equilibrated with 10 ml of QBT solution. After centrifugation the lysate was passed through the equilibrated column thus allowing the DNA to bind with the column resin. The bound DNA in the column was then washed twice with 10 ml of QC solution. Finally, the DNA was eluted with 5 ml of QF solution. To precipitate the DNA, 3.5 ml of isopropanol was added, mixed thoroughly and centrifuged at 14000xg for 30 min at 4°C. The DNA pellet was washed with 70% ethanol, air-dried and dissolved in 100 µl of H₂O.

QBT:	750 mM NaCl 50 mM MOPS (pH 7.0) 15% Ethanol 0.5% Triton X-100
QC:	1 mM NaCl 50 mM MOPS (pH 7.0) 15% Ethanol
QF:	1.25 M NaCl 50 mM Tris/HCl (pH 8.5)

2.2.1.3 Isolation of genomic DNA from tissue samples

Lysis buffer I: 100 mM Tris/HCl (pH 8.0)
 100 mM NaCl
 100 mM EDTA
 0.5% SDS

The method was performed according to Laird et al. (1991). Mouse tail or tissue biopsies were incubated each in 700 μ l of lysis buffer I containing 30 μ l Proteinase K (10 μ g/ μ l) at 55°C for overnight in a thermomixer. To the tissue lysate, equal volume of phenol was added, mixed by inverting several times, and centrifuged at 14000xg for 5 min. After centrifugation the upper aqueous layer was transferred into a new tube. Similar extraction procedure was also performed with phenol-chloroform (1:1 ratio) and followed with chloroform extraction. Finally, the DNA was precipitated with 0.7 volumes of isopropanol, and the DNA pellet was subsequently washed with 500 μ l of 70% ethanol. The DNA was resuspended in 100-200 μ l of sterile H₂O and dissolved by incubation at 60°C for 20 min, for long term storage the DNA samples were stored at 4°C.

2.2.1.4 Isolation of genomic DNA from mouse embryos

Lysis buffer II 20 mM Tris-HCl
 50 mM KCl
 0.45% NP40
 0.45% Tween-20
 200 μ g/ml Proteinase K

Mouse embryos E7.5-E8.5 were dissected in cold PBS and washed briefly two times. The embryos were collected in separate microfuge tubes containing 30 μ l of Lysis buffer II and incubated at 55°C in thermomixer for 3 hrs. After this digestion step, the samples were boiled in water bath for 10 min and centrifuged. The supernatant was transferred into a new tube and stored at -20°C.

2.2.1.5 Isolation of genomic DNA from preimplantation embryos

Blastocysts were isolated from the uterus of superovulated female mice. Individual Blastocysts were collected in a PCR tube (0.2 ml) containing 5 µl Ampuwa H₂O. Genomic DNA was extracted by freeze-thaw method. In brief, the samples were incubated on dry ice for 2 min and thawed by centrifugation at a full speed. This procedure was repeated 3 times. Finally, the tube content was collected by short centrifugation and stored at -20°C.

2.2.1.6 Isolation of genomic DNA from cultured blastocysts

DNA was prepared by incubating the cultured blastocysts with 15 µl of Lysis buffer III for 4 hrs at 55°C followed by incubation at 90°C for 10 min.

Lysis buffer III	50 mM Tris/HCl (pH 8.0)
	0.5 mM EDTA (pH 8.0)
	0.5% Tween 20
	0.2 mg/ml Proteinase K

2.2.1.7 Isolation of total RNA from tissue samples and cultured cells

Total RNA isolation (TRI) reagent is an improved version of the single-step method for total RNA isolation described first by Chomczynski and Sacchi (1987). The composition of the reagent includes phenol and guanidine thiocyanate in a monophasic solution. In order to avoid any RNase activity, the homogeniser used for RNA isolation was previously treated with RNase away and DEPC water and RNase-free Eppendorf cups were used during the preparation. Tissue samples (50-100 mg each) were homogenised in 1 ml of TRI reagent by using a glass-teflon or plastic homogeniser. The homogenate was vortexed and incubated on ice for 15 min to permit the complete dissociation of nucleoprotein complexes. Then, 0.2 ml of chloroform was added to the sample, vortexed and incubated on ice for further 10 min. After centrifugation at 12000xg for 15 min at 4°C, the colourless upper aqueous phase was transferred into a new tube. After adding of 500 µl of isopropanol, solution was mixed and the

RNA was precipitated by centrifugation at 7500xg for 10 min. Finally, the pellet was washed with 75% ethanol and dissolved in 50-100 μ l DEPC-H₂O. The RNA was stored at -80°C.

2.2.1.8 Determination of the nucleic acid concentration

The concentration of nucleic acids was determined spectrophotometrically by measuring absorption of the samples at 260 nm. The quality of nucleic acids i.e. contamination with salt and protein was checked by measuring the absorption at 280 nm. The concentration was calculated according to the formula:

$$C = (E_{260} - E_{280}) / c$$

C = concentration of sample (μ g/ μ l)

E 260 = ratio of extinction at 260 nm

E 280 = ratio of extinction at 280 nm

f = dilution factor

c = concentration (standard) / absorption (standard)

for double stranded DNA: c = 0.05 μ g/ μ l

for RNA: c = 0.04 μ g/ μ l

for single stranded DNA: c = 0.03 μ g/ μ l

2.2.2 Nucleic acid electrophoresis

Gel electrophoresis is the technique by which mixture of charged macromolecules, especially nucleic acids and proteins are separated in an electrical field according to their mobility which is directly proportional to macromolecule's charge to mass ratio. Agarose gels are used to electrophorese nucleic acid molecules from as small as 50 base pairs to more than 50 kilobases, depending on the concentration of the agarose and the precise nature of the applied electrical field (constant or pulse). Usually, 1 g of agarose was added in 100 ml of 0.5x TBE buffer and boiled in the microwave to dissolve the agarose. The agarose solution was cooled to about 60°C before adding 3 μ l ethidium bromide (10 mg/ml) and poured into a horizontal gel chamber. The 0.5x TBE buffer was used also as electrophoresis running buffer in the gel chamber. The DNA samples were mixed with 5x loading buffer and then loaded into the wells of the gel. The electrophoresis was carried out at a steady voltage (50-100 V). The size of the

DNA fragments on agarose gels was determined by extrapolating the size from a DNA size marker which was also loaded along with the samples in a separate lane of the gel. After electrophoresis, the DNA in the gel was photographed using a UV gel documentation system.

2.2.3 Isolation of DNA fragments from agarose gel

The Wizard® SV Gel and PCR Clean-Up System is designed to extract and purify DNA fragments of 100 bp to 10 kb from standard or low-melt agarose gels or to purify PCR products directly from a PCR amplification. PCR products are commonly purified to remove excess nucleotides and primers. The Wizard® SV Gel and PCR Clean-Up System is based on the ability of DNA to bind to silica membranes in the presence of chaotropic salts. After gel electrophoresis one volume of Membrane Binding Solution was added to the excised from agarose DNA fragment and incubated at 50°C for 10 min. In case of PCR products an equal volume of Membrane Binding Solution was added. After the gel slice was dissolved completely, it was transferred onto the minicolumn and incubated for 1 min. Thereafter the minicolumn was centrifuged at 16000xg for 1 min. The flow through was discarded and the column was washed two times with 700 µl and 500 µl of Membrane Wash Solution respectively. After emptying of the collection tube the minicolumn was centrifuged one more time for 1 min to remove any residual ethanol and placed into a new 1.5 ml microcentrifuge tube. To elute DNA, 20 µl of H₂O was applied directly to the centre of the column, incubated for 1 min and centrifuged for 1 min at 16000xg.

Membrane Wash Solution	10 mM Potassium acetate (pH 5.0) 80% Ethanol 16.7 µM EDTA (pH 8.0)
Membrane Binding Solution	4.5 M Guanidine isothiocyanate 0.5 M Potassium acetate (pH 5.0)

2.2.4 Enzymatic modifications of DNA

2.2.4.1 Digestion of DNA using restriction enzymes

Restriction enzyme digestions were performed by incubating double-stranded DNA with an appropriate amount of restriction enzyme in its respective buffer as recommended by the supplier and incubated at the optimal temperature for that specific enzyme. Standard digestions include 2-10 U enzyme per microgram of DNA. Reactions were usually incubated for 1-3 hrs to ensure complete digestion at the optimal temperature for enzyme activity (typically 37°C).

2.2.4.2 Dephosphorylation of 5' ends of DNA

To prevent the recircularization of restriction digested plasmids without insertion of a DNA fragment, an alkaline phosphatase treatment is necessary. Alkaline phosphatase catalyses the hydrolysis of 5'-phosphate residues from DNA. The following components were mixed:

1-5 µg vector DNA
4 µl 10x Dephosphorilation buffer
1 µl alkaline phosphatase (0.01 U)
in a total volume of 40 µl

The reaction was incubated at 37°C for 30 min, and the reaction was stopped by incubation at 85°C for 15 min. Dephosphorylated DNA was purified by agarose gel electrophoresis and gel extraction.

2.2.4.3 Ligation of DNA fragments

The ligation of an insert DNA into a vector (digested with appropriate restriction enzyme) was carried out in the following reaction:

25-50 ng vector DNA (linearized)
50-100 ng insert DNA
1 μ l 10x ligation buffer or 5 μ l 2x ligation buffer
0.3 μ l T4 DNA ligase (5U/ μ l)
in a total volume of 10 μ l

Blunt-end ligations were carried out at 16°C for overnight, whereas overhang-end ligations were carried out at 4°C overnight.

2.2.4.4 TA - cloning

(Clark, 1988; Hu, 1993)

Taq polymerases have a terminal transferase activity that results in the template-independent addition of a single nucleotide to the 3' ends of PCR products. In the presence of all 4 dNTPs, dATP is preferentially added. This terminal transferase activity is the basis of the TA-cloning strategy. For cloning of PCR products pGEM-T Easy vector system that has 5' T overhangs was used. The reaction was carried out by mixing the following ingredients:

50 ng of pGEM-T Easy Vector
PCR product (1:3, vector to insert ratio)
1 μ l T4 DNA Ligase 10x buffer
1 μ l T4 DNA Ligase
in a total volume of 10 μ l

The content was mixed by pipetting and the reaction was incubated overnight at 4°C.

2.2.4.5 Non-radioactive dye terminator cycle sequencing

Non-radioactive sequencing was performed with the Dye Terminator Cycle Sequencing-Kit (Amersham). The principle of this method is based upon chain termination reaction (Sanger et al. 1977). The reaction was carried out by adding the following components:

1 µg of plasmid DNA or 100-200 ng of purified PCR products
10 pmol of primer
3 µl ET-mix
in a total volume of 10 µl

The ET-mix contains dNTPs, dideoxy dye terminators and *Taq* DNA polymerase. Elongation and chain termination take place during the following PCR reaction:

95°C	4 min	} 25 cycles
95°C	30 sec (denaturation)	
55°C	15 sec (annealing)	
60°C	4 min (elongation)	

2.2.5 Transformation of competent bacteria

(Ausubel et al., 1994)

Transformation of the bacteria was done by gently mixing one aliquot of competent bacterial cells (50 µl) with 10 µl of ligation reaction. After incubation for 20 min on ice, bacteria were heat shocked for 1 min at 37°C and quick chilled for 2 min on ice. Thereafter, 600 µl of S.O.C. medium was added to the bacteria and recovered by incubating at 37°C with shaking for 1 h. They were plated on LB-agar plates containing appropriate antibiotic, for blue-white selection additionally 1 mM IPTG and 40 mg/ml X-Gal were added.

2.2.6 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) is one of the most important techniques in the field of molecular biology. It is a very sensitive and powerful technique (Saiki et al., 1988) that is widely used for the exponential amplification of specific DNA sequences in vitro by using sequence specific synthetic oligonucleotides (primers). The general principle of PCR starts from a pair of oligonucleotide primers that are designed so that a forward or sense primer directs the synthesis of DNA towards a reverse or antisense primer, and vice versa. During the PCR, the *Taq* DNA polymerase (a heat stable polymerase) (Chien et al., 1976) catalyses the

synthesis of a new DNA strand that is complementary to a template DNA from the 5' to 3' direction by a primer extension reaction, resulting in the production of the DNA region flanked by the two primers. It allows the rapid and unlimited amplification of specific nucleic acid sequence that may be present at very low concentration in any sample.

2.2.6.1 PCR amplification of DNA fragments

The amplification cycles were performed in an automatic thermocycler. The PCR reaction contains following components:

1 μ l	DNA (20-50 ng)
1 μ l	forward primer (10 pmol)
1 μ l	reverse primer (10 pmol)
1 μ l	2 mM dNTPs
2 μ l	10x PCR buffer
0.6 μ l	50 mM MgCl ₂
0.2 μ l	<i>Taq</i> DNA Polymerase (5U/ μ l)
Up to 20 μ l H ₂ O	

The reaction mixture was placed in a 200 μ l reaction tube and placed in thermocycler. A standard PCR program is shown here:

Initial denaturation	95°C	5 min		
Elongation	95°C	30 sec	(denaturation)	} 35 cycles
	55° - 65°C	30-45 sec	(annealing)	
	72°C	1-2 min	(extension)	
Final extension	72°C	10 min		

2.2.6.2 Genotyping PCR

For genotyping of *Sall1* and *Sall4* mouse lines three different primers were designed in such a way that amplification product of wild type and mutant allele had different length. *Sall1*

Materials and Methods

genotyping was performed with pTKneoF264, Ex2F264, SAR364 primers. *Sall4* genotyping was performed with S4MI1R1a, pGeoR6, S4I1F1. The genotyping was performed in following reaction:

1 μ l	DNA
1 μ l	primer 1 (10 pmol)
1 μ l	primer 2 (10 pmol)
1 μ l	primer 3 (10 pmol)
1 μ l	2 mM dNTPs
2 μ l	10x PCR buffer
0.6 μ l	50 mM MgCl ₂
0.2 μ l	<i>Taq</i> DNA Polymerase (5U/ μ l)
Up to 20 μ l H ₂ O	

PCR program was used for *Sall1* genotyping:

Initial denaturation	95°C	5 min	} 35 cycles
Elongation	95°C	30 sec	
	64°C	30 sec	
	72°C	1 min 30 sec	
Final extension	72°C	10 min	

PCR program was used for *Sall4* genotyping:

Initial denaturation	95°C	5 min	} 35 cycles
Elongation	95°C	30 sec	
	64°C	30 sec	
	72°C	50 sec	
Final extension	72°C	10 min	

2.2.6.3 Reverse transcription PCR (RT-PCR)

RT-PCR is a technique, which generates cDNA fragments from the RNA templates and thereafter amplifies it by PCR. It can be used to determine the tissue specific expression of genes or expression of a gene in different development stages. 1-5 µg of total RNA was mixed with 1 µl of oligo (dT)₁₈ primer (10 pmol/µl) and RNase free H₂O to a total volume of 12 µl. To disrupt the secondary structure of the RNA, which might interfere with the cDNA synthesis, the mixture was heated to 70°C for 10 min, and then quickly chilled in ice. After a brief centrifugation, the following components were added to the reaction mixture:

4 µl	5x First strand buffer
2 µl	0.1 M DTT
1 µl	10 mM dNTPs

The reaction was mixed and incubated at 42°C for 2 min. Then, 1 µl of reverse transcriptase enzyme (SuperScript II) was added and further incubated at 42°C for 50 min for the first strand cDNA synthesis. The reaction was inactivated by heating at 70°C for 15 min. 1 µl of the first strand reaction was used for the PCR reaction (as described above).

2.2.7 Real time RT-PCR using SYBR Green I dye

SYBR Green I dye intercalates into double-stranded DNA and produces a fluorescent signal. The intensity of the signal is proportional to the amount of dsDNA present in the reaction. Therefore, in each step of the PCR reaction, the signal intensity increases as the amount of product increases. Thus it provides a very simple and reliable method to monitor the PCR reactions in a real time course. Another advantage of this technique is that no modification in oligonucleotide primers are required which facilitates primer design/synthesis and more important it lowers the running cost of PCR reaction. However, optimization of the reaction conditions for each primer set is required. The expression of genes of interest was quantitated using ABI Prism Real-Time detection system 7900 (Applied Biosystems) by using 2x QuantiTest SYBR Green PCR Master Mix (Qiagen). Briefly, according to manufactures

Materials and Methods

instructions the Master mix for reactions was prepared by combining the following items on ice in a 0.5 ml microcentrifuge tube:

5 μ l 2x SYBR Green Mix
2.5 μ l Primer mix (forward and reverse, 1 pmol/ μ l)
2.5 μ l cDNA template

cDNA templates were synthesized from 1-2 μ g of total RNA from E8.5 embryo using oligo(dT)₁₈ primer and SuperScript II reverse transcriptase. The cDNA for standard curves was prepared in serial dilutions: 1; 0.4; 0.16; 0.064. All samples were amplified in duplicates. The program for ABI Prism was created according to the user manual and consists of the following steps:

- 1) Initial PCR activation step: 95 °C 15 min

- 2) Amplification

94 °C 30 sec	}	45 cycles
63 °C 30 sec		
72 °C 30 sec		

- 3) Dissociation step

95 °C 15 sec
60 °C 15 sec (ramping rate 2%)
95 °C 15 sec

Data analysis was done by SDS 2.1 software using standard curve method and Microsoft Office Excel. To verify the specificity and efficiency of real-time analysis, the PCR products were analyzed by monitoring their dissociation curves. The expression of the investigated genes was normalized to β -actin mRNA expression.

2.2.8 Protein biochemical methods

2.2.8.1 Isolation of total protein

Proteins were extracted from fresh or frozen mouse tissues by homogenization in tissue lysis buffer containing protease inhibitors.

Protein lysis buffer	50 mM Tris/HCl (pH 7.5)
	150 mM NaCl
	1 mM DTT
	1% Nonidet P-40
	2.5 mM EDTA
	10 µg/ml Aprotinin
	10 µg/ml Leupeptin

Lysates were sonicated on ice and centrifuged at 16000xg for 20 min at 4°C. Supernatant, containing membrane, organelles and cytosol proteins was collected and stored at -80°C, or used immediately for Western blotting.

2.2.8.2 Electrophoresis of proteins

The NuPAGE® Pre-Cast Gel System (Invitrogen) is a polyacrylamide gel system for high performance gel electrophoresis and is based on SDS-PAGE gel chemistry (Laemmli, 1970). It consists of NuPAGE® Bis-Tris Pre-Cast Gels and specially optimized buffers which have an operating pH of 7.0, giving the system advantages over existing polyacrylamide gel systems with an operating pH of 8.0. The neutral pH increases the stability of the proteins and provides better electrophoretic results. To 10 µl of whole protein lysate 10 µl of 2x Protein sample buffer was added. The samples were denatured by boiling in the water bath for 10 min, cooled at RT for 5 min and loaded in SDS-PAGE (NuPage 4-12% Bis-Tris gel). The gel electrophoresis was run in 1x SDS MES buffer (Invitrogen). To determine the molecular weight of the proteins on the gel, 10µl of a pre-stained molecular weight standard (See Blue Plus2, Invitrogen) was also loaded. The gel was run at 100 V for 2 hrs at RT.

2.2.8.3 Transfer of protein onto nitrocellulose membranes

(Gershoni and Palade, 1982)

After the electrophoresis of proteins in SDS-PAGE, proteins were transferred onto a nitrocellulose membrane by semi-dry system using an electro-blotter (Biometra). Eight pieces of GB004 Whatman filter paper were cut similar to the size of the gel. First, four papers soaked with transfer solution were placed on semi dry transfer machine's lower plate and then equilibrated nitrocellulose membrane was placed over them. Next the gel was placed avoiding any air bubbles. Another four Whatman paper soaked with transfer solution were placed over to complete the sandwich model. The upper plate was placed over this sandwich and the protein transfer was carried out at 10 W (150 – 250 mA, 39 V) for 1 h.

10x Transfer buffer	250 mM Tris
	1.5 M Glycine

Transfer solution	1x Transfer buffer
	10% Methanol

2.2.8.4 Staining of polyacrylamide gels

To assess transfer efficiency of proteins onto nitrocellulose membranes, the gel was incubated overnight in Coomassie blue staining solution. The methanol and acetic acid in the staining solution cause the proteins to precipitate and thus be fixed in the gels. In acidic solution Coomassie Blue dye interacts chiefly with arginine residues of the proteins, resulting in colour development in the gel. Gel was destained in Coomassie destaining solution for 2-3 hrs at RT.

2.2.8.5 Western blotting

The membrane was first incubated in P1 buffer with 5% milk powder for 1 h at RT in order to block non-specific binding sites. Membrane was then incubated with a primary antibody at the recommended antibody dilution in P2 (P1 buffer with 2% milk powder) overnight at 4°C. The unbound antibody was removed by washing three times in P2 for 20 min. Thereafter the

immunoblot was incubated with the alkaline phosphatase conjugated secondary antibody at 1:10000 dilution in P2 for 1 h at RT. Again non-specific bound antibody was washed 3 times in P2 and then immunoblot was equilibrated with AP buffer for 5 min at RT. Finally, alkaline phosphatase activity was detected by incubation in 10 ml of staining solution and the coloring reaction was stopped by washing in H₂O.

P1 buffer: 150 mM NaCl
 100 mM Tris/HCl (pH 7.5)
 0.1% Tween-20

AP buffer 100 mM Tris/HCl (pH 8.5)
 100 mM NaCl
 5 mM MgCl₂

Staining Solution: 66 µl NBT
 33 µl BCIP
 in 5 ml of AP buffer

2.2.8.6 Generation of polyclonal antibody

2.2.8.6.1 Sequence analysis of Sall4 protein

Different computational tools were applied to select potential antigenic domains. The primary amino acid sequence of the Sall4 protein (GenBank NP_780512) was obtained from the NCBI database (<http://www.ncbi.nlm.nih.gov>). In order to select unique sequences for antibody generation, the Sall4 peptide sequence was compared with peptide sequences of the other members of the SALL gene family using the Multalin programme (<http://prodes.toulouse.inra.fr/multalin/multalin.html>). Before synthesis of the fusion protein, a hydrophobicity profile analysis using the Kyte and Doolittle scale and polarity analysis by Gratham scale was carried out (<http://us.expasy.org/tools/protscale.html>). A unique sequence from the C-terminus of Sall4 was selected to synthesize a fusion protein.



Figure 2.1. Schematic representation of Sall4 protein and C-terminus part of the protein with was used for the fusion protein generation (red).

2.2.8.6.2 Amplification and cloning of *Sall4* fusion construct

Sall4 cDNA C-terminal domain (1080 bp) was amplified using primers with 5' overhangs restriction sites sequences (dom1F1 and dom1R1). PCR was performed using *PfuI* DNA polymerase (Promega) with proof-reading activity to avoid mismatches in amplification. Touch-down PCR conditions were as follows:

95 °C	4 min	} 7 cycles
95°C	30 sec	
62°C-56°C	45 sec	
72°C	2 min	
95°C	30 sec	} 30 cycles
56°C	45 sec	
72°C	2 min	
72°C	10 min	

PCR products were digested with *EcoRI* and *XhoI* enzymes, purified from a 1% agarose gel and ligated into pET41a expression vector and constructs were transformed into competent *E.coli* DH5 α cells.

2.2.8.6.3 Expression of recombinant proteins in the pET vector

The recombinant pET41a construct was transformed into the host bacterial strain *E.coli* BL-21 (DE3). The BL-21 strain (Novagen) is lysogenic for a prophage that contains an IPTG

inducible T7 RNA polymerase. A single colony was picked from a freshly streaked plate into 50 ml of LB culture medium containing kanamycin for overnight culture at 37°C. Next day overnight culture was diluted into 500 ml fresh LB medium and cultured further at 37°C until OD₆₀₀ reached 0.6-0.8. A non-induced sample was collected as a control. Induction was performed by adding IPTG to a final concentration of 1 mM and cultured for 4 hrs. Finally, the bacterial cells were harvested by centrifugation at 5000xg for 10 min at 4°C and frozen at -80°C.

2.2.8.6.4 Purification of GST fusion proteins

The pellet of bacterial cells was resuspended in 10 ml of Lysis buffer (pH 8.0) and sonicated on ice. The cell debris from protein extract were removed by centrifugation for 20 min at 10000xg at 4°C. The supernatant was filtrated through 0.45 µm filter and incubated with pre-equilibrated Glutathione-agarose (Sigma) for 1 h at 4°C with gentle mixing. The bound resin was washed four times with PBT at 4°C. Thereafter fusion protein was eluted with 3 ml Elution buffer (10mM reduced glutathione in 50 mM Tris/HCl, pH 7.5) and eluate was collected into three fractions. Purified protein was dialysed overnight at 4°C against PBS. Protein quality and quantity were checked by SDS-PAGE and Western blotting.

2.2.8.6.5 Immunisation of rabbits

Two rabbits were immunised each with 500 µg of Sall4 fusion protein mixed with Freund's complete adjuvant in 1:1 ratio. Before injection, preimmune sera were collected from the animals. First booster immunisation was performed after 14 days after first injection with 1:1 ratio of antigen with Freund's incomplete adjuvant. Second booster was given after 28 days and a third booster after 56 days from the first immunisation. Final bleeding was done after 86 days from first immunisation. The antiserum was aliquoted and stored at -80°C.

2.2.8.6.6 Affinity purification of polyclonal antibody

For affinity purification, HiTrap NHS-activated 1 ml columns (Amersham) were used, the columns are made of medical grade polypropylene, which is biocompatible and non-

interactive with biomolecules. NHS-Activated Sepharose is designed for the covalent coupling of ligands containing primary amino group. Non-specific adsorption of proteins to HiTrap columns is negligible due to the hydrophilic properties of the base matrix. The activated gel is supplied in 100% isopropanol to preserve the stability of the activated gel prior to coupling.

2.2.8.6.6.1 Ligand coupling

Coupling buffer: 0.2 M NaHCO₃
 0.5 M NaCl (pH 8.3)

Sall4-GST fusion protein (1 mg) was dialysed overnight against Coupling buffer. Isopropanol present in column was washed with 1 M HCl (ice cold). The flow rate during the pumping was adjusted to about 1 ml/min. Immediately after washing, 1 ml of the ligand solution was injected onto the column, then column was sealed and let it stand for 4 hrs at 4°C.

2.2.8.6.6.2 Washing and deactivation

Buffer A: 0.5 M Ethanolamine, 0.5 M NaCl (pH 8.3)
Buffer B: 0.1 M Acetate, 0.5 M NaCl (pH 4.0)

A series of alternate washing (injection 3x2 ml) with buffer A and buffer B was done. Finally, two ml of PBS was injected.

2.2.8.6.6.3 Affinity purification of antiserum

The column was equilibrated with 10 column volumes of start buffer (PBS). The antiserum was diluted two times with PBS and filtered through a 0.45 µm filter and then applied onto the column. During pumping, a constant flow rate of 0.5 ml/min was maintained. The column was washed with 10 column volumes of PBS. Elution was done with three volumes of 100 mM glycine/HCl (pH 2.7). The purified antibody fraction was desalted by dialysis against PBS. The column was re-equilibrated with 10 volumes of PBS. The antiserum solution was recovered from the dialysis tubing and concentrated using Centriscart columns to about 0.5 ml.

2.2.8.6.7 Protein G purification of antibody

Protein G sepharose beads (Santa Cruz) were equilibrated by washing five times in binding buffer (20 mM sodium phosphate buffer, pH 7.0, 0.15 M NaCl). Antiserum diluted in ratio 1:1 with binding buffer was mixed with protein G sepharose and incubated at RT for 1.5 h. After brief centrifugation step, the supernatant was removed and protein G was washed four times with binding buffer. Elution was done with four protein G volumes of Elution buffer (100 mM glycine/HCl, pH 2.7). Eluted antibody fractions were neutralized by adding 3.5 μ l of 1 M Tris/HCl (pH 7.5). Finally, antibody was dialyzed overnight against PBS.

2.2.9 Eukaryotic cell culture methods

2.2.9.1 Cell culture conditions

NIH3T3 cells and F9 cells were grown in DMEM medium containing 10% fetal bovine serum and 1% penicillin/streptomycin solution. The cells were cultured at 37°C in a humidified incubator with 5% CO₂.

2.2.9.2 Trypsinisation of eukaryotic cells

Cells were washed twice with sterile PBS and incubated in minimal amount trypsin-EDTA (0.5 g/l trypsin, 0.2 g/l EDTA) at 37°C until they had detached from the dish. The process was controlled under an inverted microscope. Trypsin was inhibited by addition of growth medium in which the cells were subsequently resuspended. The trypsin was removed by centrifugation at 2000xg for 3 min. Cells were resuspended in an appropriate volume of cell culture medium and transferred into a new flask with medium.

2.2.9.3 Cryopreservation and thawing of eukaryotic cells

Trypsinased cells were spun down (1000xg for 5 min) in 4 ml of growth medium. The supernatant was aspirated and the cells were resuspended in ice-cold freezing medium (DMEM, 20% FCS, 10% DMSO). The aliquots of the cells were kept for 2 days at -80°C and

then stored in liquid nitrogen. For revitalization, frozen cells were quickly thawed and cells were inoculated in a suitable amount of growth medium.

2.2.9.4 Transient transfection of the cells

Approximately 4×10^5 fibroblast cells (NIH3T3) were plated on a cell culture slide (Falcon) and cultured overnight in 1 ml of DMEM medium containing 10% FCS and penicillin/streptomycin at 37°C and 5% CO₂. The construct DNA (2 µg of pcDNA3.1-Sall1tr) was diluted in 100 µl of Opti-MEM medium. Lipofectamine 2000 was diluted in 100 µl of Opti-MEM medium in ratio 1:3 per µg of DNA. After 10 min of incubation, the diluted DNA was mixed with diluted Lipofectamine 2000 and incubated for 30 min at RT to allow DNA complex formation. Meanwhile the cells were washed twice with PBS. After DNA complex formation 800 µl of cell medium was added to the reaction tube, mixed vigorously and applied to the cells in culture slide, The transfection was carried out by incubating the cells at 37°C for 3 hrs in a CO₂ incubator, than 1 ml of Opti-MEM medium containing 20% FBS was added and further incubated for 2.5 hrs. Finally, cells were washed with PBS and supplemented with 2 ml of fresh complete growth medium. Cells were incubated for 24-48 hrs at 37°C and 5% CO₂.

2.2.9.5 Immunocytochemical analysis

Transiently transfected cells attached to the cell culture slide were rinsed twice with PBS and fixed with 4% PFA in PBS for 30 min at RT, then rinsed again with PBS. Fixed cells were permeabilized and blocked with GSDB for 1 hr. For the immunodetection cells were incubated with primary antibody at the recommended dilution in GSDB antibody solution overnight at 4°C. The unbound primary antibody was removed by washing three times in PBT for 10 min. Thereafter the cells were incubated with either FITC or Cy3-coupled secondary antibody at 1:200 dilution in GSDB antibody solution for 1.5 hrs at RT. The cells were washed three times in PBT. After washing, the cells were mounted with Vectashild mounting medium containing DAPI and were visualised with fluorescence microscope (BX-60, Olympus).

2.2.10 Histological techniques

2.2.10.1 Tissue preparation for paraffin-embedding

The freshly prepared tissue was fixed for 24-72 hrs in 4% PFA to cross link free amino groups of proteins and to prevent the alterations in the cellular structure. The dehydration process was accomplished by passing the tissue through a series of increasing alcohol concentrations. For this purpose, the tissue was serially incubated in 70%, 80%, 90%, 96% and 100% ethanol for 30 min at RT each step. The ethanol was replaced from the tissue by incubating in isopropanol for overnight. Tissue was then serially incubated in different compositions of isopropanol/xylene in ratios 3:1, 1:1 and 1:3 for 30 min-1 h at RT each step. Then tissue was incubated in 100% xylol overnight. Then the tissue was incubated in paraplast at 60°C overnight and next day the paraplast was changed at least three times. Finally, the tissue was placed in embedding mould and melted paraffin was poured into the mould to form a block. The block was allowed to cool and then stored at 4°C.

2.2.10.2 Sections of the paraffin block

The paraffin blocks were first cut to the optimal size before clamping them into the microtome (Hn 40 Ing.). The cut-thickness used for paraffin embedded sections was 3-7 µm. The sections were floated on 42°C water bath to allow complete spreading and subsequently collected onto Superfrost slides. A fine brush was used to transfer the sections to slides. Slides were then dried at 40°C, to remove the excess of water. The sectioned slides were then stored at RT for further analysis.

2.2.10.3 Staining of the histological sections (Hematoxylin-Eosin staining)

The sectioned slides were first incubated three times in xylene for 3 min each, followed by incubation in 100%, 96%, 80%, 70% and 50% ethanol for 2 min each. Thereafter slides were washed for 5 min in H₂O and stained for 3 min in hematoxylin. This staining was followed by washing in running tap water for 10 min. The treated slides were destained in acid ethanol (1% acetic acid, 70% ethanol) for 1 min, washed in H₂O for 2 min and stained with eosin (0.1% +

2% acetic acid) for 1 min, then in H₂O for 1 min and incubated in 50%, 70%, 80%, 90%, 96% and 100% ethanol for 2 min in each. Finally the stained slides were incubated two times in xylene for 3 min and mounted with cover slides.

2.2.10.4 Immunohistochemistry

Fixation and subsequent treatment of mouse tissue was performed as described above. Tissue sections (5-7 µm) were deparaffinized 3 times with xylene and rehydrated by 50%, 70%, 80%, 90%, 96% and 100% ethanol for 2 min in each. For immunostaining, sections were washed 3 times in PBS and were then incubated in GSDB blocking solution for 1 h at RT. After blocking, the tissue sections were incubated with primary antibody (1:50-1:200) overnight at 4°C in humidified chamber. Next day, the sections were rinsed three times in PBS and incubated with secondary antibody (e.g. FITC-conjugated or Cy3-conjugated, 1:200) for 1 h at RT. After incubation with secondary antibody, sections were washed again in PBS for three times 10 min each. The slides were air-dried and mounted with DAPI to stain the nuclei. The immunostained sections were examined using a fluorescence-equipped microscope (BX60, Olympus).

2.2.11 Techniques for recovery and culture of preimplantation embryos

2.2.11.1 Superovulation

Adult female mice were superovulated by intraperitoneal injections of 5 IU of pregnant mare's serum gonadotropin (PMSG, Sigma). After 44-48 hrs the females were injected by 5 IU of human chorionic gonadotrophin (HCG; Sigma) and bred with males. Next morning the females were checked for vaginal plug (VP). The E0.5 was considered to be 12:00 noon at the day of vaginal plug.

2.2.11.2 Preparation of blastocyst stage embryos

The VP positive female mice at E3.5 stage were killed by cervical dislocation. The skin and peritoneum were opened with the large transverse incision to expose the abdominal cavity.

The oviducts with the upper part of the uterus were dissected and placed into a drop of M16 medium. Blastocyst stage embryos were flushed from the uterus of superovulated females. Under dissection microscope a needle attached to a 1 ml syringe was inserted in the fimbrial end of the oviduct. The needle was then held with forceps and oviducts were flushed with 0.05 ml of M16 medium. Embryos were collected with a pipette and washed several times with M16 medium. The collected embryos were washed two times in a drop of PBS and each single embryo was transferred into a PCR cup (0.2 ml) containing 5 μ l H₂O. The single isolated blastocysts were used for genotyping.

2.2.11.3 In vitro culture of blastocyst stage embryos

The blastocysts were flushed out from the uterus of plugged females at day 3.5 (as described at 2.2.11.2) and placed in gelatinized 96-well plate. Embryos were cultured at 37°C and 5% CO₂ in ES cell medium. Embryonic stem cell growth was monitored daily under inverted microscope for 5 days.

2.2.12 Whole mount *in situ* hybridisation

(Wilkinson, 1992; modified)

Whole mount *in situ* hybridisation method was employed in order to determine the expression pattern of genes during embryogenesis. The principle behind *in situ* hybridisation is that a specific labelled antisense mRNA probe binds to complementary mRNA sequences in the fixed tissue, which can be detected by enzymatic colouring reaction.

2.2.12.1 Embryo dissection

After breeding, the E7.5 and E8.5 stage pregnant female mice were dissected and the embryos were isolated and placed in the ice-cold PBS. Embryos were then fixed overnight at 4°C in 4% PFA in PBS. The embryos were washed twice in ice-cold PBT for 10 min and then dehydrated by incubating for 5 min in 25%, 50%, 75%, 2x100% methanol in PBT, respectively. Embryos can be stored at this point for several months in -20°C.

Materials and Methods

The probes for Pax3 and brachyury genes (with the permission of Prof. Herrmann, Berlin) were kindly provided by Max-Plank Institute of Biophysical Chemistry Plasmid Library.

	sense polymerase	antisense polymerase	sense restriction	antisense restriction	length of the probe
Pax3		T7	PstI	HindIII	519 bp
Brachyury		T3		NotI	1 kb

An aliquot of the linearised plasmids was tested on agarose gels, and the plasmid DNA was precipitated. The transcription was carried out as described:

2 μ l 10x transcription buffer
1 μ g linearised plasmid
2 μ l 0.2 M DTT
2 μ l dNTP labelling mix
0.5 RNAase inhibitor (100 U/ μ l)
2 μ l RNA polymerase
in 20 μ l H₂O

Incubation was performed overnight at 37°C. Precipitation of the product was done as follows:

20 μ l transcription reaction
80 μ l H₂O
1 μ l glycogen (20 μ g/ μ l)
35 μ l ammonium acetate (10 M)
250 μ l ethanol (cold)

The mixture was incubated for 15 min on dry ice, centrifuged for 10 min and washed with 500 μ l of 70% ethanol. The dried pellet was dissolved in 80 μ l DEPC-H₂O and stored at -20°C. 3-4 μ l of DIG-labeled probe was used per ml of hybridisation mix as described below.

2.2.12.2 Pre-hybridisation treatment and hybridisation of the embryos

Rehydration of embryos was performed by incubating the embryos for 5 min in 100%, 75%, 50% methanol in PBT. They were washed twice in PBT and bleached with 6% hydrogen peroxide in PBT for 1 hr. This treatment decreased non-specific background (Wilkinson, 1992). The embryos were then washed three times with PBT and treated with 10 µg/ml Proteinase K in PBT for 2 min (E7.5 and E8.5). The reaction was immediately stopped by adding 2mg/ml glycine in PBT, and then the embryos were washed three times for 5 min in PBT. The embryos were postfixed in 4% paraformaldehyde/0.2% glutardialdehyde in PBT for 20 min and then washed twice in PBT. The hybridisation steps were carried out in 6 well cell culture dishes, and solutions were exchanged using a fine titration pipette. Prehybridisation was carried out by incubating in prehybridisation mix (PreHyb: 50% deionised formamide, 5x SSC, pH4.5, 1x Denhardt's Solution, 100 µg/ml yeast tRNA, 1 mg/ml salmon sperm DNA) overnight at 70°C. Hybridisation solution (50% deionised formamide, 5x SSC, pH4.5, 100 µg/ml yeast tRNA, 1 mg/ml salmon sperm DNA, 50 µg/ml heparin) and DIG-labeled RNA probe were mixed, denatured at 80°C for 5 min. Hybridisation was carried out overnight at 70°C in a humidified chamber.

2.2.12.3 Detection of hybridisation signals

Posthybridisation washes were performed to remove non-specifically bound probe. Washing was done two times for 30 min with Solution I (50% formamide, 5x SSC, 1% SDS) and two times for 45 min with Solution II (50% formamide, 2x SSC, 0.2% SDS). The embryos were then incubated for 5 min in freshly prepared MAB buffer (100 mM maleic acid, 150 mM NaCl, pH 7.5). Antibody blocking solution (2% blocking reagent in MAB, 10% sheep serum) was added to the embryos and incubated for 2-3 hrs at RT. Anti-DIG alkaline phosphatase conjugate antibodies were used for immunodetection of DIG-labeled probe. The embryos were incubated in blocking solution with 1% sheep serum overnight at 4°C with shaking. Post antibody washes were carried out in MAB buffer every hour during day at RT and overnight at 4°C.

	5 mM EGTA
	2 mM MgCl ₂
	in solution A
Solution D	0.5 mg/ml X-gal
(staining solution)	10 mM K ₃ [Fe(CN) ₆]
	10 mM K ₄ [Fe(CN) ₆]
	in solution C

2.2.13.2 X-gal staining of blastocyst stage embryos

The blastocyst stage embryos were rinsed twice with PBS (pH 7.4) and then fixed for 10 min at 4°C with 0.25% glutaraldehyde in PBS. After two times washing in PBS, the embryos were incubated in staining solution. Incubation was carried out in humidified chamber overnight in the dark at 37°C. Stained embryos were washed two times in PBS then placed over a glass slide under the mineral oil. The embryos on slide were observed under light microscope (BLX12, Olimpus).

Staining solution	0.04% X-gal
	1 mM MgCl ₂
	10 mM K ₃ [Fe(CN) ₆]
	10 mM K ₄ [Fe(CN) ₆]
	in PBS (pH 7.4)

2.2.14 BrdU staining

Bromodeoxyuridine (BrdU), an analog of thymidine, is a uridine derivative that can be incorporated specifically into DNA in place of thymidine. Those cells that are synthesizing DNA (in S-phase of the cell cycle) will incorporate BrdU into the DNA. Cells, which have incorporated BrdU into DNA, can be quickly detected using an antibody against BrdU. The binding of the antibody is achieved by denaturation of the DNA. This is usually achieved by exposing the cells to acid or heat.

Pregnant mouse females were injected with 100 µg/g body weight of mouse BrdU dissolved in 0.9% NaCl. To detect BrdU incorporation, the mouse embryos were dissected after 2 hrs and washed twice in cold PBS. After that, embryos were fixed in 4% PFA overnight, then paraffin embedded, sectioned and mounted onto slides. To decrease antigen masking by chromatin proteins, the sections were digested in 1mg/ml trypsin in PBS for 10 min and then washed 3 times in PBS. Denaturation of the DNA was carried out with 2.5 N HCl at 37°C for 15 min; slides were washed two times in PBS. Blocking was performed with 2% horse serum in PBS. Primary antibody rat anti-BrdU was incubated for 1 h (1:200 dilution) with following washing and incubation with secondary antibody (horse anti-rat IgG HRP-conjugated, 1:500 dilutions). After washing with PBS 100 µL of DAB substrate was added onto each slide and incubated for 15 min in the darkness. Slides were counterstained and mounted with HistoKit.

2.2.15 Computer analysis

For the analysis of the nucleotide sequences, programs like BLAST, BLAST2, MEGABLAST and other programs from National Centre for Biotechnology Information (NCBI) were used (www.ncbi.nlm.nih.gov). For restriction analysis of DNA sequence NEBcutter V2.0 programme was used (<http://tools.neb.com/NEBcutter2/index.php>). Information about mouse alleles, phenotypes and strains was obtained from Jackson Laboratory (www.informatics.jax.org). For proteins studies ExPASy tools (www.expasy.ch) were used.

3. RESULTS

3.1 Gene trap cell line

3.1.1 Identification of *Sall4* gene trap cell line

In order to establish an appropriate mouse model system for Okihiro syndrome and to study the function of *Sall4* gene we have searched for the *Sall4* gene trap ES cell line clone in the database of German Gene Trap Consortium (GGTC) (Kohlhase, 2002c). The GGTC has established a resource of mutant ES cell lines, which have been obtained by random integration of the gene trap vector (Hansen et al., 2003). Partial sequencing results from 5' RACE PCR of trapped exon sequences are available in the database at www.genetrap.de.

We have identified the ES cell line W097E01 in the GGTC database and used it for further analysis, because 5' RACE PCR result showed that the gene trap vector pT1 β geo was spliced to exon 1 of the *Sall4* gene (Fig 3.1). This ES cell line was used for blastocyst injection and generation of chimeric mice.

```

1      AGCCGCTAGTCCTCTTAGTTAGGGGTCTTCTTCCCTTGGTTTTTCGGGACCT
51     GGGACCGTGTTTCCCTCCGCCGACAGCGCACTAAGCTGAGTTAGAAACTCTG
101    GAGCTGCTGAGGCTGCTCTCCCTAGAGCCTGCTCCCAGTGATGGGATGG
151    GGCTTCGCCTGCTTGCGCCTCGACATGGTGCGCGCATTGGGGCGCCGGGA
201    AGAGCCCTGGTGACTTGCCCCCTCTCCGCCACCAATTCTGGAGGTGGGA
251    AATTTTAGCCCCCTTCGGCCGAAACGCGCATGGCCCAGTAATTCTCCCC
301    TCAATAATGCATTGGGATTTATCATGAGCTGTGACAGATGGGTAACGACA
351    CCCCCCCCCCGGGGCGGAAAGATGCCAAGGCAAACAAGGGGAAAAAC
401    GGGAACGGGAAGGAGAAGGGGGGGGGGGGGAGGGGGGGGGGGAGGGGGG
451    GGGGGGAGGGGGCAAGGNCAGCCNCGGGGGGCAGGGCNCAGGGGTTCGGG
501    GAGGGGGGGGGAGGGGANNCGCCAGGGGGGCAAGGGGGGTGTNTTNTACA
551    AGCTCGGGCCCAAGACAGTGGAGGGCGCGGGGGGAGGGAAAGGGGGAAAG
601    GGGGGGAAGAGGGTGGGGCAGAAGGGGGGAGGCCGGGGACTAGAGGGGGC
651    AGAGGGAAGGGACTGGGGACAGGGGGGGGAGAGGGGGGGCAAGGGGAA
701    AAGAAAAGAGGGGGGGGAAGGGGGGGGGGGCGGGGGGGGAACGGGGG
751    GCGGGGGGACGGGGACGGGGGACAGGGGAGGACGACGGGGGACGG
801    GGGGGGGGACGGGGACAGGGAAAGGGGACGGGGACAGGGGGAAACGG
851    GGGGGGGGGGGGGGACAAGGGGGGCGGGGGCAGGAAAGCACGGGGGGA
901    CGGCGGGGAAAAGGGAACGGGGAGGGGGGGGGGAAACGGGCAAGGGGGG
951    GGGGGTGGGGGAAACGGGGGGGAAGGGGCGAACACGGGGAACGGGGGA
1001   CAGCGGGGGGGGGGGGGGACGGGGGGGAAGGGGACGGGGACAGGGGGG
1051   GGACCGGGGAAGGGAAACGGGGGAAGGGGGGACCCGGGGAAGGGGGCA
1101   GGGGACGGAGGGGGAA

```

Figure 3.1: Sequence of the 5'RACE PCR product (3'-5' orientation) from the W097E01 ES cell line. Blue colour: 5' part of gene trap vector (pT1 β geo), green: exon 1 of *Sall4*, red: unspecific sequence generated by 5' RACE.

3.1.2 Localisation of the gene trap vector integration site in *Sall4* gene

The integration site of the gene trap vector was determined by comparison of the *Sall4* cDNA sequence with the sequence of the 5' RACE PCR. The gene trap insertion occurred within intron 1, between exon 1 and exon 1a (Fig. 3.2). To identify the exact insertion site of the gene trap vector in the genomic DNA, a PCR-based approach was used. The mutant locus was amplified using primers derived from exon 1, intron 1, exon 2 and gene trap vector (Fig 3.2B). Primers in the intron 1 were designed every 1000 bp in order to cover the full length of 10 kb genomic sequence (Table 3.1).

5' part of the insertion		3' part of the insertion	
Forward primer	Reverse primer	Forward primer	Reverse primer
S4ME1F1	pGeoR1	pGeoF1	S4ME2R1
S4ME1F2	pGeoR2	pGeoF2	S4ME2R2
S4ME1F3		pGeoF3	S4ME2R3
S4MI1F4		pGeoF4	S4MI1R1
S4MI1F5			S4MI1R2
S4MI1F6			
S4MI1F7			
S4MI1F8			
S4MI1F9			
S4MI1F10			
S4MI1F11			
S4MI1F12			
S4MI1F13			
S4MI1F14			

Table 3.1: Primers which were used in different combination to amplify 5' part of the insertion and 3' part of the insertion.

Amplification with primers S4ME1F2 and pGeoR1 revealed a fragment of 878 bp. After sequencing of this fragment the insertion of gene trap vector was localized 208 bp 3' from exon 1 (Fig 3.2). In the integration site of the gene trap vector it was small deletion of the

Results

Sall4 intron 1 sequence (14 bp). Also it was found that 1340 bp Engrailed 2 intron sequence of the pT1 β geo vector was deleted.

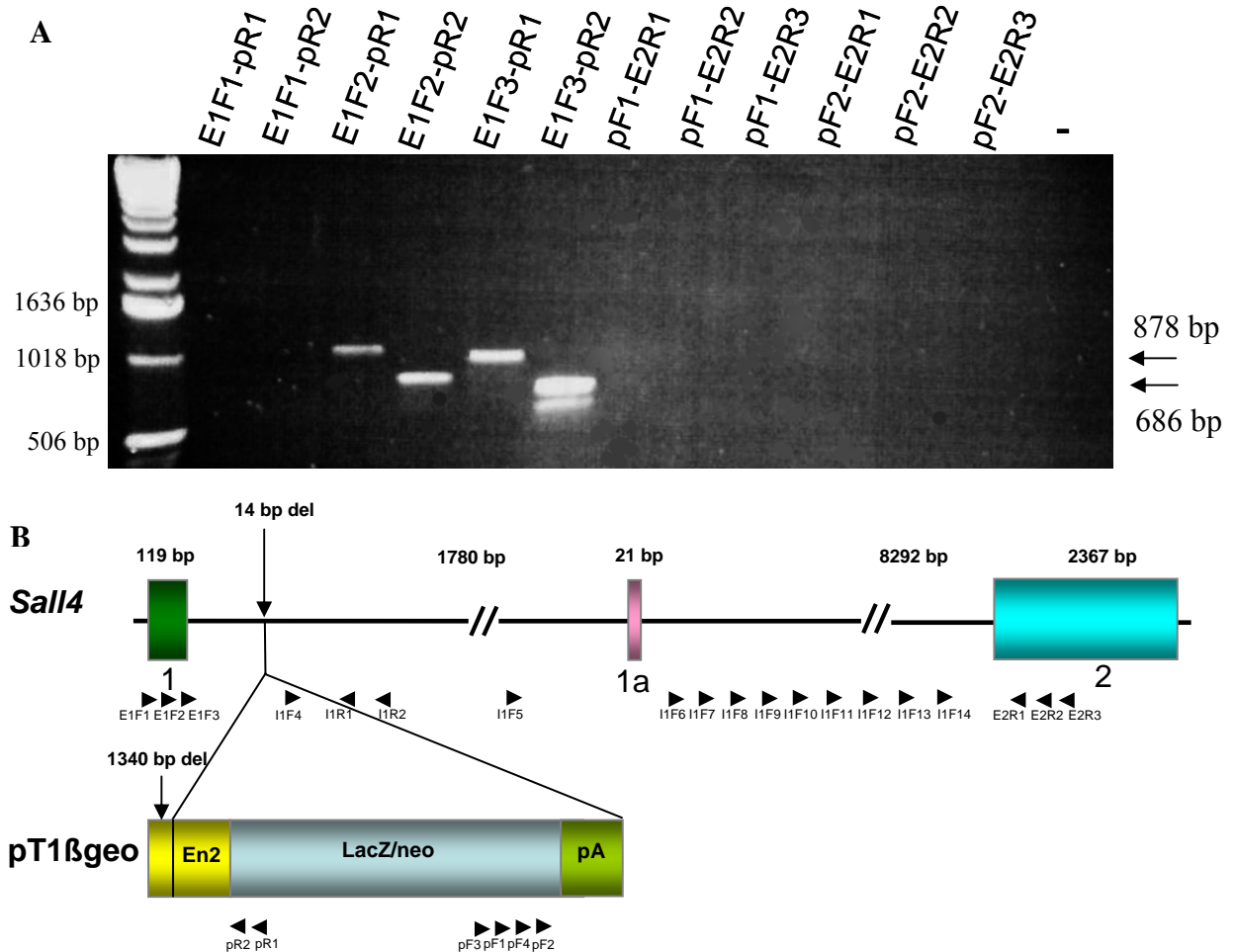


Figure 3.2: (A) Localisation of gene trap vector by PCR amplification of 5' part of the insertion. PCR products were used for sequencing and further analysis. (B) Schematic representation of gene trap vector pT1 β geo integration locus in *Sall4* gene. Only first three exons are shown. Integration occurred in the intron 1 between exon 1 and exon 1a. Transcription of mutated *Sall4* is presumably stopped by a polyadenylation signal at the 3' end of gene trap vector. Deletions of the intronic sequences in the genomic DNA and in the vector are marked with arrows. Positions of primers are marked by arrowheads.

3.2 Characterisation of *Sall4* splicing forms

The *Sall4* gene was first described by Kohlhase et al. (2002). At that time only one transcript variant of *Sall4* was known with 4 exons encoding a protein of 1067 aa in length. Comparison of mouse ESTs with the mouse *Sall4* coding sequence (NM_175303) revealed several highly homologous EST sequences. The 3' part of the BY729264 sequence showed similarity to exon 2 of *Sall4* but the 5' end showed similarity to a putative exon in intron 1. Moreover this sequence contained a predicted translation initiation site and possible splice acceptor sequences. These data have suggested the presence of another 5' exon localized in intron 1, named 1a, which was downstream of exon 1 in *Sall4* (Fig. 3.3). An extensive analysis of the all reported EST, of *Sall4* in the database detected transcripts containing either exon 1 or exon 1a suggesting the presence of two alternative spliced forms of *Sall4*.

```

1  ggggtcggagc  ccggcctctg  gcgtgcgggc  gccctctgga  gaccggccat gtgcggaactg
61  cgacacacac  ctggtgctcc  agtgaactcc  cctgggaact  gcgatgaagc  ctgagaggac
121 tccataccgg  tgaagcggcc  ccggcgggag  gacactcaca  tctgcaacaa  atgctgtgcc
181 gagttcttta  gtctctctga  attcatggaa  cacaagaaaa  gttgcaactaa  aacctctct
241 gtctctcatca  tgaatgacag  cgagggggcca  gtgccttcag  aggacttttc  cagagctgcc
301 ctgagccacc  agctgggcag  cccaagcaat  aaagacagtc  tccaggagaa  cggcagcagc
361 tcggggggact  tgaagaagct  gggcacggac  tccatctgt  acttgaagac  agaggctacc
421 cagccatcca  cccccagga  cataagctat  ttaccxaaag  gcaaagtagc  caacaccaat
481 gtgactctgc  aggcgctccg  gggaccaagg  tggcctgtaa  ccaacggggg  gcagaggcac
541 ccatggcgcc  catgcctgct  gccagggcat  cccttggggc  ctggagcaga  tcttgggctt
601 gcagcaggag  ccacttccgc  aaagccgctt  acggaacaga  tcccgtccag  gtgacatgtg
661 ggca

```

Figure 3.3: EST sequence (BY729264) showing a new alternative splice variant of *Sall4*. Black colour: sequence of the predicted alternative exon 1a, red: exon 2 sequence. Note that the exon 2 sequence is in frame with the exon 1a sequence harbouring an alternative predicted start codon ATG (underlined).

To confirm the presence of predicted two alternative first exons, RT-PCR with total RNA from E8.5 embryos was performed. Primers S4ME1F1 and S4ME2R2 were used to amplify transcript with exon 1 and primers S4ME1aF1 and S4ME2R2 – to amplify transcript with exon 1a. (Fig. 3.4). Two RT-PCR products of different length were purified from the agarose gel and sequencing of the two products confirmed the presence of two *Sall4* splice variants.

Results

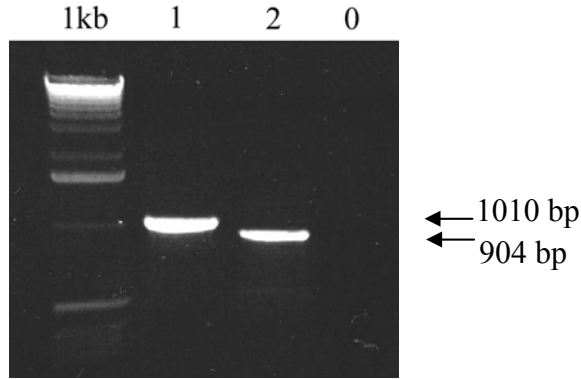


Figure 3.4: RT-PCR analysis with total RNA isolated from E8.5 mouse embryos using *Sall4* specific primers. Lane 1: PCR product from exon 1 – exon 2 transcript (1010 bp), lane 2: PCR product from exon 1a – exon 2 transcript (904 bp), lane 0: negative control for PCR.

The *Sall4* major transcript consists of 4 exons. The *Sall4* minor transcript variant contains alternative exon 1a instead of exon 1 (Fig 3.5), which was spliced in frame to the exon 2 sequence.

```

2   gtttcggccgaagggggctaaaatttcccaactccaggaattggtggcggagagggggca
62  agtcaccagggtctcttcccggcgcccccaatgcgcgccaccatgutcgagggcgcaagcaggcg
                                     M S R R K Q A
122 aagccccagcacatcaactgggaggaggggccagggcgagcagcctcagcagctaccgagc
    K P Q H I N W E E G Q G E Q P Q Q L P S
182 cccgacctcgccgaggcgctggcggcgagggaacccgggtgagtgcggtgctgcgccacgtc
    P D L A E A L A A E E P G
242 cgtgccccgactgtccgtgcggtcccgggagggcgctcctagcaactgccgcgctccctc
362 gagggcgcggttccgggagccgattctggaatggggattccggactctgttcaggaaa
422 ggagagatgctggaggaggaggcctcaatccccacctcggaggagcctcaaccgccc
482 acccct*ttgcaaatccccctagttttctccctcttattgtaaagttgattccacattctg
542 cctctccctccaacttctctctcctcgtccacagtcaccataattgggttccctcctttgg
602 atacataagcaccacctacgcccgttaacttgggcttccctcaccctctctagtg...
1802 ...ttggggtgagcgcatactagtggttttttgtttttgttttttctcctattttagctggt
1862 ggtggggactctgggggtgggcgcatggagaggtggcttgctatagaagggttccgaa
1922 atctccgctggctggagtcgtagttaacctggctgcaccctgaggccgctgggcgcgcg
1982 tgccaccagccttgggaagcgagttttctgcgatcactcgcggtcggagcccggcctc
2042 tggcgtgccccgcccctctggagaccggccatgutgcggactgcgacacacacctggtgag
                                     M C G L R H T P G
2102 tgaggttagctaggccggagaccgcacacgcgctgatctctgtgggggtccccatctt...
10271 ...tggtcacttgacatcccttctcagggcatcctgctcacctttttctcctcttcaca
10328 ggtgctccagtgaaactcccctgggaactgcgatgaagcctcagaggactccataaccggtg
    G A P V N S P G N C D E A S E D S I P V
10388 aagcggccccggcgggaggacactcacatctgcaacaaatgctgtgcccagagttctttagt
    K R P R R E D T H I C N K C C A E F F S
10448 ctctctgaattcatggaacacaagaaaagttgcaactaaaaccctcctgtcctcatc...
    L S E F M E H K K S C T K T P P V L I

```

Figure 3.5: Genomic sequence of *Sall4* gene including exon 1, intron 1, exon 1a and partial sequence of exon 2. The corresponding amino acid sequence of exon 1 (blue), exon 1a (red) and exon 2 (orange) is represented below the sequence. Transcription start sites are underlined. Gene trap vector integration site is marked with asterisk. The deletion of the genomic sequence in the site of integration is marked with green.

Results

Search of the GenBank database at NCBI (<http://www.ncbi.nlm.nih.gov>) revealed the description of two additional transcript variants ([NM_201395](#), [NM_201396](#)). These transcripts are either missing 3' sequence of exon 2 (isoform b) or the complete sequence of exon 2 (isoform c). In the context of our project, we focused our investigation on the *Sall4* major and *Sall4* minor transcript variants only (Fig. 3.6).

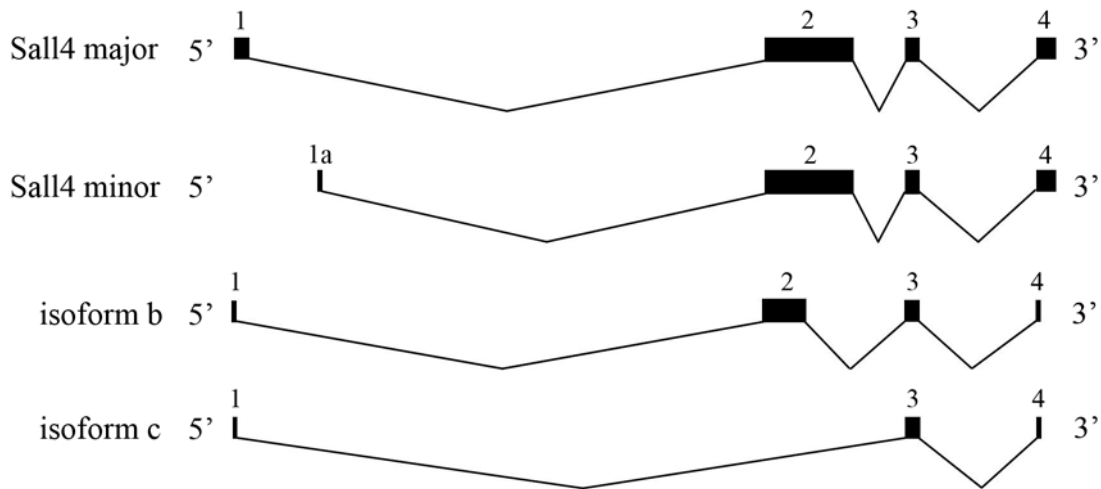


Figure 3.6: Schematic representation of the genomic organization and splicing forms of the *Sall4* gene. The *Sall4* major transcript consists of 4 exons. The *Sall4* minor transcript variant contains alternative exon 1a instead of exon 1. Isoform b consists of 4 exons, but includes only 5' part of exon 2. Isoform c includes 3 exons completely missing exon 2.

3.3 Expression analysis of *Sall4*

3.3.1 Expression of *Sall4* gene in preimplantation embryos

The expression of *Sall4* gene in blastocyst stage embryos was studied by X-gal staining (Fig. 3.7). Embryos were obtained from crosses between *Sall4* heterozygous males and wild type females. Blue β -galactosidase staining in blastocyst stage embryos was detected in the inner cell mass and in the trophectoderm of heterozygous embryos. The staining was not detected in wild type embryos.

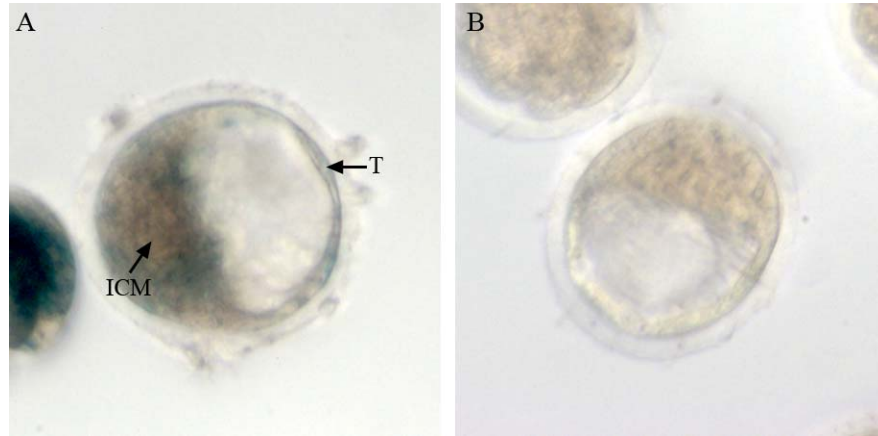


Figure 3.7: Expression of *Sall4* in mouse preimplantation stage embryos. (A) Expression of β -galactosidase was detectable both in the inner cell mass (ICM) and in the trophoblast (T) of heterozygous embryos. (B) Wild type embryos were used as a negative control.

3.3.2 Expression of *Sall4* gene in postimplantation embryos

Sall4 expression studies using whole mount *in situ* hybridisation on mouse embryos were previously described by Kohlhasse et al. (2002b). X-gal staining was performed in order to confirm the expression pattern of *Sall4* and to show the proper function of the gene trap vector in tissues where the “trapped” gene is known to be expressed.

X-gal staining was detectable as early as E7.5 stage (Fig. 3.8). Strong *Sall4* expression was found in the primitive streak, cardiac crescent and lateral plate mesoderm. At E8.5 strong expression was found in the cephalic neural fold, early somites, primitive streak and allantois. At E9.5 stage, when the neural tube begins to close at the level of forebrain/midbrain, *Sall4* expression became more confined to the midbrain, but strong expression remains in the more caudal region of neural tube. Transient *Sall4* expression was observed in the region of first branchial arch, region of future heart outflow tract (Fig 3.8). By the E11.5 stage, diminished *Sall4* expression was observed in several embryonic regions, including midbrain, rhombomers, limb bud mesenchyme and genital papilla.

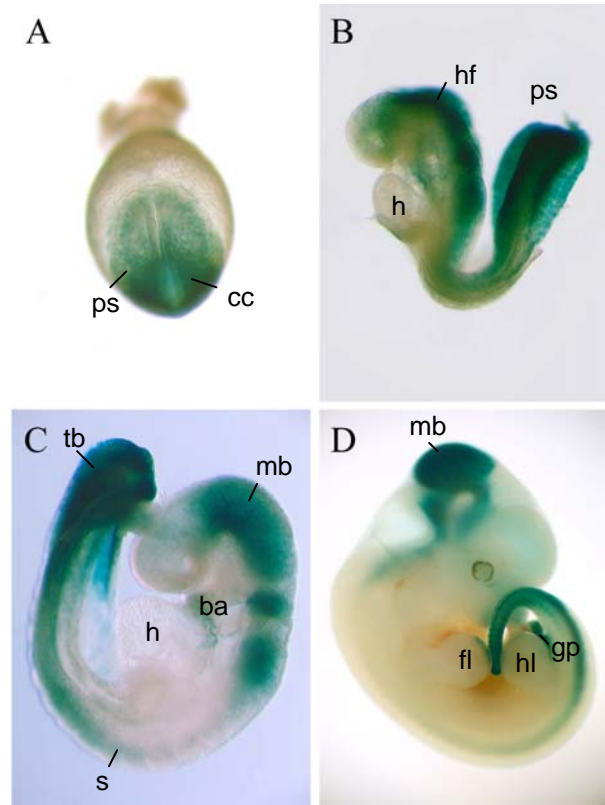


Figure 3.8: Whole mount X-gal staining of *Sall4* heterozygous embryos during early postimplantation development. (A) E7.5, *Sall4* expression is in the primitive streak. (B) E8.5, *Sall4* expression is in the neural fold, early somites, primitive streak and allantois. (C) E9.5, *Sall4* expression became more confined to the midbrain and strong expression remains in the caudal region of neural tube. (D) E11.5, *Sall4* expression remains in the midbrain, limb buds mesenchyme and genital papilla. ps – primitive streak, cc – cardiac crescent, hf – head fold, h – heart, tb – tail bud, s – somites, mb – midbrain, ba – branchial arch, fl – fore limb bud, hl – hind limb bud, gp – genital papilla.

To verify in which cell types *Sall4* is expressed, embryos were sectioned after whole mount X-gal staining on a cryotome and counterstained with Fast red (Fig. 3.9). In E7.5 embryos X-gal staining was observed in the embryonic mesoderm (region of future myocardial plate) and allantois. Some amnionic cells were also X-gal positive. *Sall4* expression was also seen in the neural epithelium and was particularly notable in the neural crest cells immigrating from the hindbrain of E8.5 stage. The strongest expression of *Sall4* during neural crest cell migration was found in the dorsal and lateral portion of the branchial arch. At the spinal cord level *Sall4* is expressed throughout the neural tube, mostly in the dorsal region. At later E8.75 stage, X-gal staining was found only in those neural crest cells, which will form pericardial cavity and

Results

outflow tract of the heart. In the neural fold *Sall4* is expressed in the medial part in neural epithelium and caudal dorsal neural tube.

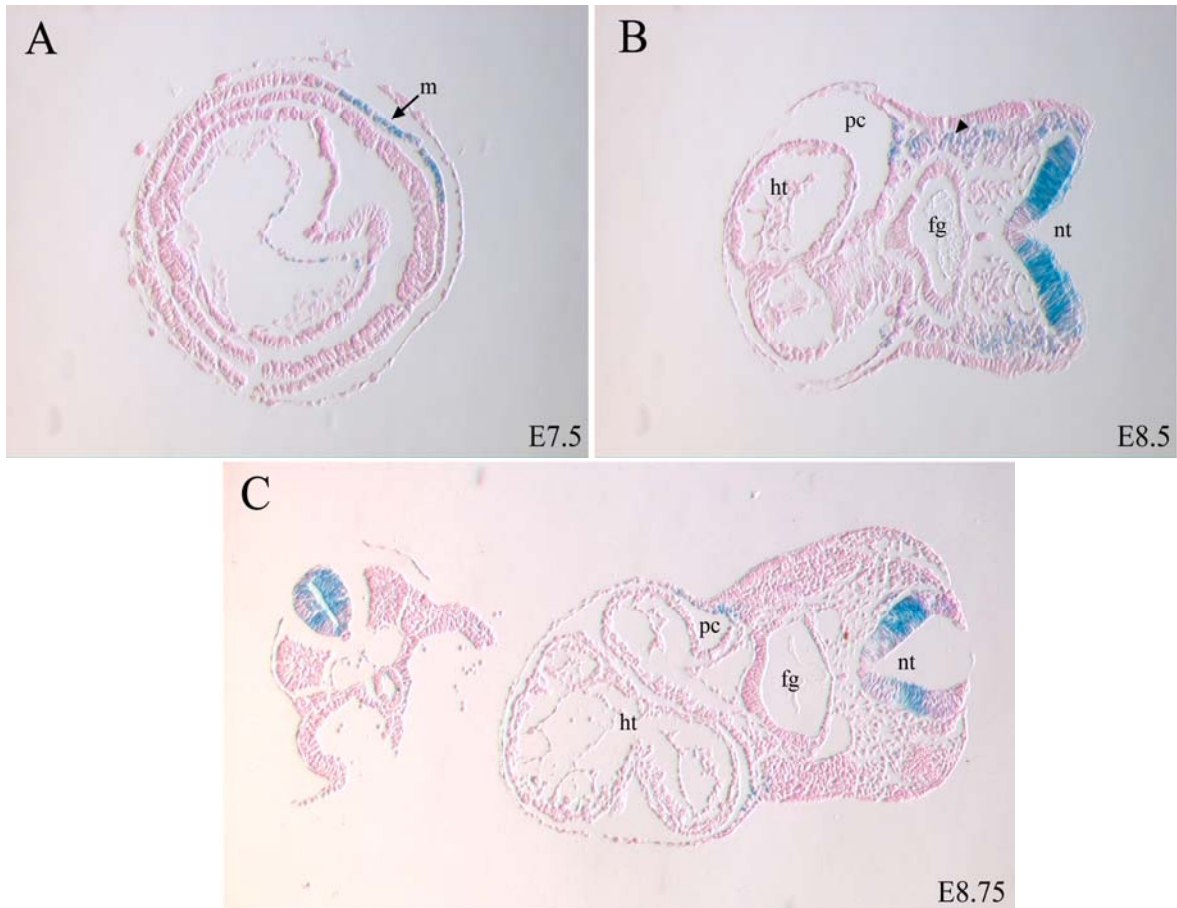


Figure 3.9: Transverse cryosections through whole mount *X-gal* stained *Sall4* embryos, counterstained with Fast Red. (A) E7.5, *Sall4* expression in the mesoderm and allantois. (B) E8.5, *Sall4* expression is notable in the neural epithelium and in migrating neural crest cells. (C) E8.75, *Sall4* is expressed in the medial part on neural epithelium, caudal neural tube and in the migrating neural crest cells likely forming pericardial cavity and outflow tract. Migrating neural crest cells are marked with arrowhead. m; mesoderm, pc; pericardial cavity, ht; heart, fg; foregut, nt; neural tube.

3.3.3 RT-PCR analysis of *Sall4* expression during embryonic development

The tissue-specific expression of the mouse *Sall4* gene was first determined by Northern blot analysis with total RNA prepared from different tissues. A single transcript of size 5.4 kb was observed only in testis and ovary (Kohlhase et al., 2002b). Here, expression analysis of *Sall4*

Results

during embryonic development was performed by RT-PCR with total RNA of mouse embryos from different stages (Fig. 3.10).

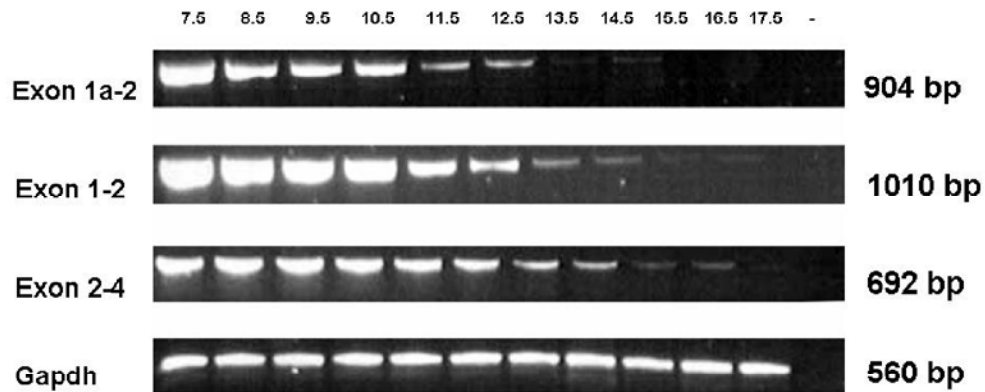


Figure 3.10: RT-PCR analysis with total RNA isolated from whole mouse embryos of different developmental stages using *Sall4* specific primers. Primers S4ME1F1 and S4ME2R2 were used to amplify exon 1-2 transcript, primers S4ME1aF1 and S4ME2R2 – to amplify exon 1a-2 transcript and primers S4ME2F2 and S4ME4RTR – to amplify exon 2-4 transcript. The splice variants *Sall4* major (exon 1-2) and *Sall4* minor (exon 1a-2) are compared. Exon 2-4 RT-PCR detects both variants. RT-PCR for *Gapdh* transcript was performed as a control for RNA amount and integrity. Note that strongest expression of *Sall4* minor variant was found at E7.5 and almost not detectable at E14.5. *Sall4* major expression was observed until E16.5. Exon 2-4 transcript expression was detectable until E16.5.

Analysis of embryonic mRNAs by RT-PCR confirmed expression of transcripts derived from the two alternative 5' coding exons. Expression for both transcripts was detectable as early as E7.5 and gradually decreases until E17.5. *Sall4* minor variant showed a limited expression period and was almost not detectable at E14.5 stage. The expression of *Sall4* major transcript variant was observed until E16.5. However the expression of exon 2-4 transcript which detects both variants was detectable until E16.5.

3.3.4 Expression of splicing variants in *Sall4* mutant mice

In order to determine whether the gene trap vector insertion leads to a null mutation or to a hypomorphic allele, quantitative real-time RT-PCR was performed. Total RNA was isolated from E8.5 stage embryos and cDNA was prepared by reverse transcription using oligo(dT) primers. The expression of *Sall4* transcripts was analysed by using a SYBR Green assay and a

Results

standard curve method as described in chapter 2.2.7. The expression level of investigated transcripts was normalised to β -actin mRNA expression to compensate for difference in sample amounts (Fig. 3.11).

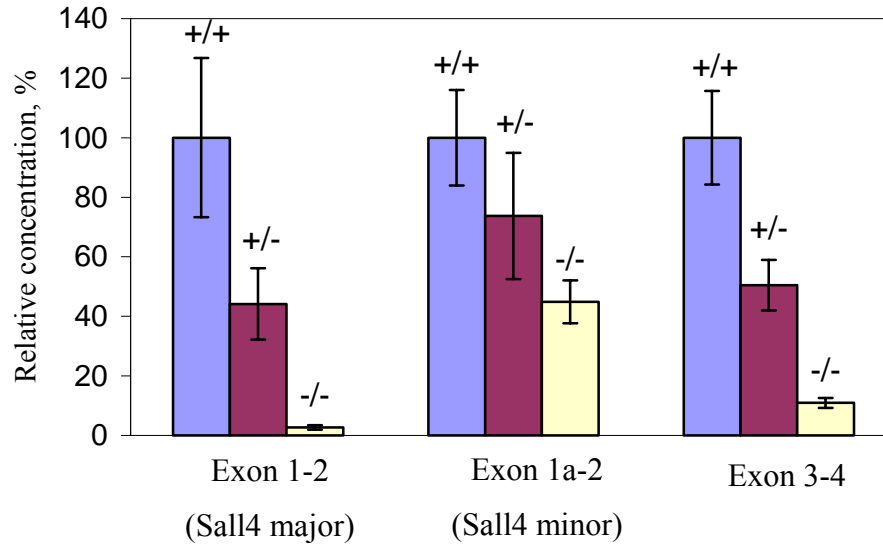


Figure 3.11: Total RNA was isolated from E8.5 wild type (+/+), heterozygous (+/-) and homozygous (-/-) *Sall4* embryos and analysed by real-time RT-PCR. The expression level was normalised to β -actin expression. Expression levels in wild type embryos for each transcript are set to 100%. Exon 3-4 product detects both transcripts. Note that expression of major transcript in +/- embryos is 58% less and in -/- embryos is 98% less than in wild type. The expression level in case of minor transcript is reduced by 27% in +/- embryos and by 55% in -/- embryos as compare to wild type. Exon 3-4 transcript represents total amount of both transcripts and is reduced in +/- embryos to 50% and in -/- embryos to 10%.

Relative expression of *Sall4* major and *Sall4* minor transcripts was compared between wild type, heterozygous and homozygous mutant mice. Real-time RT-PCR analysis revealed that the level of the major transcript in heterozygous animals was reduced by 58% as compared to wild type; in contrast, the level of reduction in case of minor transcript was only 27% in heterozygous animals. Real-time RT-PCR showed presence of approximately 2% of major transcript and 45% of minor transcript in *Sall4* homozygous embryos. Exon 3-4 transcript represents total amount of both transcripts and was reduced in +/- embryos to 50% and in -/- embryos to 10%.

3.4 Phenotype of *Sall4* mutant mice

After identification of the genomic localization of the gene trap insertion, primers were designed for a genotyping PCR. The gene trap construct had been electroporated into ES cells originating from the mouse strain 129X1/SvJ, but the mouse line created was crossed to the C57BL/6J background. Therefore, the genome from the *Sall4* mouse line was of the mixed 129X1/SvJxC57BL/6J background.

ES cells from the GGTC line W097E01 were injected into inner cell mass of blastocysts derived from 129X1/SvJ mice at the GGTC facility. Chimeras were bred with C57BL/6J mice to obtain F1 animals. We received such F1 animals. Further breeding was performed by mating F1 animals with C57BL/6J wild type animals. The germline transmission was checked by genomic PCR with S4MI1R1a, pGeoR6, S4I1F1 primers (Fig. 3.12).

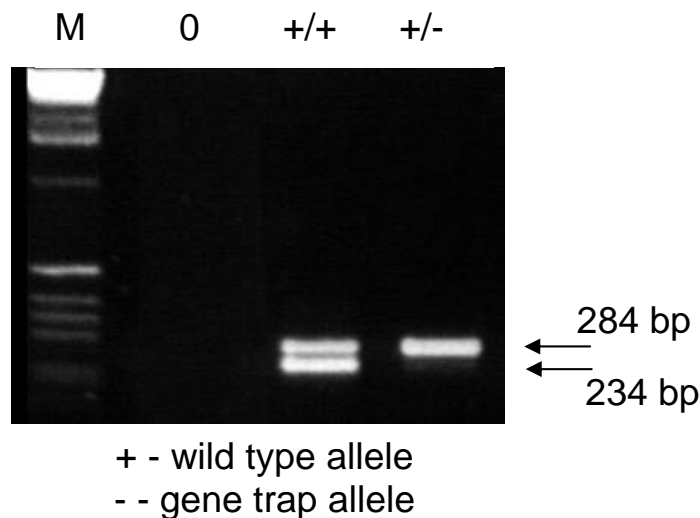


Figure 3.12: PCR genotyping of the *Sall4* mouse gene trap line W097E01. The wild type allele for *Sall4* is amplified by S4I1F1 and S4MI1R1a primers resulting in a PCR product of 284 bp. The trapped *Sall4* allele expected size is 234 bp using the primers S4I1F1 and pGeoR6. +/+ - wild type; +/- - heterozygous; 0-negative control, M-1 kb marker.

3.4.1 Early embryonic lethality of *Sall4* mice

We chose the *Sall4* gene trap mouse line in order to investigate the function of *Sall4* and to use it as a model system for Okihiro syndrome. *Sall4* heterozygous mice obtained from backcrossing with wild type animals were firstly checked for the presence of anomalies

Results

similar to the symptoms observed in Okihiro syndrome patients. This would imply defects of the upper limb, especially on the radial side, heart defects, renal anomalies and a missing 6th cranial nerve or its nucleus. The morphological and histological analysis of the brain as well as a histological analysis of heart, liver, muscle and kidney was done by Dr. Walter Schulz-Schaeffer, Dept. of Neuropathology, University of Göttingen, and revealed no abnormalities. Gross morphological analysis of the heterozygous mice revealed no abnormalities in skeleton, brain, kidney, muscles and heart. Heterozygous mice were also fertile. Genotyping analysis of the F2 generation revealed that 37% of the offspring were wild type and 63% were heterozygous. No homozygous mutant mice were detected suggesting embryonic lethality (Table 3.2).

	+/+	+/-	-/-	total
♀	24	32	0	56
♂	22	44	0	66
total	46 (37%)	76 (63%)	0 (0%)	122

Table 3.2: Genotyping of newborn offspring resulting from intercross breeding of *Sall4* heterozygous mice.

To determine the time of embryonic lethality, embryos of different developmental stages were isolated from heterozygous intercrosses and genotypes were determined by PCR (Table 3.3).

	+/+	+/-	-/-	total
E3.5	14	20	8	42
E7.5	7	9	1*	17
E8.5	10	30	9*	49
E9.5	8	21	8†	37
E10.5	1	3	2†	6

Table 3.3: Genotyping of the embryos from different developmental stages from breeding of *Sall4* heterozygous mice. * - malformed embryos, † - reabsorbed embryo tissue.

At stage E3.5, 8 homozygous embryos were found, indicating that development of the embryos during preimplantation step occur normal. At the later stages (E7.5-E8.5)

homozygous embryos were smaller and malformed. After stage E9.5, only residual parts of the amnion of homozygous embryos were present in the uterus. *Sall4* homozygous mutant embryos were developing normally up to E7.0 stage, thereafter development was severely retarded and they were reabsorbed by E9.5 stage.

3.4.2 Phenotype of *Sall4* homozygous embryos

Sall4 homozygous embryos appeared to develop normally until E7.0, but failed to develop properly during later stages of development. Homozygous embryos were smaller and showed a lack in the development and formation of primary structures in the embryos (Fig. 3.13).

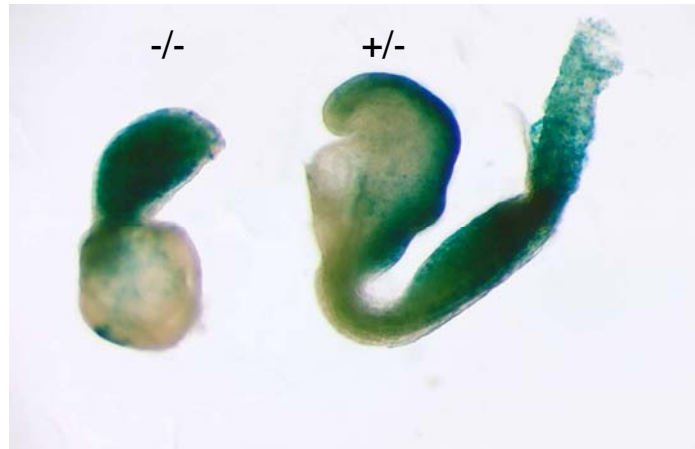


Figure 3.13: Morphology of heterozygous and homozygous mutant *Sall4* embryos at E8.5 using X-gal staining. The homozygous mutant embryo appears severely malformed and smaller than the heterozygous due to developmental arrest.

3.4.3 Histological analysis of *Sall4* embryos

In order to understand the underlying cause of the early embryonic lethality, *Sall4* homozygous mutant embryos were analysed histologically at different stages and compared with wild type littermates. E7.5 embryos were smaller than their littermates and demonstrated abnormalities in primitive streak, amnionic cavity and allantois (Fig. 3.14)

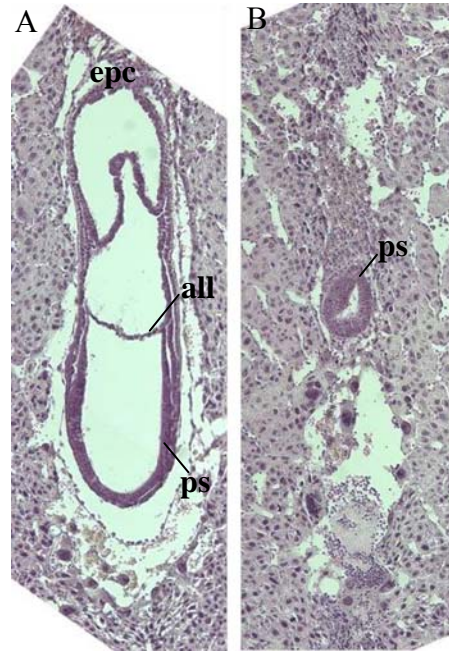


Figure 3.14: Histological analysis of *Sall4* E7.5 embryos. Sagittal sections of paraffin embedded E7.5 wild type (A) and homozygous (B) embryos were stained with haematoxylin-eosin. all, allantois; epc, ectoplacental cone; ps, primitive streak.

Homozygous mutant E8.5 embryos appeared developmentally retarded as compared to their littermates (Fig. 3.15). The neural tube of the homozygous E8.5 embryos remained open. The foregut region which plays an important role during heart development was deformed in shape. However the most severe phenotypic abnormality was in the heart region. Normal heart development starts with two heart fields at E7.5 which are fusing at later stages to form a tubular heart. In E8.5 wild type embryos, parts of the heart were already fused forming a primitive heart with a pericardial cavity (Fig. 3.15A). In contrast, in homozygous mutant embryos the heart was not formed properly, the heart bulbs were not fusing and the pericardial cavity was underdeveloped. The embryonic lethality of *Sall4* homozygous mutant embryos during early gestation could therefore result from cardiac bifida together with the deformation of the foregut and missing neural tube closure.

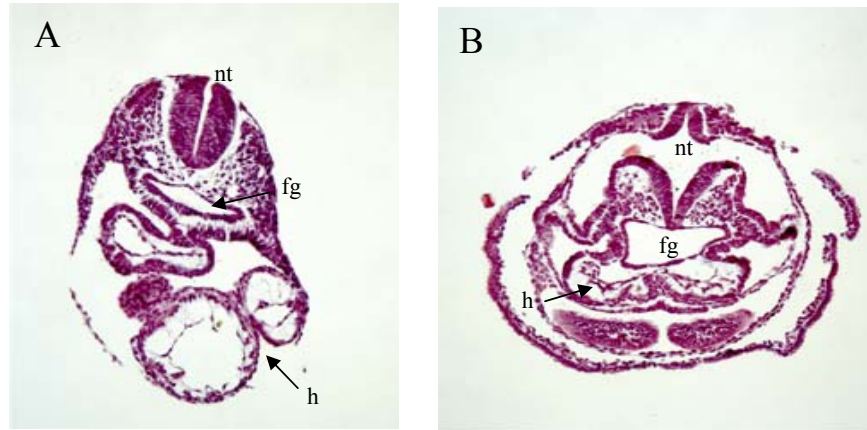


Figure 3.15: Haematoxylin-eosin staining of E8.5 embryos. Transverse sections through wild type (A) and *Sall4* homozygous mutant (B) embryos. Homozygous embryos fail to form a primitive heart by fusion of two heart bulbs. nt, neural tube; fg, foregut; h, heart..

3.4.4 Analysis of cellular proliferation in *Sall4* embryos

To identify whether the *Sall4* mutant phenotype was due to a cell proliferation problem, BrdU labelling was performed. BrdU labelling provides a simple technique for the identification of proliferating cells. To measure the BrdU incorporation in mouse embryos, pregnant females were given a single dose of BrdU through intraperitoneal injection (chapter 2.2.14). No difference in cellular proliferation between wild type and homozygous mutant embryos was observed (Fig. 3.16).

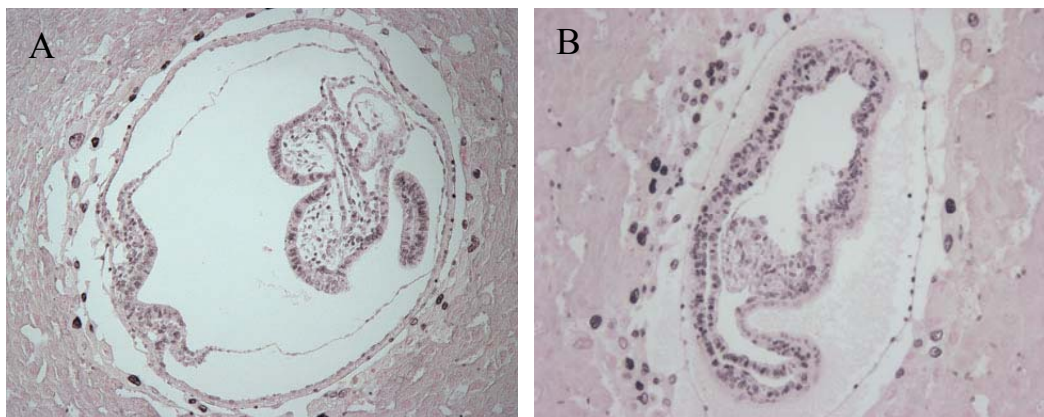


Figure 3.16: BrdU staining of formalin fixed, paraffin embedded E8.5 embryos. No obvious difference between wild type (A) and homozygous (B) embryos was observed.

Results

To check whether *Sall4* mutant phenotype was caused by developmental arrest or by a proliferation problem, E3.5 embryos were cultivated for 5 days. The phenotype of the cultured blastocysts was documented daily by inverted microscope (Fig. 3.17). In total, 25 blastocysts were analysed. There was no significant difference in cell growth and development of both inner cell mass and trophoblast between wild type, heterozygous and homozygous blastocysts. This result suggests that during preimplantation step *Sall4* mutant embryos are developing normal.

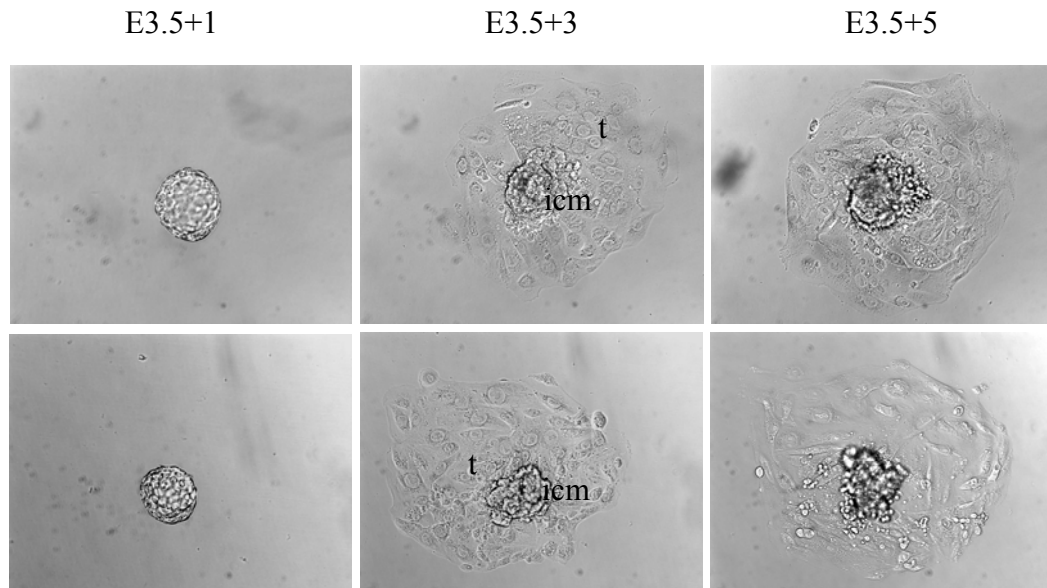


Figure 3.17: Growth of *Sall4* embryos in the culture. Blastocysts (E3.5) derived from heterozygous intercrosses were cultured for 5 days and photographed at the 2nd (E3.5+1), 3^d (E3.5+2) and 5th day (E3.5+5) of culture. icm, inner cell mass, t, trophoblast.

3.5 *Sall4* regulates the expression of *Pax3* and T-box genes

Numerous signalling molecules and pathways regulate early embryogenesis. To determine which genes are misregulated in *Sall4* mutant embryos, we analysed their expression using whole mount *in situ* hybridisation and real-time RT-PCR approach.

Pax3 could serve as early embryonic marker of neuroepithelium for *in situ* hybridisation studies. Mutations in *Pax3* cause open neural tube phenotype (Epstein et al., 1991), which was also observed in *Sall4* mutant embryos. In the wild type embryos *Pax3* expression was detected in the dorsal part of the neural tube and in the lateral epithelium of the proencephalon (Fig. 3.18). In the homozygous mutant embryos *Pax3* expression was not detectable.

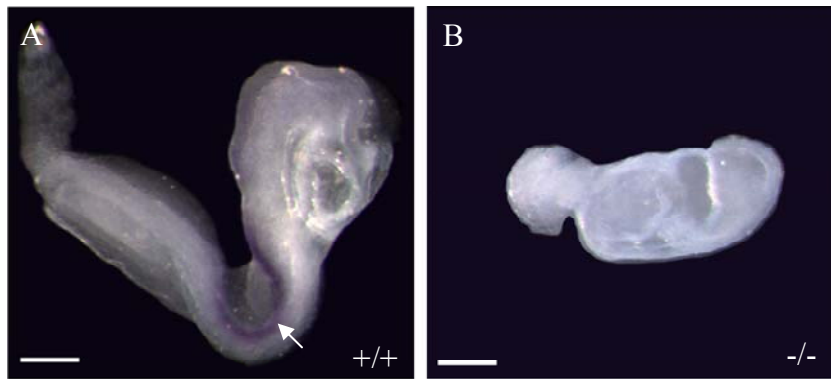


Figure 3.18: *Pax3* expression in wild type (A) and in *Sall4* homozygous mutant (B) E8.5 embryos analyzed by *in situ* hybridization. (A) *Pax3* expressed in the dorsal part of the neural tube (arrow) and in the proencephalon. (B) No *Pax3* expression in mutant embryo was observed. Scale bar 300 μ m.

The *brachyury* (*T*) gene is an early mesodermal marker involved in notochord formation (Wilkinson et al., 1990). The phenotype of *brachyury* mice includes disruption of trunk development and embryonic death in homozygotes. In wild type and heterozygous E7.5 embryos *brachyury* is exclusively expressed in the mesoderm of the primitive streak (Fig. 3.19A, B). In contrast, in homozygous *Sall4* mutant embryos *brachyury* expression was diffuse and ectopically localised (Fig. 3.19C, D).

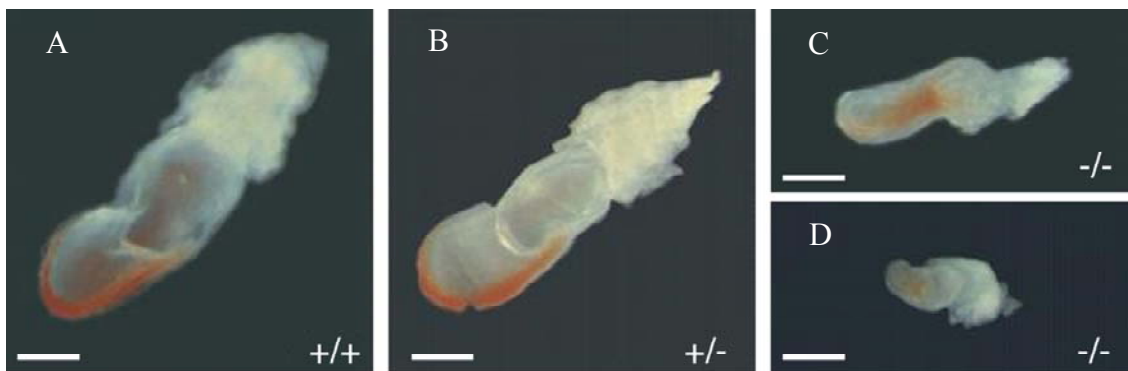


Figure 3.19: Expression of *brachyury* in E7.5 embryos by whole mount *in situ* hybridization (detection with INT/BCIP). Wild type (A) and heterozygous *Sall4* mutant (B) embryos show similar expression patterns in the mesoderm of primitive streak. Note diffuse misexpression in homozygous mutants (C, D). Scale bar 200 μ m.

The genes *Tbx1* and *Tbx5* like *brachyury* belong to the T-box gene family (Tbx) and have been implicated in early cardiac lineage determination, chamber specification, and valvuloseptal development (Chapman et al., 1996; Simon, 1999). Heterozygous mutation of

Results

mouse *Tbx5* causes aberration of cardiac development (linear heart) and embryonic lethality as well as severe upper limb defects (Bruneau et al., 2001; Zhou et al., 2005). *Tbx1* mutations have been implicated in the etiology of cardiac outflow tract and pharyngeal arch anomalies (Merscher et al., 2001). The regulation of these genes plays an important role during primitive heart formation. Therefore, expression of the T-box genes *brachyury*, *Tbx1* and *Tbx5* was examined in *Sall4* mutant embryos. At E8.5 stage, embryos were collected in separate tubes. RNA and DNA of single embryos were isolated using TRI reagent. Embryos were genotyped and cDNA was synthesised with oligo(dT) primers. The expression of *brachyury*, *Tbx1* and *Tbx5* was analysed by quantitative real-time RT-PCR using SYBR Green dye. Expression level in all samples was normalised to β -actin expression (Fig. 3.20).

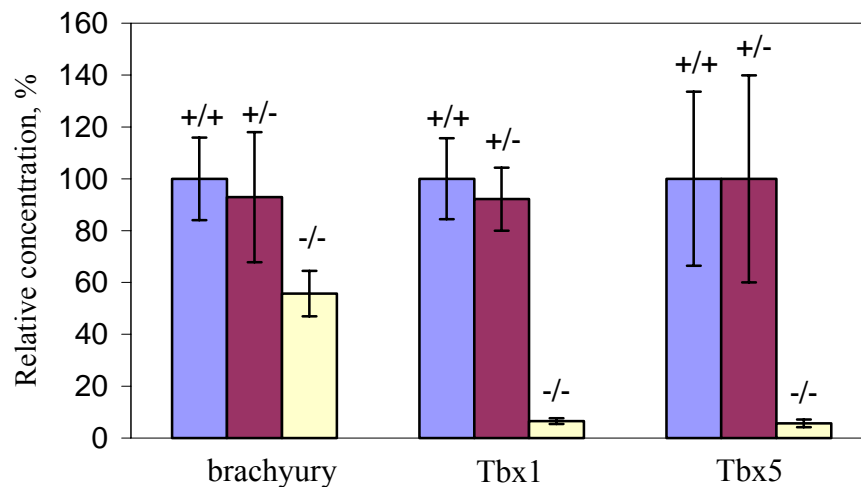


Figure 3.20: Total RNA was isolated from E8.5 wild type, heterozygous and homozygous *Sall4* mutant embryos and analysed by real-time RT-PCR for *brachyury*, *Tbx1* and *Tbx5* expression relative to β -actin expression. Expression levels in wild type embryos for each gene are set to 100%.

Expression of the three T-box genes was largely reduced in the *Sall4* homozygous embryos. *Brachyury* expression level was diminished by 45%. Expression of *Tbx1* and *Tbx5* was lowered to around 6% of the wild type level. In contrast, expression of these genes was almost unaffected in *Sall4* heterozygous embryos.

3.6 *Sall4* antibody generation

Polyclonal antibody was generated against *Sall4* fusion protein as described in methods part. In order to determine the affinity and specificity of the antiserum, Western blot analysis was

Results

performed. Total protein extracts from several tissues and from whole E9.5 embryos was used to detect Sall4 expression. The antiserum derived from rabbit showed a lot of unspecific bands and the affinity of the antibody against the Sall4 epitope was very weak (Fig. 3.21A). Protein G purification was performed to enrich IgG fraction of antiserum and to increase the concentration of the Sall4 antibody (Fig. 3.21B), but the titer of antibody was too weak to get a consistent result.

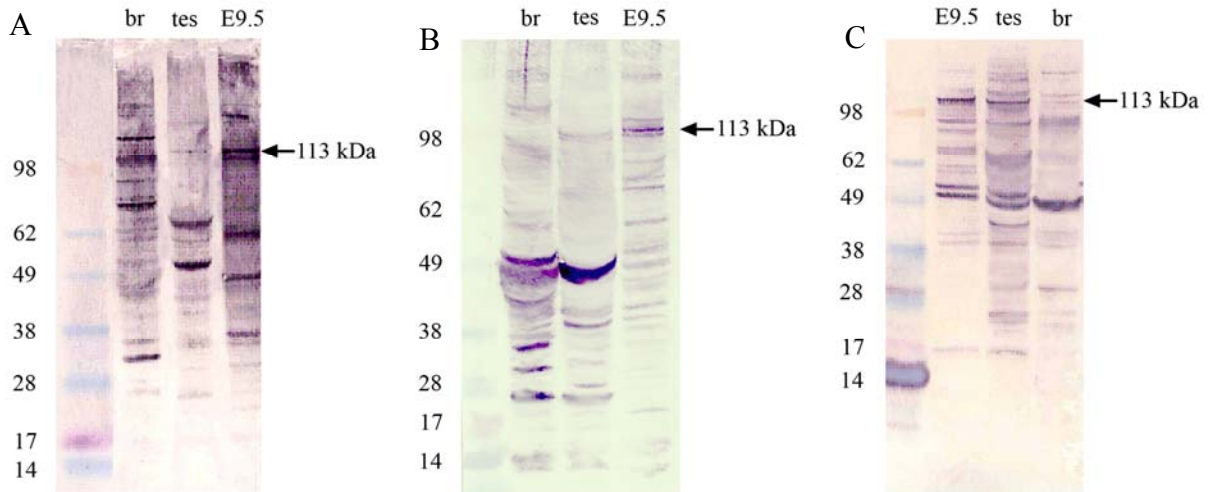


Figure 3.21: Western blot analysis with polyclonal antibody against Sall4 protein. In this analysis antibody dilution 1:150 was used. (A) Unpurified antiserum showed high cross-reactivity to other proteins. (B) Protein G purified antibody. (C) Affinity purified antibody detected a 113 kDa band corresponding to the predicted Sall4 protein size. br, brain; tes, testes; E9.5, embryo E9.5 stage.

In order to obtain a high titre antibody, affinity purification of antiserum was performed. The fusion protein used for rabbit injections was coupled with NHS-HiTrap column and antiserum was passed through the column. Antibody fractions were eluted from the column, dialysed against PBS and concentrated using a Centriscart column. A specific band corresponding to molecular weight of 113 kDa of Sall4 was detected by Western blot (Fig. 3.21C). To check the specificity of anti-Sall4 antibodies, a Western blot analysis with the F9 cell line, which expresses Sall4 endogeneously, was performed (Fig. 3.22). A single specific band of 113 kDa was detected.

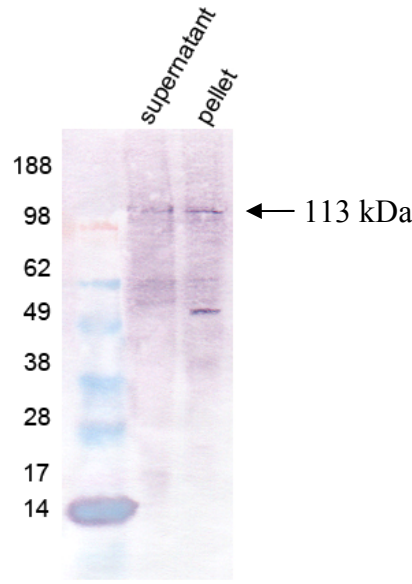


Figure 3.22: Western blot analysis of F9 cell line with polyclonal antibody against Sall4 protein. In this analysis an antibody dilution of 1:150 was used. Affinity purified antibody detected a 113 kDa band corresponding to the Sall4 protein.

3.7 Sall4 and Sall1 interaction studies

3.7.1 Interaction on the genetic level

The genes *Sall1* and *Sall4* belong to the family of SAL-like genes, named after the gene *sal* (*spalt*) of *D. melanogaster* (Kühnlein et al., 1994). They show an overlapping expression pattern during development (Buck et al., 2000, Kohlhase et al., 2002b). From the previous studies it was known that some of the *SALL* family genes can interact *in vitro* with other *SALL* family members in mouse and chicken (Kiefer et al., 2003, Sweetman et al., 2003) and form homo- and heterodimeric complexes.

To further evaluate the functional interaction between *Sall1* and *Sall4* we initiated generation of double knockout mice. A *Sall1* knockout mouse line on 129X1/SvJ background had been previously established in our group (Buck, 2002). *Sall1* heterozygous mice appeared to be normal and homozygous mice were dying in the perinatal period because of severe kidney defects leading to uremia, analogous to the phenotype of the published *Sall1* knock-out mice (Nishinakamura et al., 2001). In *Sall4* heterozygous mutant mice no obvious phenotype was observed and homozygous mice were dying before E7.5. The deviation of genotype to

Results

Mendelian inheritance can reflect an interaction between two genes. In double heterozygous breedings it was expected to observe changes in the lethality or phenotype of *Sall1*(-/-)/*Sall4*(+/-) and *Sall1*(+/-)/*Sall4*(-/-) offspring (Fig. 3.23).

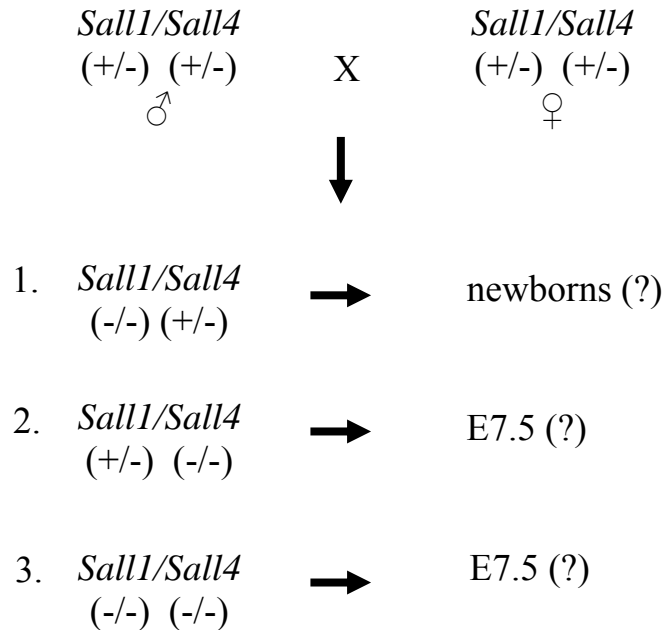


Figure 3.23: Schematic representation of the strategy to generate *Sall1*/*Sall4* double mutant mice.

To generate double heterozygous animals, *Sall1* heterozygous mutants and *Sall4* heterozygous mutants were crossed. However, genotyping analysis of *Sall1*/*Sall4* offspring revealed an obvious deviation from Mendelian inheritance ratio (Table 3.4).

Sall1 (+/-) X *Sall4* (+/-)

Genotype	<i>Sall1</i> / <i>Sall4</i> (+/+) (+/+)	<i>Sall1</i> / <i>Sall4</i> (+/+) (+/-)	<i>Sall1</i> / <i>Sall4</i> (+/-) (+/+)	<i>Sall1</i> / <i>Sall4</i> (+/-) (+/-)
Mendelian ratio	25%	25%	25%	25%
Offspring number	24	24	25	2
Offspring ratio	32%	32%	33.3%	2.7%

Table 3.4: Genotyping analysis of *Sall1* +/- mutant/ *Sall4* +/- mutant breeding. The total amount of genotyped animals was 76. P<0.0001

Sall1(+/-)/*Sall4* (+/-) was noticeably reduced. The two-tailed P value was less than 0.0001. By conventional criteria, this difference is considered to be extremely statistically significant.

3.7.2 Co-localization of *Sall4* with truncated *Sall1* protein in F9 cells

The N-terminally truncated *Sall1* protein, expected to be generated by typical *SALL1* mutations in Townes-Brocks syndrome patients, can exert a dominant negative effect (Sato et al., 2004). Truncated *Sall1* protein is able to interfere with the correct function of native *Sall1* protein by causing its displacement from the nucleus (Sweetman et al., 2003). For the interaction study, a *Sall1* N-terminal part encoding amino acids 1-435 was cloned into pcDNA3.1 vector. The subcellular distribution of the endogenous *Sall4* protein and the possible co-localisation with the N-terminally truncated *Sall1* (*Sall1tr*) protein in mouse F9 teratocarcinoma cells were analysed by immunocytochemical studies. F9 cells were transiently transfected with the myc-tagged *Sall1tr* construct and the expression of proteins was determined by staining with anti-*Sall4* and anti-c-myc antibodies. A signal of the *Sall4* protein was detected exclusively in the nucleus of the stained cells (Fig. 3.24).

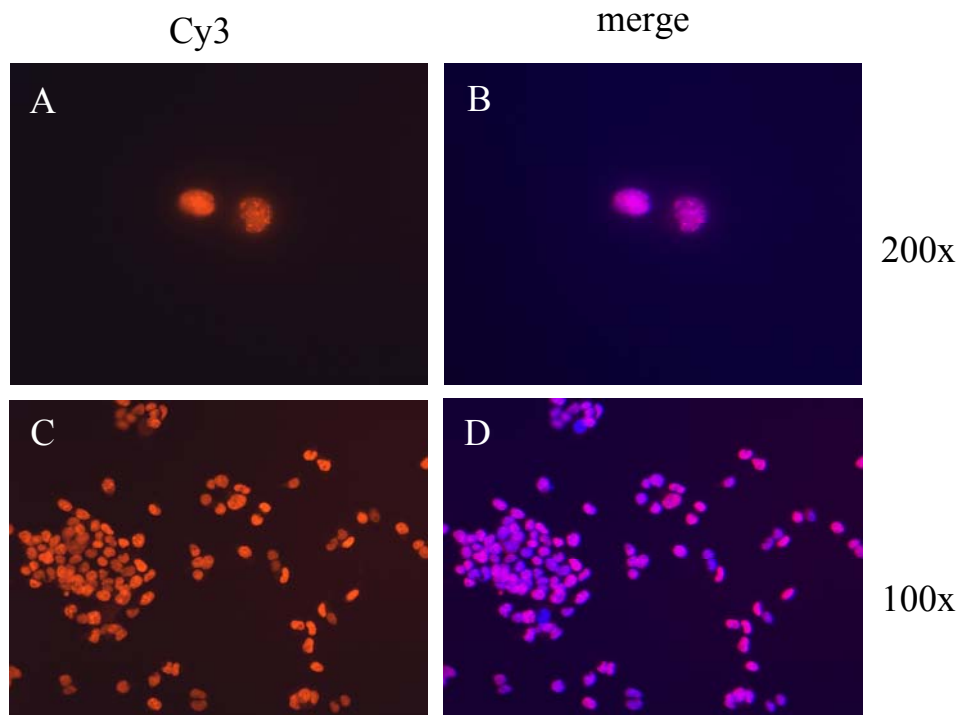


Figure 3.24: Subcellular localization of the *Sall4* protein in F9 cells. (A, C) F9 cells stained with polyclonal anti-*Sall4* antibody and anti-rabbit Cy3-coupled antibody; (B, D) merged picture with DAPI staining.

Results

The staining pattern for the Sall4 in the nucleus is similar to that known for other members of Sall family present in a punctuate pattern probably representing pericentromeric heterochromatin (Netzer et al., 2001; Kiefer et al., 2002; Sato et al., 2004). In contrast, transient transfection of F9 cells with myc-tagged Sall1tr construct revealed a clear change in the subcellular localization of Sall4 protein. In the presence of Sall1tr, Sall4 was detected in the cytoplasm co-localising with Sall1tr protein (Fig. 3.25). This result suggests that Sall1tr can interact with Sall4 protein by changing localisation of the endogenous Sall4 protein.

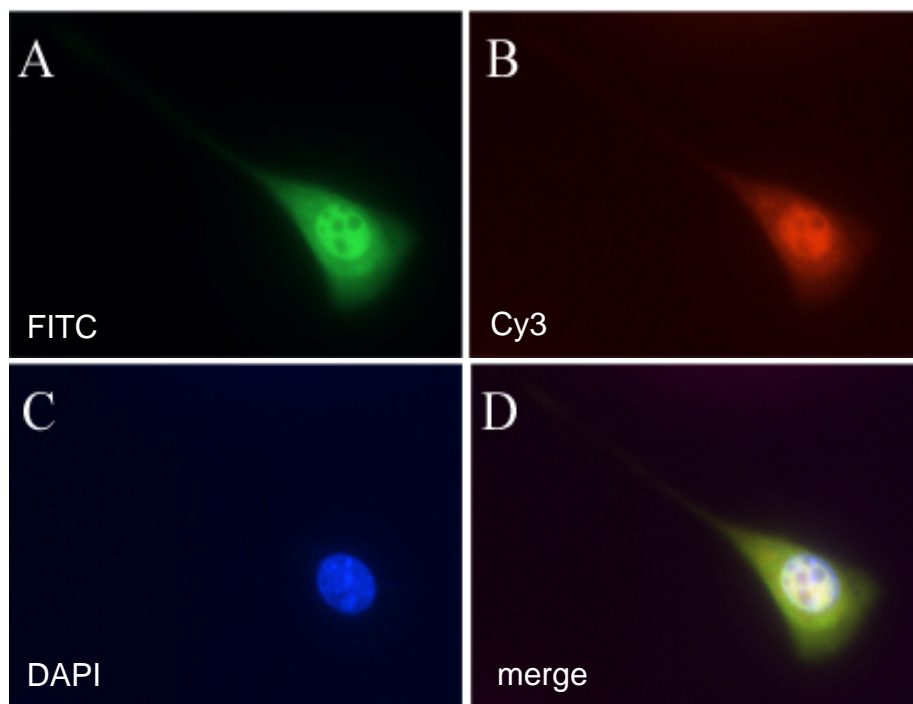


Figure 3.25: Endogenous Sall4 co-localised in the cytoplasm with truncated Sall1 protein in F9 cells. (A) Truncated Sall1 protein was detected by staining with polyclonal anti-myc and anti-rabbit FITC-conjugated antibody. (B) Sall4 protein with polyclonal anti-Sall4 antibody and anti-rabbit Cy3-coupled antibody. (C) DAPI. (D) merged picture.

4. DISCUSSION

The present work discusses the identification and characterisation of *Sall4* gene trap mouse line. The following results were obtained:

- *Sall4* gene trap ES cell line was identified by comparing *Sall4* sequences to the database of the German Gene Trap Consortium. This ES cell line was used for blastocyst injection and generation of *Sall4* gene trap mouse line.
- New splicing isoform of *Sall4*, named *Sall4* minor, was isolated by EST database analysis and confirmed by RT-PCR and sequencing.
- Expression of *Sall4* major and *Sall4* minor transcript variants during embryonic development was analysed. It was demonstrated that expression for both transcripts was detectable as early as E7.5 and gradually decreases until E17.5. *Sall4* minor variant was expressed until E14.5 stage, while the expression of *Sall4* major transcript variant was observed until E16.5.
- During embryonic development *Sall4* expression was found at blastocyst stage in the inner cell mass and in the trophectoderm. At later stages *Sall4* expression was found in the primitive streak, neural tube, developing brain, branchial arches and limb buds.
- X-gal staining revealed expression of *Sall4* in the intraembryonic mesoderm, neural epithelium and migrating neural crest cells.
- Expression of *Sall4* major and *Sall4* minor transcript variants in *Sall4* mutant mice was analysed by quantitative real-time RT-PCR. The expression level of major transcript in heterozygous embryos was 58% less and in homozygous embryos was 98% less than in wild type. The expression level in case of minor transcript was reduced by 27% in heterozygous embryos and by 55% in homozygous embryos as compared to wild type.
- *Sall4* heterozygous mice revealed no abnormalities, but homozygous embryos were severely retarded and die around E7.5.
- The embryonic lethality of *Sall4* homozygous mutant embryos during early gestation results from cardiac bifida together with the deformation of the foregut and missing neural tube closure.
- Whole mount *in situ* hybridisation revealed absence of the *Pax3* expression in *Sall4* homozygous mutant embryos and ectopical expression of *brachyury*.

- By real-time RT-PCR it was demonstrated that expression of T-box genes *brachyury*, *Tbx1* and *Tbx5* was downregulated in *Sall4* homozygous mutant embryos, but not in the heterozygotes.
- Generated polyclonal anti-Sall4 antibody can specifically detect Sall4 expression in the Western blot analysis and in immunocytochemical analysis.
- Sall4 and truncated Sall1 interact and co-localise in mouse F9 teratocarcinoma cells.
- Sall4-Sall1 interaction on the genetic level was demonstrated by crossing experiments. The number of double heterozygous Sall1/Sall4 offspring was significantly reduced.

4.1 *Sall4* expression analysis

It was previously reported that *Sall4* consists only of four exons and has one transcript variant (Kohlhase et al., 2002b). In the mammalian SALL family alternative splicing was previously described for mouse *Sall2* (Kohlhase et al., 2000b), human *SALL2* (Ma et al., 2001), *Msal-1* (*Sall3*) and human *SALL3* (Ott et al., 2001, Kohlhase et al., 1999b). Evidence is provided here for the existence of alternative splice variants for the murine *Sall4* gene. Four mRNA isoforms (*Sall4* major, *Sall4* minor, *Sall4* isoforms b, *Sall4* isoform c) are produced by alternative splicing of *Sall4*. Database analysis suggested the presence of an additional exon 1a in *Sall4* gene. Analysis of embryonic mRNA by RT-PCR confirmed that exon1 and exon 1a are utilised in the production of two distinct transcripts (Fig. 3.4). Two additional alternative *Sall4* transcripts were reported in the database. In the *Sall4* isoform b and isoform c 3' part of exon 2 and complete exon 2 are skipped, respectively (Fig. 3.6).

Polyclonal anti-Sall4 antibody generated in the present study was raised against C-terminal domain (360 aa) of Sall4 and should detect all isoforms. However, Western blot analysis revealed only one single band of 113 kDa corresponding to the predicted size of the *Sall4* major transcript (Fig. 3.22). Probably due to the small protein size difference between *Sall4* major and minor it is not possible to detect *Sall4* minor transcript. It remains to be proven whether isoform b and isoform c lead to functional proteins or whether they have tissue specific or different temporal expression pattern.

Sall4 major transcript includes exon 1 and is expressed as early as E7.5 together with *Sall4* minor (Fig. 3.10). Expression of *Sall4* minor with exon 1a is restricted to the E14.5 stage while the *Sall4* major transcript is expressed until E16.5. However, the expression of exon 2-4

transcript which detects both variants was detectable until E16.5. Overall during embryonic development the expression level of *Sall4* decreases. This observation could imply an important role of *Sall4* in initial steps of development like gastrulation and early organogenesis. The contribution of the two transcript variants to the *Sall4* expression level is not equal. The presence of two different regulatory elements which control the expression of *Sall4* major and *Sall4* minor could be proposed. The presence of two distinct tissue specific promoters to control transcription of *SALL2* splice variants was shown by Ma et al. (2001). They demonstrated that the P1 promoter exhibits a weaker activity only in the thymus, testis and colon, whereas P2 activity was significantly higher in the wide spectrum of tissues like fetal liver, leukocytes, thymus, tonsil, spleen, colon, prostate and testis (Ma et al., 2001). In conclusion, it has to be confirmed by additional experiments if alternative *Sall4* promoters exist, and which role they play in the tissue specific and stage specific expression of *Sall4*.

By X-gal staining, *Sall4* expression was detected in preimplantation stage embryos at the blastocyst stage. *Sall4* expression was observed in the inner cell mass of the embryo which gives rise to the three germ cell layers ectoderm, mesoderm and endoderm, and also in the trophoctoderm cells which form extraembryonic tissues including placenta. *Sall4* homozygous blastocysts look normal at E3.5 and after 5 days in culture which would imply that *Sall4* is not required for trophoctoderm formation. However, the observed *Sall4* expression in trophoctoderm might be sufficient to support trophoctoderm differentiation on later developmental stages. This observation might play an important role in interpretation of the *Sall4* mutant phenotype.

Expression analysis in postimplantation embryos by means of X-gal staining revealed a pattern similar to the previously reported expression of *Sall4* (Kohlhase et al., 2002b). During embryonic development the expression pattern of *Sall4* in the nervous system, branchial arches and tail tip is comparable to the expression in the chick and *Xenopus* embryos (Barembaum and Bronner-Fraser, 2004, Neff et al., 2005). At E8.5 *Sall4* is expressed in the medial part in neural epithelium and in the caudal dorsal neural tube. At later stages *Sall4* expression is confined to the midbrain region and mesenchyme of the limb buds. Limb expression was seen with all murine *Sall* genes (Buck et al., 2001, Kohlhase et al., 2002a, Ott et al., 1996, Kohlhase et al., 2002b). *Sall4* as well as *Xenopus XISALLA* (Neff et al., 2005) is expressed transiently in limb bud mesenchyme. The *Sall4* limb expression pattern is interesting in the light of observations that human *SALL4* mutations also cause severe limb

malformations (Kohlhase et al., 2002b, Al-Baradie et al., 2002, Kohlhase et al., 2005). Therefore, *Sall4* could be an important transcriptional regulator which involved in limb patterning and specification together with other *Sall* genes.

4.2 *Sall4* is required for early heart development

In E7.5 embryos *Sall4* expression was found in the cardiac crescent and on the sections of the mesoderm, the region of future myocardial plate. The heart arises from cells originated from anterior lateral plate of mesoderm, the bilateral heart fields (Cohen-Gould et al., 1996). Cardiomyocytes first appear in the cardiac crescent (Fig. 4.1). The endocardial tubes fuse in the ventral midline to form the tubular heart with arterial and venous pole. At the arterial pole of the heart, the outflow tract connects with the aortic arch arteries in the pharyngeal region. Recently, a population of myocardial precursor cells in chick and mouse embryos has been identified in pharyngeal mesoderm anterior to the early heart tube. This anterior heart-forming field gives rise to myocardium of the outflow region or arterial pole of the heart (Kelly et al., 2002). In previous reports it was shown that *Xenopus xsal-3*, chicken *csal1* and *csal4*, and mouse *Sall1* and *Sall3* are expressed in the heart or in the structures necessary for heart development.

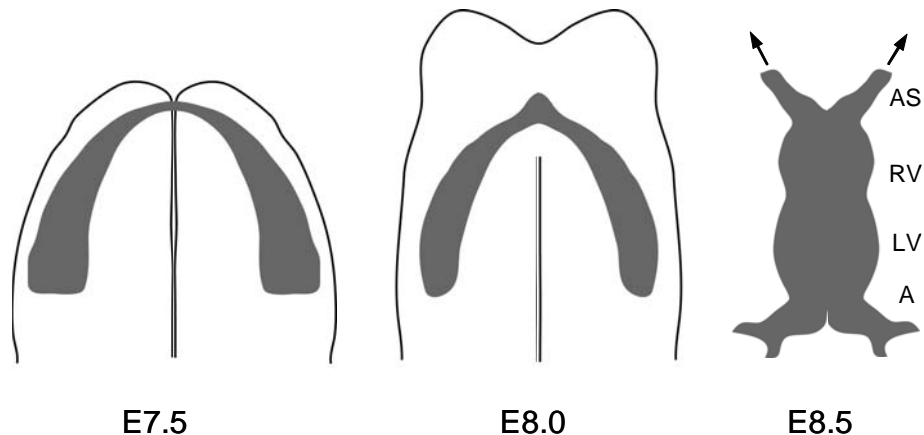


Figure 4.1: Stages of mouse cardiogenesis. The major heart forming field is shown in black. Bilateral mesodermal heart fields fuse in the ventral midline to form the tubular heart with arterial and venous pole. AS, atria sinus venosa; RV, right ventricle; LV, left ventricle; A, atria (modified from Zaffran et al., 2002).

Sall4 expression was notable in the neural crest cells migrating from the hindbrain of E8.5 stage embryo. During the formation of the neural tube, the ectodermal cells which are at the junction of the neural plate form neural crest cells. Neural crest-derived cells have stem cell characteristics because they have the capacity to proliferate and differentiate into various types of the cells, including sensory and sympathetic neurons, glial cells in the PNS, and smooth muscle cells in the blood vessels (Hall, 2000, Gammill and Bronner-Fraser, 2003, Sakai et al., 2005). In addition, there are contributions of neural crest cells to the wall of the aortic arches and the aorticopulmonary septum (Tomita et al., 2005). There is also a speculation that defects in neural crest cell migration might be involved in tetralogy of Fallot (Gittenberg-de Groot et al., 2005). Chicken *csal4* is involved in specification of avian cranial neural crest cell lineages (Barembaum and Bronner-Fraser, 2004). Our observation that *Sall4* is expressed in the branchial arches and migrating neural crest cells which play an essential role during early heart development is consistent with the finding that *SALL4* mutations can also be associated with severe heart defects (Borozdin et al., 2004a). Therefore, *Sall4* could be an important transcriptional regulator which plays an essential role during cardiogenesis. In conclusion, it remains to be proven how *Sall4* contributes to the development of the heart and in which signalling pathways it is involved.

4.3 Hypomorphic phenotype of *Sall4* gene trap mice

In the present study a mouse mutant for the *Sall4* gene is described which was generated by a gene trap approach. Based on the observation that all of the *SALL4* mutations in Okihiro syndrome mutations are predicted to cause *SALL4* haploinsufficiency (Borozdin et al., 2004b), the *Sall4* heterozygous mutant mice were expected to show at least some features similar to Okihiro syndrome patients which include defects of the upper limbs, heart defects, renal anomalies and a missing 6th cranial nerve or its nucleus. However, no phenotype was observed in the heterozygous mutants. They had normal appearance, and both male and female heterozygotes were fertile. No abnormalities of skeleton, brain, kidney, muscles and heart were found.

There are several ways to interpret this finding. The gene trap vector integration site lies between exon 1 and exon 1a, and quantitative RT-PCR showed presence of the minor transcript of approximately 73% in heterozygous embryos and 45% in *Sall4* homozygous

embryos as compared to wild type. The integration led to a reduction of the major transcript to 42% in heterozygous embryos and down to 2% in *Sall4* homozygous embryos. The expression level of exon 3-4 transcript which represents the total amount of transcript was reduced in heterozygous embryos to 50% and in homozygous embryos to 10%. This result indicates that the *Sall4* major transcript is largely disrupted by the gene trap vector integration. Despite of a strong splice acceptor sequence in the gene trap vector associated with the LacZ reporter and polyadenylation signal, the residual amount of transcript (2%) in homozygous embryos can be explained by a “bypass” of the gene trap vector. Bypass of gene trap vector which results in hypomorphic allele was discussed by Shawlot et al. (1998).

The *Sall4* minor transcript is expressed in homozygous embryos, but the level is strongly reduced. It can be assumed that the integration interferes with a possible promoter upstream of exon 1a or isolates exon 1a from a promoter upstream of exon 1. It remains unclear if presence of about 50% of the *Sall4* major and of 70% of *Sall4* minor transcripts is sufficient to avoid phenotypic consequences. Taking in consideration that expression level of exon 3-4 transcript is highly decreased in homozygous embryos the contribution of *Sall4* minor transcript variant in overall *Sall4* expression profile seems very low. The deletion of *Sall4* minor transcript could confirm the involvement of this transcript in phenotype of *Sall4* mice. However, it is unclear if *Sall4* haploinsufficiency would result in a phenotype at all.

Some interesting results about molecular mechanisms of the disease and mouse models were obtained from *Sall1* studies in human and mouse. It has been proposed that molecular mechanism of Townes-Brocks syndrome (TBS) is haploinsufficiency, because most of the *SALL1* mutations lead to the premature termination codons (Kohlhase, 2000a, Botzenhart et al., 2005). The recent study on *SALL1* deletions in human confirmed *SALL1* haploinsufficiency as one of the mechanism for TBS development (Borozdin et al., 2005). Nevertheless, mouse model for TBS with deletion of all zinc fingers, reported by Nishinakamura et al., failed to recapitulate TBS phenotype. Like *Sall4* heterozygous mice, *Sall1* heterozygotes did not display any abnormalities. In homozygous condition *Sall1* knockout mice revealed kidney agenesis or severe kidney dysgenesis, but not other changes in any organs (Nishinakamura et al., 2001). TBS causing mutations in *SALL1* could also lead to a truncated protein with a dominant negative activity of the mutated protein (Kiefer et al., 2003, Sweetman et al., 2003). Based on this hypothesis another transgenic mouse model which mimics the phenotype of TBS patients was described by Kiefer et al. (2003). Heterozygous

mice, expressing *Sall1* transcriptional repression domain and lacking the DNA binding domain recapitulate the abnormalities found in TBS patients (Kiefer et al., 2003). These studies demonstrated that haploinsufficiency of murine *Sall1* could not lead to the phenotype similar to TBS, while haploinsufficiency of human *SALL1* cause mild TBS phenotype.

Sall4, *Sall2* and *Sall3* heterozygous mutants are indistinguishable from wild type animals (Sato et al., 2003, Parrish et al., 2004). *Sall2* knockout mice did not reveal any abnormalities neither in homozygous nor in heterozygous animals suggesting a possible alternative mechanism of compensation by other *Sall* genes (Sato et al., 2003). *Sall3* homozygous mutant mice do not survive because of inability to feed due to palate and epiglottis anomalies (Parrish et al., 2004).

To summarise, mouse models for *Sall* genes with haploinsufficiency of all zinc fingers do not display any abnormalities in heterozygotes. Homozygous mutants, except *Sall2*, die shortly after birth (*Sall1* and *Sall3*) or during early embryonic development (*Sall4*). Molecular mechanisms described for human syndromes (TBS and Okihiro syndrome) most probably are not fully applicable for mouse models because of differences in functions in mouse and human.

4.4 *Sall4* homozygous mutant mice die during embryonic development

Genotyping analysis of the F2 generation of *Sall4* mouse line revealed that distribution of the genotypes in newborns was deviated from the Mendelian inheritance. Failure to recover mice homozygous for the *Sall4* mutation suggested that the gene is essential for embryonic development. No changes in cellular proliferation or developmental arrest in *Sall4* homozygous mutant embryos were found during the preimplantation period. By genotyping analysis we could demonstrate that homozygous embryos die around E7.5 (Table 3.3). At E3.5 stage *Sall4* homozygous embryos were indistinguishable from heterozygous or wild type littermates (Fig. 3.17). By E7.5, however, mutants could be easily distinguished from normal embryos. Mutant embryos lacked the development of primitive streak and extraembryonic cavities or membranes and exhibited morphology similar to E6.5 embryos (Fig. 3.14). At E8.5 normal embryos have completed gastrulation and start with organogenesis. In contrast, mutants were smaller and severely delayed as compared to their littermates (Fig. 3.13). At E8.5 stage embryos the most severe phenotypic abnormalities were in the heart region. The

heart is the first organ to form during early embryogenesis and its circulatory function is critical for the viability of the mammalian embryo.

Several signalling pathways are involved in the specification of the heart precursors including BMP, Wnt and FGF (Shi et al., 2000, Marvin et al., 2001, Lavine et al., 2005, Zaffran et al., 2002). Mouse models with severe heart defects such as knockout mice for *Nkx2-5*, *Gata4*, *Hand1*, or *Tbx5* (Lyons et al., 1995, Narita et al., 1997, Srivastava et al., 1997, Bruneau et al., 2001) die during early gestation period because of heart failure. In *Nkx2-5* mutant embryos, the heart tube formation occurs normally, but looping morphogenesis is not initiated at the linear heart tube stage (E8.25-E8.5) (Lyons et al., 1995). *Gata4* knockout mice die by E9.5 and exhibit profound defects in ventral morphogenesis, including abnormal foregut formation and a failure of fusion of the bilateral myocardial primordial fields (Narita et al., 1997). The embryonic lethality of *Sall4* homozygous mutant embryos during early gestation could therefore result from cardiac bifida together with the deformation of the foregut and missing neural tube closure. At the same time cardiogenic defects are unlikely to be a result of a general delay in the growth and development of *Sall4* mutant embryos. The embryonic death because of heart failure occurs in later stages, starting from E8.25. Another possibility would be that the cardiac defects could be secondary to insufficient nutrition because of abnormalities of the developing placenta or the ectoplacental cone. A support for this hypothesis could be that *Sall4* is expressed during blastocyst stage in the trophoblast which later gives rise to the embryonic part of the placenta and ectoplacental cone. A similar mouse model was described by Firulli et al. (1998). Homozygous *Hand1* mutant embryos die around E8.5 because of yolk sac abnormalities combined with heart defects. It might therefore well be that *Sall4* homozygous embryos are developmentally delayed and die around E7.5 of heart failure, foregut deformation and open neural tube. The possible contribution of an extraembryonic component to the *Sall4* phenotype remains to be investigated.

4.5 *Sall4* regulation during embryonic development

Previous analysis showed that the expression of *D. melanogaster sal* and *salr* genes in the wing imaginal disc is under direct regulation by the TGF- β molecule Dpp, ortholog of vertebrate BMP family (de Celis et al., 1996). During sensory organ development *sal/salr* expression is regulated by the signalling molecules *hedgheg* (*hh*), *wingless* (*wnt*) and

decapentaplegic (dpp) (de Celis et al., 1999). It was also found that *sal* is downstream of the *wingless* signal in the *Drosophila* tracheal system (Chihara and Hayashi, 2000). Medaka *sal* expression was found in all known *hh* signaling centres and appears to be a general target of *hh* in the medaka fish (Köster et al., 1997). Chicken *csall* expression in the developing limbs is dependent on Shh signals and also is regulated by molecules of FGF and Wnt signalling pathways (Farrel et al., 2000).

To study the mechanism through which *sall* genes carry out these functions several downstream targets for *sall* proteins were identified in *D. melanogaster*. *sal* inhibits expression of *iriduois* and *kni* in the developing veins (de Celis et al., 2000). *C. elegans* homologue *sem-4* represses the activity of Hox genes *mec-3* and *egl-5* through binding to the promoter regions of these genes (Toker et al., 2003). In *Sall1*-deficient mice remarkable downregulation of some kidney mesenchymal markers like BMP7, Wnt4 and Pax2 was observed (Nishinakamura et al., 2001). Murine *Sall1* interacts with HDAC components and strongly repress the transcription in the luciferase reporter assay (Kiefer et al., 2002). *Sall1* can also interact with UBE2I and undergoes covalent modification with SUMO-1 molecules (Netzer et al., 2002). *Sall1* can synergistically enhance reporter activity of the Wnt pathway (Sato et al., 2004). Human SALL1 repression function is mediated through the interaction with PIN2/TRF1 (Netzer et al., 2001). These experiments have demonstrated that SALL proteins are transcriptional repressors and most probably inhibit their downstream targets directly.

Identification of downstream targets of *Sall4* is important for an understanding of the phenotype of *Sall4* mutant mice. The expression pattern of several well known molecular markers which expressed during early embryogenesis was examined. At E8.5 *Pax3* is expressed in the neural tube and by pre-migratory neural crest cells (Goulding et al., 1991) and loss of *Pax3* leads to the spina bifida phenotype and displays cardiac outflow abnormalities and defective melanocyte development (Epstein et al., 1991). *Sall4* homozygous embryos did not express *Pax3* (Fig. 3.18). A possible explanation for the lack of *Pax3* expression is that mutant embryos in their development do not reach the stage when *Pax3* expression initiates. At the same time *Pax3* expression pattern and influence on the migration behaviour of cardiac neural crest cell (Epstein et al., 2000) fits with the expression pattern and heart phenotype of *Sall4* mice. *Sall4* phenotype is more severe than *Pax3* and occurs earlier. From this observation it can be speculated that *Pax3* expression might be controlled by *Sall4*.

The T-box gene (Tbx) family plays an essential role in controlling many aspects of embryogenesis. Mice homozygous for T-box gene *brachyury* die shortly after gastrulation because of complete loss of the posterior mesoderm (Wilkinson et al., 1990). Misregulation of *brachyury* expression in homozygous *Sall4* embryos can be explained by developmental delay and abnormal gastrulation (Fig. 3.19C, D). The cells of the cardiogenic mesoderm are brought to the cardiac crescent by gastrulation movements (Zaffran et al., 2002) and misregulation of this mechanism could lead to heart formation failure. This result indicates that axis of the development is disrupted possibly because of problems in mesoderm formation.

Several members of T-box family have been implicated in early cardiac lineage determination, chamber specification and valvuloseptal development (Chapman et al., 1996, Simon, 1999). Haploinsufficiency of human *TBX1* leads to the DiGeorge syndrome (DGS) which comprises hypocalcaemia, thymic hypoplasia, and outflow tract defects of the heart with tetralogy of Fallot, truncus arteriosus, and ventricular septal defects (Yagi et al., 2003). Disturbance of cervical neural crest cell migration into the derivatives of the pharyngeal arches can account for the phenotype (Kochilas et al., 2002). Loss of *Tbx1* in mice has been implicated in the etiology of cardiac outflow tract and pharyngeal arch anomalies (Merscher et al., 2001, Jerome et al., 2001). Mutations in human *TBX5* are responsible for Holt-Oram syndrome (HOS) with highly variable phenotypes including both upper limb and cardiac defects (Basson et al., 1999). There are overlapping clinical features between Okinhiro syndrome and HOS. Mutations of *SALL4* gene were reported in some HOS patients (Kohlhase et al., 2003). Absence of mouse *Tbx5* causes embryonic lethality (E10.5) and aberration of cardiac development (Bruneau et al., 2001, Zhou et al., 2005).

The expression profile of *brachyury*, *Tbx1* and *Tbx5* in *Sall4* mice was analysed by real-time RT-PCR. Strong reduction in the expression level of all three T-box genes was observed in homozygous embryos. Expression of *Tbx1* and *Tbx5* was downregulated to around 6% of the wild type level, while *brachyury* expression was reduced by 45%. Expression of these genes was almost unaffected in *Sall4* heterozygous embryos. This result indicates that T-box genes expression might be regulated by *Sall4* and gene dosage is important for this regulation. *Sall4* function might be mediated through some other genes which interact with T-box genes directly, for example *Gata4* or *Nkx2.5* (Hatcher et al., 2003, Bruneau, 2002). *Sall4* gene might be involved in some signalling pathways which play a role during neural crest cell migration and heart formation. Direct involvement of BMP, Wnt and FGF signaling in early

Discussion

cardiogenesis was previously shown in *Drosophila* and vertebrates (Delot et al, 2003, Barron et al., 2000, Marvin et al., 2001). For example, BMP signals control several cardiac regulatory genes, which include *Gata4* (Nemer et al., 2003) and *Hand* genes (Schlange et al., 2000). The regulation of *Sall4* seems to be a complex process in which both transcriptional activators and repressors interact with particularly regulatory regions to exert the signals during specification and patterning of the heart. To summarise, we can propose a general model for *Sall4* regulation during cardiac formation (Fig. 4.2). BMP, FGF and Wnt pathways trigger the interaction of several cardiac transcription factors like *Hand1*, *Hand2*, *Nkx2.5*, *Gata4*, *Tbx5*. *Sall4* might be involved in this cascade by regulating the activity of *Tbx1* and *Tbx5* genes. The essential role of *Sall4* in this pathway, potential interaction partners and target genes might be studied in future.

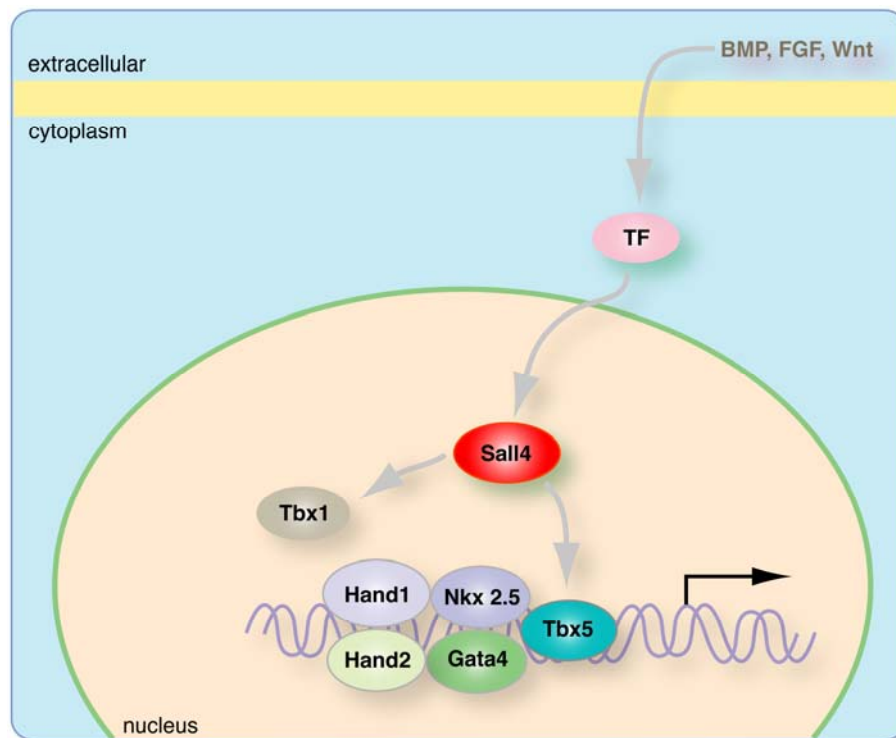


Figure 4.2: General pathway of cardiac transcription factors regulation. Signaling molecules of BMP, Wnt and FGF pathways mediate the cascade of the reactions which lead to the activation of the protein-protein interactions of cardiac transcription factors (TF). These complexes regulate transcriptional activity of the target genes. BMP, bone morphogenic protein, FGF, fibroblast growth factor, TF, transcription factor.

4.6 Future perspectives

New *Sall4* isoforms characterised in the present study address the function of these isoforms during embryonic development and in the adult tissues. By additional experiments the existence of an alternative promoter region for *Sall4* minor transcript has to be confirmed and which role it plays in the tissue specific and stage specific expression of *Sall4*.

Polyclonal anti-Sall4 antibody opens new possibilities in detecting novel protein-protein interactions as they can be used for the co-localisation and co-immunoprecipitation assays. The main purpose of these experiments is to find out new interaction partners and to expand the knowledge about the function of Sall4 and other Sall proteins *in vivo*.

In the present study a hypomorphic *Sall4* gene trap mouse line was characterised. Although the mutant *Sall4* mice do not show similar phenotype to Okihiro syndrome, the hypomorphic mouse model gives the opportunity to study the role of *Sall4* major transcript as transcriptional regulator which is involved in cardiac development. Overall developmental delay of *Sall4* homozygous embryos raises the question about possible contribution of an extraembryonic component to the *Sall4* phenotype which can be investigated by additional markers or tetraploid rescue.

To establish real mouse model for Okihiro syndrome some important points have to be considered. *Sall4* gene has several isoforms and to get *Sall4* loss-of-function all of them should be disrupted. New gene trap line with integrations in intron 2 and 3 or knockout strategy will be the necessary tool for this analysis. Another important point is that *Sall4* heterozygous mice are normal and phenotype of *Sall4* homozygous mice is different from Okihiro syndrome features. This observation may suggest the compensation effect by other *Sall* genes. Deletion of other *Sall* family members in combination could overcome this problem and help to elucidate their functions.

Of additional interest are the interaction studies of Sall4 with both Sall1 and truncated Sall1 proteins. Both *Sall4* and *Sall1* are co-expressed during embryonic development in the primitive streak, brain, limbs and branchial arches (Buck et al., 2001, Kohlhase et al., 2002b). From previous studies it is known that some of the *Sall* family members can interact *in vitro* with other family members in mouse and chicken and form homo- and heterodimeric complexes (Kiefer et al., 2003, Sweetman et al., 2003). The important interaction domain with dominant negative activity was mapped to the N-terminal part of the Sall1 (Kiefer et al.,

2003). It contains single zinc finger domain and the conserved glutamine-rich domain which is responsible for the interaction (Sweetman et al., 2003). Furthermore, it was found that truncated chicken csall1 and murine Sall1, similar to the truncated protein expressed in TBS patients, can interfere with normal function of full-length proteins by preventing the native proteins from localizing to the nucleus (Sweetman et al., 2003, Sato et al., 2004). Sall1 and truncated Sall1 proteins can bind specifically with all four Sall proteins, suggesting that this might be one of the Sall proteins function (Kiefer et al., 2003).

Nuclear localisation of Sall4 in F9 cells is predicted by other Sall family members and probably localised to pericentromeric heterochromatin (Netzer et al., 2001, Kiefer et al., 2002, Sato et al., 2004). Transfection experiments with myc-tagged Sall1 truncated construct revealed a clear displacement of Sall4 protein from the nucleus and cytoplasmic co-localisation with Sall1tr protein (Fig. 3.25). These preliminary studies suggest that dominant negative effect of Sall1tr can be also expanded to Sall4 protein. Another evidence of Sall4-Sall1 interaction was provided from the breeding experiments of heterozygous mice from Sall4 gene trap line and Sall1 knockout mouse line (Buck, 2002). The obvious deviation from the Mendelian inheritance ratio of Sall1/Sall4 offspring reflects an interaction between two genes (Table 3.4). The amount of Sall1/Sall4 double heterozygous mice was not significant and could be explained by genetic background effect. The absence of double heterozygous Sall1/Sall4 mice indicates the synergistic functional role and interaction between Sall1 and Sall4 gene. The Sall4-Sall1 interaction studies presented in this work could not give conclusive evidence about the mechanism of interaction and which domains or regulatory elements of Sall4 protein are responsible for maintaining specific protein-protein interactions. Co-localisation and co-immunoprecipitation experiments of Sall4 with full length Sall1 and truncated Sall1 will confirm the specificity and affinity of the interaction. It will be also of great interest to verify Sall4 protein interaction with other Sall family members. This will help to check the hypothesis of compensating the function of *Sall* genes by other family members. To confirm Sall4-Sall1 interaction on the genetic level it is necessary to find out when double heterozygous Sall1/Sall4 mice are dying and what is the underlying cause of the lethality.

There is also a clear need for further studies to confirm the reported involvement of Sall4 in the cardiac development regulation. Although the regulation of several important cardiac transcription factors by Sall4 was demonstrated, it remains to investigate additional upstream

Discussion

and downstream genes and targets which are responsible for early cardiac formation. *Sall4* mouse gene trap line and real-time PCR could be the powerful tool in determining such genes.

5. SUMMARY

In the present study the function of *Sall4* gene was investigated. Mutations in human homologue *SALL4* gene cause Okhiro syndrome, an autosomal dominant disorder characterised by radial ray abnormalities and Duane anomaly and some other anal, renal, cardiac, ear, foot malformations and hearing loss. All of the mutations in *SALL4* lead to preterminal stop codons and are thought to cause the phenotype via haploinsufficiency. Up to now no mouse model for *Sall4* gene was available. The main goal of this work was to characterize the mutation and the phenotype of the *Sall4* gene trap mouse line which can reflect the possible mechanisms of development of human syndrome.

Sall4 consists of 5 exons with ORF of 3203 bp. New splice variant *Sall4* minor which utilises alternative exon 1a instead of exon 1 was characterised. Expression for both transcripts was detectable as early as E7.5 and gradually decreases until E17.5. The contribution of the two transcript variants to the *Sall4* expression level is not equal. The presence of two different regulatory elements which control the expression of *Sall4* major and *Sall4* minor can be proposed.

By X-gal staining *Sall4* expression was detected in the blastocyst stage embryos in the inner cell mass and in the trophectoderm cells. During embryonic development *Sall4* is expressed in the nervous system, branchial arches, limb buds and tail tip. In E7.5 embryos *Sall4* expression was found in the cardiac crescent, the region of future myocardial plate. *Sall4* expression was also notable in the migrating neural crest cells which are involved in generation of outflow tract of the heart. Therefore, *Sall4* could be an important transcriptional regulator which plays an essential role during cardiogenesis.

In order to establish an appropriate mouse model system for Okhiro syndrome and to study the function of *Sall4* gene the gene trap ES cell line with pT1 β geo gene trap vector integration between exon 1 and exon 1a was used to generate *Sall4* mouse line. Quantitative RT-PCR showed that the *Sall4* major transcript was largely disrupted by the gene trap vector integration while the *Sall4* minor transcript is expressed in homozygous embryos, but the level is strongly reduced. No apparent phenotype was observed in the heterozygous mutants. *Sall4* homozygous embryos were developmentally delayed and die around E7.5 because of heart failure, foregut deformation and open neural tube.

Summary

Some downstream targets of *Sall4* were used to describe the phenotype of *Sall4* mutant mice. *Sall4* homozygous embryos did not express *Pax3*, an early marker for neural epithelium, indicating that *Pax3* expression might be controlled by *Sall4*. Misregulation of *brachyury* expression in homozygous *Sall4* embryos can be explained by developmental delay and problems in gastrulation. The expression profile of T-box genes *brachyury*, *Tbx1* and *Tbx5* in *Sall4* mice was analysed by real-time RT-PCR. Strong reduction of the expression level of all three T-box genes was observed in homozygous embryos; in contrast, the expression of these genes was almost unaffected in *Sall4* heterozygous embryos. This result indicates that expression of T-box genes which are involved in early cardiogenesis might be regulated by *Sall4* and gene dosage is important for this regulation.

Sall4 can interact with other *Sall* family members. Transfection experiments with myc-tagged *Sall1* truncated construct revealed a clear displacement of *Sall4* protein from the nucleus and cytoplasmic co-localisation with truncated *Sall1* protein. Another evidence of *Sall4*-*Sall1* interaction was provided by breeding experiments of heterozygous mice from *Sall4* gene trap line and *Sall1* knockout mouse line. The severe reduction of double heterozygous *Sall1*/*Sall4* mice indicates the synergistic functional role and interaction between *Sall1* and *Sall4* gene on the genetic level.

6. REFERENCES

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., Okihira M.M., Engle E.C. (2002). Duane radial ray syndrome (Okihira syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. *Am J Hum Genet* **71**: 1195-9.

Austin C.P., Battey J.F., Bradley A., Bucan M., Capecchi M., Collins F.S., Dove W.F., Duyk G., Dymecki S., Eppig J.T., Grieder F.B., Heintz N., Hicks G., Insel T.R., Joyner A., Koller B.H., Lloyd K.C., Magnuson T., Moore M.W., Nagy A., Pollock J.D., Roses A.D., Sands A.T., Seed B., Skarnes W.C., Snoddy J., Soriano P., Stewart D.J., Stewart F., Stillman B., Varmus H., Varticovski L., Verma I.M., Vogt T.F., von Melchner H., Witkowski J., Woychik R.P., Wurst W., Yancopoulos G.D., Young S.G., Zambrowicz B. (2004). The knockout mouse project. *Nat Genet* **36**: 921-4.

Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., and Struhl K. (1994). *Current Protocols in Molecular Biology*, (John Wiley & Sons Inc., USA).

Barenbaum M. and Bronner-Fraser M. (2004). A novel spalt gene expressed in branchial arches affects the ability of cranial neural crest cells to populate sensory ganglia. *Neuron Glia Biology* **1**: 57-63.

Barron M., Gao M., Lough J. (2000). Requirement for BMP and FGF signaling during cardiogenic induction in non-precardiac mesoderm is specific, transient, and cooperative. *Dev Dyn* **218**: 383-93.

Basson C. T., Huang T., Lin R. C., Bachinsky D. R., Weremowicz S., Vaglio A., Bruzzone R., Quadrelli R., Lerone M., Romeo G., Silengo M., Pereira A., Krieger J., Mesquita S. F., Kamisago M., Morton C. C., Pierpont M. E. M., Muller C. W., Seidman J. G., Seidman C. E. (1999). Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proc Nat Acad Sci* **96**: 2919-2924.

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

Borozdin W., Steinmann K., Albrecht B., Bottani A., Devriendt K., Leipoldt M., Kohlhase J. (2005). Detection of heterozygous SALL1 deletions by quantitative real time PRC proves the contribution of a SALL1 dosage effect in the pathogenesis of Townes-Brocks syndrome. *Hum Mutat* (in press).

Borozdin W., Wright M.J., Hennekam R.C., Hannibal M.C., Crow Y.J., Neumann T.E., Kohlhase J. (2004a). Novel mutations in the gene SALL4 provide further evidence for acro-renal-ocular and Okihira syndromes being allelic entities, and extend the phenotypic spectrum. *J Med Genet* **41**: e102.

References

- Botzenhart E.M., Green A., Ilyina H., Konig R., Lowry R.B., Lo I.F., Shohat M., Burke L., McGaughran J., Chafai R., Pierquin G., Michaelis R.C., Whiteford M.L., Simola K.O., Rosler B., Kohlhase J. (2005). SALL1 mutation analysis in Townes-Brocks syndrome: twelve novel mutations and expansion of the phenotype. *Hum Mutat* **26**: 282.
- Bruneau B.G. (2002). Transcriptional regulation of vertebrate cardiac morphogenesis. *Circ Res* **90**: 509-19.
- Bruneau B.G., Nemer G., Schmitt J.P., Charron F., Robitaille L., Caron S., Conner D.A., Gessler M., Nemer M., Seidman C.E., Seidman J.G. (2001). A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell* **106**: 709-21.
- Buck A. (2002). Studies on the *sal*-like genes of the mice. PhD thesis.
- Buck A., Archangelo L., Dixkens C., Kohlhase J. (2000). Molecular cloning, chromosomal localization, and expression of the murine SALL1 ortholog Sall1. *Cytogenet Cell Genet* **89**: 150-153.
- Buck A., Kispert A., Kohlhase J. (2001). Embryonic expression of the murine homologue of SALL1, the gene mutated in Townes--Brocks syndrome. *Mech Dev* **104**: 143-6.
- Camp E., Hope R., Kortschak R.D., Cox T.C., Lardelli M. (2003). Expression of three spalt (*sal*) gene homologues in zebrafish embryos. *Dev Genes Evol.* **213**: 35-43.
- Chapman D.L., Garvey N., Hancock S., Alexiou M., Agulnik S.I., Gibson-Brown J.J., Cebra-Thomas J., Bollag R.J., Silver L.M., Papaioannou V.E. (1996). Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. *Dev Dyn* **206**: 379-90.
- Chien A., Edgar D.B., Trela J.M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol* **127**: 1550-57.
- Chihara T., Hayashi S. (2000). Control of tracheal tubulogenesis by Wingless signaling. *Development* **127**: 4433-42.
- Chomczynski P., Sacchi N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem* **162**: 156-159.
- Clark J.M. (1988). Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res* **16**: 9677-86.
- Cohen-Gould L., Mikawa T. (1996). The fate diversity of mesodermal cells within the heart field during chicken early embryogenesis. *Dev Biol* **177**: 265-73.
- de Celis J.F. (1998). Positioning and differentiation of veins in the *Drosophila* wing. *Int J Dev Biol* **42**: 335-43.

References

de Celis J.F., 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. und 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.

de Celis J.F., Barrio R., Kafatos F.C. (1996). A gene complex acting downstream of *dpp* in *Drosophila* wing morphogenesis. *Nature* **381**: 421-4.

Delot E.C., Bahamonde M.E., Zhao M., Lyons K.M. (2003). BMP signaling is required for septation of the outflow tract of the mammalian heart. *Development* **130**: 209-20.

Domingos P.M., Mlodzik M., Mendes C.S., Brown S., Steller H., Mollereau B. (2004). Spalt transcription factors are required for R3/R4 specification and establishment of planar cell polarity in the *Drosophila* eye. *Development* **131**: 5695-702.

Epstein D.J., Vekemans M., Gros P. (1991). Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* **67**: 767-74.

Epstein J.A., Li J., Lang D., Chen F., Brown C.B., Jin F., Lu M.M., Thomas M., Liu E., Wessels A., Lo C.W. (2000). Migration of cardiac neural crest cells in Splotch embryos. *Development* **127**: 1869-78.

Farrell E.R. and Munsterberg A.E. (2000). *csal1* is controlled by a combination of FGF and Wnt signals in developing limb buds. *Dev Biol* **225**: 447–458.

Farrell E.R., Tosh G., Church E. und Munsterberg A.E. (2001). Cloning and expression of CSAL2, a new member of the spalt gene family in chick. *Mech Dev* **102**: 227–230.

Firulli A.B., McFadden D.G., Lin Q., Srivastava D., Olson E.N. (1998). Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. *Nat Genet* **18**: 266-70.

Freceer V., Miertus J., Borozdin W., Kohlhase J., Amoroso A., Miertus S. (2004). Molecular modeling of *c2h2* zinc finger mutation of putative human transcription factor SALL4. *Internet Electron J Mol Des* **3**: 295–307.

Gammill L.S., Bronner-Fraser M. (2003). Neural crest specification: migrating into genomics. *Nat. Rev. Neurosci.* **4**: 795–805.

Gershoni J.M., Palade G.E. (1982). Electrophoretic transfer of proteins from sodium sulfate-polyacrylamide gels to a positively charged membrane filter. *Anal Biochem* **124**: 396-405.

Gittenberger-de Groot A.C., Bartelings M.M., Deruiter M.C., Poelmann R.E. (2005). Basics of cardiac development for the understanding of congenital heart malformations. *Pediatr Res* **57**: 169-76.

References

- Goulding M.D., Chalepakis G., Deutsch U., Erselius J.R., Gruss P. (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J* **10**:1135-47.
- Hall B.K. (2000). The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evol. Dev.* **2**: 3–5.
- Hansen J., Floss T., Van Sloun P., Fuchtbauer E.M., Vauti F., Arnold H.H., Schnutgen F., Wurst W., von Melchner H., Ruiz P. (2003). A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *Proc Natl Acad Sci USA* **100**: 9918-22.
- Hasuwa H., Kaseda K., Einarsdottir T., Okabe M. (2002). Small interfering RNA and gene silencing in transgenic mice and rats. *FEBS Lett* **532**: 227-30.
- Hatcher C.J., Diman N.Y., McDermott D.A., Basson C.T. (2003). Transcription factor cascades in congenital heart malformation. *Trends Mol Med* **12**: 512-5.
- Holleman T., Schuh R., Pieler T. und Stick R. (1996). Xenopus Xsal-1, a vertebrate homolog of the region specific homeotic gene spalt of Drosophila. *Mech Dev* **55**: 19–32.
- Hu G. (1993). DNA polymerase-catalyzed addition of non-templated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol* **12**: 763–770.
- Jerome L.A., Papaioannou V.E (2001). DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nature Genet* **27**: 286-291.
- Jürgens G. (1988). Head and tail development of the Drosophila embryo involves spalt, a novel homeotic gene. *EMBO J* **7**: 189–196.
- Kelly R.G., Buckingham M.E. (2002). The anterior heart-forming field: voyage to the arterial pole of the heart. *Trends Genet* **18**: 210-6.
- Kiefer S.M., McDill B.W., Yang J., Rauchman M. (2002). Murine Sall1 represses transcription by recruiting a histone deacetylase complex. *J Biol Chem* **277**: 14869-76.
- Kiefer S.M., Ohlemiller K.K., Yang J., McDill B.W., Kohlhase J., Rauchman M. (2003). Expression of a truncated Sall1 transcriptional repressor is responsible for Townes-Brocks syndrome birth defects. *Hum Mol Genet* **12**: 2221-7
- Kochilas L., Merscher-Gomez S., Lu M.M., Potluri V., Liao J., Kucherlapati R., Morrow B., Epstein J.A. (2002). The role of neural crest during cardiac development in a mouse model of DiGeorge syndrome. *Dev Biol* **251**: 157-66.
- Kohlhase J. (2002c). Personal communication.
- Kohlhase J, Wischermann A, Reichenbach H, Froster U, Engel W. (1998). Mutations in the *SALL1* putative transcription factor gene cause Townes-Brocks syndrome. *Nat Genet* **18**: 81-83.

References

Kohlhase J. (2000). SALL1 mutations in Townes-Brocks syndrome and related disorders. *Hum Mutat* **16**: 460-6.

Kohlhase J., Altmann M., Archangelo L., Dixkens C., Engel W. (2000). Genomic cloning, chromosomal mapping, and expression analysis of msal-2. *Mamm Genome* **11**: 64-8.

Kohlhase J., Chitayat D., Kotzot D., Ceylaner S., Froster U.G., Fuchs S., Montgomery T., Rosler B. (2005). SALL4 mutations in Okihiro syndrome (Duane-radial ray syndrome), acrorenal-ocular syndrome, and related disorders. *Hum Mutat* **26**: 176-83.

Kohlhase J., Hausmann S., Stojmenovic G., Dixkens C., Bink K., Schulz-Schaeffer W., Altmann M., Engel W. (1999b). SALL3, a new member of the human spalt-like gene family, maps to 18q23. *Genomics* **62**: 216-22.

Kohlhase J., Heinrich M., Liebers M., Frohlich Archangelo L., Reardon W., Kispert A. (2002b). Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. *Cytogenet Genome Res* **98**: 274-7.

Kohlhase J., Heinrich M., Schubert L., Liebers M., Kispert A., Laccone F., Turnpenny P., Winter R.M., Reardon W. (2002a). Okihiro syndrome is caused by SALL4 mutations. *Hum Mol Genet* **11**: 2979-87.

Kohlhase J., Schubert L., Liebers M., Rauch A., Becker K., Mohammed S.N., Newbury-Ecob R., 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, acrorenal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. *J Med Genet* **40**: 473-8.

Kohlhase J., Schuh R., Dowe G., Kühnlein R.P., Jackle H., Schroeder B., Schulz-Schaeffer W., Kretzschmar H.A., Kohler A., Muller U., Raab-Vetter M., Burkhardt E., Engel W., Stick R. (1996). Isolation, characterization, and organ-specific expression of two novel human zinc finger genes related to the Drosophila gene spalt. *Genomics* **38**: 291-8.

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., Seidel J., Kirkpatrick S.J., Pauli R.M., Wargowski D.S., Devriendt K., Proesmans W., Gabrielli O., Coppa G.V., Wesby-van Swaay E., Trembath R.C., Schinzel A.A., Reardon W., Seemanova E., Engel W. (1999a). Molecular analysis of SALL1 mutations in Townes-Brocks syndrome. *Am J Hum Genet* **64**: 435-45.

Köster R., Stick R., Loosli F. und Wittbrodt J. (1997). Medaka spalt acts as a target gene of hedgehog signalling. *Development* **124**: 3147-3156.

Kühnlein R.P., Frommer G., Friedrich M., Gonzalez-Gaitan M., Weber A., Wagner-Bernholz J.F., Gehring W., Jackle H. und Schuh R. (1994). spalt encodes an evolutionary 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.

References

- Kühnlein R.P., Schuh R. (1996). Dual function of the region-specific homeotic gene spalt during *Drosophila* tracheal system development. *Development* **122**: 2215-23.
- Kunath T., Gish G., Lickert H., Jones N., Pawson T., Rossant J. (2003). Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. *Nat Biotechnol* **21**: 559-61.
- Laird P.W., Zijderveld A., Linders K., Rudnicki M.A., Jaenisch R., Berns A. (1991). Simplified mammalian DNA isolation procedure. *Nucleic Acids Res* **19**: 4293.
- Lavine K.J., Yu K., White A.C., Zhang X., Smith C., Partanen J., Ornitz D.M. (2005). Endocardial and epicardial derived FGF signals regulate myocardial proliferation and differentiation in vivo. *Dev Cell* **8**: 85-95.
- Li D., Dower K., Ma Y., Tian Y., Benjamin T.L. (2001). A tumor host range selection procedure identifies p150(sal2) as a target of polyoma virus large T antigen. *Proc Natl Acad Sci USA* **98**: 14619-24.
- Lyons I., Parsons L.M., Hartley L., Li R., Andrews J.E., Robb L., Harvey R.P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev* **9**: 1654-66.
- Ma Y., Li D., Chai L., Luciani A.M., Ford D., Morgan J., Maizel A.L. (2001). Cloning and characterization of two promoters for the human HSAL2 gene and their transcriptional repression by the Wilms tumor suppressor gene product. *J Biol Chem* **276**: 48223-30.
- Marvin M.J., Di Rocco G., Gardiner A., Bush S.M., Lassar A.B. (2001). Inhibition of Wnt activity induces heart formation from posterior mesoderm. *Genes Dev* **15**: 316-327.
- Merscher S., Funke B., Epstein J.A., Heyer J., Puech A., Lu M.M., Xavier R.J., Demay M.B., Russell R.G., Factor S., Tokooya K., Jore B.S., Lopez M., Pandita R.K., Lia M., Carrion D., Xu H., Schorle H., Kobler J.B., Scambler P., Wynshaw-Boris A., Skoultchi A.I., Morrow B.E., Kucherlapati R. (2001). TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* **104**: 619-29.
- Miller J., McLachlan A.D. und Klug A. (1985). Repetitive zinc-binding domains in the protein transcription factor TFIIIA from *Xenopus* oocytes. *EMBO J* **4**: 1609-1614.
- Narita N., Bielinska M., Wilson D.B. (1997). Wild-type endoderm abrogates the ventral developmental defects associated with GATA-4 deficiency in the mouse. *Dev Biol* **189**: 270-4.
- Neff A.W., King M.W., Harty M.W., Nguyen T., Calley J., Smith R.C., Mescher A.L. (2005). Expression of *Xenopus* XISALL4 during limb development and regeneration. *Dev Dyn* **233**: 356-67.
- Nemer G., Nemer M. (2003). Transcriptional activation of BMP-4 and regulation of mammalian organogenesis by GATA-4 and -6. *Dev Biol* **254**: 131-48.

References

- Netzer C., Rieger L., Brero A., Zhang C-D., Hinzke M., Kohlhase J, 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.
- Nishinakamura R., Matsumoto Y., Nakao K., Nakamura K., Sato A., Copeland N.G., Gilbert D.J., Jenkins N.A., Scully S., Lacey D.L., Katsuki M., Asashima M., Yokota T. (2001). Murine homolog of SALL1 is essential for ureteric bud invasion in kidney development. *Development* **128**: 3105-15.
- Ohgane J., Wakayama T., Senda S., Yamazaki Y., Inoue K., Ogura A., Marh J., Tanaka S., Yanagimachi R., Shiota K. (2004). The Sall3 locus is an epigenetic hotspot of aberrant DNA methylation associated with placentomegaly of cloned mice. *Genes Cells* **9**: 253-60.
- Onai T., Sasai N., Matsui M., Sasai Y. (2004). Xenopus XsalF: anterior neuroectodermal specification b attenuating cellular responsiveness to Wnt signaling. *Dev Cell* **7**: 95-106.
- Onuma Y., Nishinakamura R., Takahashi S., Yokota T. und Asashima M. (1999). Molecular cloning of a novel Xenopus spalt gene (Xsal-3). *Biochem Biophys Res Commun* **264**: 151–156.
- Ott T., Kaestner K.H., Monaghan A.P. and Schütz G. (1996). The mouse homolog of the region specific homeotic gene spalt of Drosophila is expressed in the developing nervous system and in mesoderm-derived structures. *Mech Dev* **56**: 117–128.
- Ott T., Parrish M., Bond K., Schwaeger-Nickolenko A., 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**:203-7.
- Parrish M., Ott T., Lance-Jones C., Schuetz G., Schwaeger-Nickolenko A., Monaghan A.P. (2004). Loss of the Sall3 gene leads to palate deficiency, abnormalities in cranial nerves, and perinatal lethality. *Mol Cell Biol* **24**:7102-12.
- Powell C.M., Michaelis R.C. (1999). Townes-Brocks syndrome. *J Med Genet* **36**: 89-93.
- Prawitt D., Brixel L., Spangenberg C., Eshkind L., Heck R., Oesch F., Zabel B., Bockamp E. (2004). RNAi knock-down mice: an emerging technology for post-genomic functional genetics. *Cytogenet Genome Res* **105**:412-21.
- Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., Erlich H.A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-91.
- Sakai D, Wakamatsu Y. (2005). Regulatory mechanisms for neural crest formation. *Cells Tissues Organs* **179**: 24-35.
- Sambrook J., Fritsch E. F., Maniatis T. Molecular cloning: a laboratory manual (2nd edition). (1989) Cold Spring Harbour, New York, USA.

References

- Sanger F., Nicklen S., Coulson A.R. (1977). DNA sequencing with the chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463-67.
- Sato A., Kishida S., Tanaka T., Kikuchi A., Kodama T., Asashima M., 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**:103-13.
- Sato A., Matsumoto Y., Koide U., Kataoka Y., Yoshida N., Yokota T., Asashima M., Nishinakamura R. (2003). Zinc finger protein sall2 is not essential for embryonic and kidney development. *Mol Cell Biol* **23**:62-9.
- Schlange T., Andree B., Arnold H.H., Brand T. (2000). BMP2 is required for early heart development during a distinct time period. *Mech Dev* **91**:259-70.
- Shawlot W., Deng J.M., Fohn L.E., Behringer R.R. (1998). Restricted beta-galactosidase expression of a hygromycin-lacZ gene targeted to the beta-actin locus and embryonic lethality of beta-actin mutant mice. *Transgenic Res* **7**:95-103.
- Shi Y., Katsev S., Cai C., Evans S. (2000). BMP signaling is required for heart formation in vertebrates. *Dev Biol* **224**: 226–237.
- Simon H. (1999). T-box genes and the formation of vertebrate forelimb- and hindlimb specific pattern. *Cell Tissue Res* **296**:57-66.
- Skarnes W.C., von Melchner H., Wurst W., Hicks G., Nord A.S., Cox T., Young S.G., Ruiz P., Soriano P., Tessier-Lavigne M., Conklin B.R., Stanford W.L., Rossant J. (2004): A public gene trap resource for mouse functional genomics. *Nat Genet* **36**: 543-4.
- Smithies O., Gregg R.G., Boggs S.S., Koralewski M.A., Kucherlapati R.S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**: 230-4.
- Srivastava D., Thomas T., Lin Q., Kirby M.L., Brown D., Olson E.N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nat Genet* **16**: 154-60.
- Stanford W.L., Cohn J.B., Cordes S.P. (2001). Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* **2**: 756-68.
- Sweetman D., Smith T., Farrell E.R., Chantry A., Munsterberg A. (2003). The conserved glutamine-rich region of chick csall and csal3 mediates protein interactions with other spalt family members. Implications for Townes-Brocks syndrome. *J Biol Chem* **278**:6560-6.
- Sweetman D., Smith T.G., Farrell E.R., Munsterberg A. (2005). Expression of csall in pre limb-bud chick embryos. *Int J Dev Biol* **49**: 427-30.

References

Toker A.S., Teng Y., Ferreira H.B., Emmons S.W., 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**:3831-40.

Tomita Y., Matsumura K., Wakamatsu Y., Matsuzaki Y., Shibuya I., Kawaguchi H., Ieda M., Kanakubo S., Shimazaki T., Ogawa S., Osumi N., Okano H., Fukuda K. (2005). Cardiac neural crest cells contribute to the dormant multipotent stem cell in the mammalian heart. *J Cell Biol* **170**: 1135-46.

Wilkinson D.G. (1992). *In situ hybridisation: A Practical Approach*. Oxford, IRL Press

Wilkinson D.G., Bhatt S., Herrmann B.G. (1990). Expression pattern of the mouse *T* gene and its role in mesoderm formation. *Nature* **343**:657-9.

Wurst W., Rossant J., Prideaux V., Kownacka M., Joyner A., Hill D.P., Guillemot F, Gasca S., Cado D., Auerbach A., et al. (1995). A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. *Genetics* **139**: 889-99.

Yagi H., Furutani Y., Hamada H., Sasaki T., Asakawa S., Minoshima S., Ichida F., Joo K., Kimura M., Imamura S., Kamatani N., Momma K., Takao A., Nakazawa M., Shimizu N., Matsuoka R. (2003). Role of *TBX1* in human del22q11.2 syndrome. *Lancet* **362**: 1366-1373.

Zaffran S., Frasch M. (2002). Early signals in cardiac development. *Circ Res* **91**: 457-69.

Zhou Y.Q., Zhu Y., Bishop J., Davidson L., Henkelman R.M., Bruneau B.G., Foster F.S. (2005). Abnormal cardiac inflow patterns during postnatal development in a mouse model of Holt-Oram syndrome. *Am J Physiol Heart Circ Physiol* **289**: H992-H1001.

ACKNOWLEDGMENTS

First of all, I would like to thank Prof. Dr. W. Engel for giving me the opportunity to work in the Institute of Human Genetics, for providing excellent experimental facilities, permanent scientific supervision and critical review of this work.

I wish to express my appreciation to my supervisor PD Dr. Jürgen Kohlhase for inviting me in his project, giving me an introduction into exciting field of mouse genetics and proofreading of the manuscript.

I want to warmly thank my *in situ* supervisor Dr. Ashraf ul-Mannan for his leadership, discussion of the results, providing me criticism and advices in the right proportions and for giving me the opportunity to work independently.

I sincerely thank PD Dr. S. Hoyer-Fender to be my co-referee and Prof. Dr. R. Heinrich and Prof. Dr. R. Ficner for accepting to be my examiners.

I am also deeply grateful to Prof. Dr. T. Pieler for accepting me to the graduate program “Molekulare Genetik der Entwicklung” and generous support for the scientific work and taking part in the conferences.

I wish to thank all my institute colleagues for the advices, support and patience. Particulary, I want to warmly thank my lab members Lucia, Wiktor, Anne, Agusia, Pei, Anitha and Rashid for their friendship and amazing multicultural atmosphere.

I sincerely thank to Prof. Dr. W. Wurst and Dr. T. Floss from the GSF Reseach Center (Munich) for the help in generation of Sall4 gene trap mice, teaching me X-gal staining and value discussions. I want to acknowledge Prof. Dr. M. Kessel for providing me with brachyury clone and Prof. Dr. B. Crist and Dr. R. Huang for introducing me into embedding techniques, BrdU staining and *in situ* hybridisation. I am indebted to Prof. Dr. T. Brandt and Alexander Froese for the help in interpretation of the results.

I also want to thank all my belarussian friends, especially Natalia Bogdanova, Lena Smirnova, Andrey Gurachevski, Katia Torbashevich, Alena Liavonchanka, Andrey Matsulka, Yulia Butscheid for the special belarussian spirit.

Most of all I would like to thank my husband Konstantin Glebov. I hope once I finish my PhD thesis I can repay you support, guidance, love and affection in some way.

Acknowledgments

Finally, I want to express my deepest gratitude to my parents Alena and Iosif Malinouski for their constant help, encouragement and support at hard moments, faith and love. Also I would like to thank my brother Mikalai who cheered me from afar.

CURRICULUM VITAE

Personal data

Name: Lina Malinouskaya

Date of birth: 17th December 1977

Place of Birth: Minsk, Belarus

Nationality: belarussian

Educational background

1985-1995 Secondary school Nr. 53, Minsk, Belarus. Awarded with the silver medal

1995-2000 International Sakharov Environmental University, Faculty of Radiobiology, Minsk, Belarus. Diploma in Radiobiology and Radiation Medicine

2000-2002 Experimental work as a student assistant in the Laboratory of Molecular Markers of Ecological Effects, Faculty of Radiobiology, International Sakharov Environmental University, Minsk, Belarus

2002-2005 PhD studies at the George-August University Göttingen, Institute of Human Genetics, under the supervision of Prof. Dr. W. Engel.

Title of thesis: Functional analysis of murine *Sall4*