Molecular and functional analysis of two gene trap mouse lines

Dissertation

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

vorgelegt von

Batjargal Gundsambuu

aus Ulaanbaatar, Mongolei

Göttingen 2004

D7
Referent: **Prof. Dr. W. Engel**
Korreferentin: **PD Dr. S. Hoyer-Fender**
Tag der mündlichen Prüfungen:

INDEX

ABBREVIATIONS

a.a.	amino acids
AAP	Abridged anchor primer
ABI	Applied Biosystem Instrument
APS	Ammonium peroxodisulfate
ATG	Translation intiation codon
ATP	Adenosintriphosphate
AUAP	Abridged universal amplification primer
βa	β-actin promoter
BAC	Bacterial Artificial Chromosome
BCP	1-bromo-3-chloropropane
BFA	Brefeldin A
β-gal	β-galactosidase
β-geo	a fusion gene between the β-galactosidase (β-gal) and neomycinphosphotransphearse (neo) genes
β-ME	β-Mercaptoethanol
bp	base pair
BSA	Bovine serum albumin
°C	Degree Celsius
CBF	Ciliary beat frequency
cDNA	complementary DNA
CIP	Calf intestine phosphatase
Cy3	indocarbocyanine
dATP	Desoxyriboadenosintriphosphate
dH2O	distilled Water
DAPI	Diamidino-2-phenylindole dihydrochloride
dCTP	Desoxyribocytosintriphosphate
DEPC	Diethylpyrocarbonate
DMF	Dimethylformamide

DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotidetriphosphate
dpc	day post coitum
dsRNA	double-stranded RNA
dT	deoxythymidinate
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
EGFP	Enhanced green fluorescence protein
EGL	External granular layer
En2	Engrailed 2
ES	Embryonic stem
EST	Expressed sequenced tags
F	Filial generation
FCS	Fetal calf serum
FISH	Fluorescence in Situ Hybridisation
FITC	Fluorescence isothiocyanate
g	gravity
GFP	Green fluorescence protein
GITC	Guanidine-Isothiocyanate
gm	gram
GT	Gene trap
hEF	Human Elongation Factor
HEPES	N-(-hydroxymethyl) piperazin,N'-3-propansulfoneacid
HPLC	High performance liquid chromatograpy
hr(s)	hour(s)
IGL	Internal granular layer
IFN	Interferon
IPTG	Isopropyl-ß-thiogalactopyranoside

IRES	Internal ribosomal entry sites
IVF	In vitro fertilisation
JL	Jackson Laboratory
kb	kilobase
LB	Luria-Bertrani
LIF	Recombinant leukaemia inhibitory factor
LPS	lipopolysaccharides
M	molarity
Mbp	Mega base pair
MCS	Multicloning site
MEF	Mouse embryonic fibroblast
ML	Molecular layer
MoCo	Molybdenum cofactor
MOCS	Molybdenum cofactor synthesis step
MOPS	3-[N-Morpholino]-Propanesulfate
mRNA	messenger Ribonucleic acid
mg	milligram
µg	microgram
ml	milliliter
µl	microliter
µm	micrometer
min	minute
N	normal
NaAc	Sodium acetate
NBT	Nitro-blue tetrazolium
NCBI	National Center for Biotechnology Information
Neo	Neomycin
ng	nanogram
NLS	Nuclear localisation sequence
nm	nanometer

NTP	Nucleotidetriphospate
OD	Optimal density
ORF	Open Reading Frame
pA	polyadenylation signal
PAC	Bacteriophage P1 Artificial Chromosome
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
PGK-neo	Phosphoglycerate kinase/bacterial neomycin resistance gene
PGK-tk	Phosphoglycerate kinase/viral thymidine resistance gene
pH	Preponderance of hydrogen ions
pmol	picomol
PBS	Phosphatebuffersaline
PBT	Phosphatebuffersaline + Tween 20
PCD	Primary ciliary dyskinesia
PMSF	Phenylmethylsulfonyl fluoride
RACE	Rapid Amplification of cDNA Ends
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
Rnase	Ribonuclease
rpm	revolution per minute
RT	Room temperature
RT-PCR	Reverse transcriptase-PCR
SA	Splice acceptor
SD	Splice donor
SDS	Sodium Dodecylsulfate
SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis
sec	second
siRNA	small interfering RNA
Sox	Sry box

SV 40	Simian Virus 40
Taq	*Thermus aquaticus*
TA	Thymidine-adenine
TBE	Tris-Borate-EDTA-Electrophoresis buffer
TE	Tris-EDTA buffer
TEMED	Tetramethylethylene diamine
Tris	Trihydroxymethylaminomethane
U	Unit
UAP	Universal amplification primer
UTR	Untranslated region
UV	Ultra violet
V	Voltage
v/v	volume/volume
w/v	weight/volume
WT	wild type
X-Gal	5-bromo-4-chloro-3-indolyl-ß-galactosidase
ZP	zona pellucida

Symbol of amino acids

A	Ala	Alanine
B	Asx	Asparagine or Asparatic acid
C	Cys	Cystine
D	Asp	Asparatic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine

M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
Z	Glx	Glutamine or Glutamic acid

Symbols of nucleic acid

A	Adenosine
C	Cytidine
G	Guanosine
T	Thymidine

1. INTRODUCTION

The Institute of Human Genetics is interested in those genes which are expressed in testis. For example the following genes have been characterized recently:

Acrosin, an endoprotease with trypsin-like substrate specificity, is localized in the acrosomal matrix as an enzymatically inactive zymogen, proacrosin, which is converted into the active form as a consequence of the acrosome reaction. Proacrosin is located in the sperm acrosome and has long been believed to be included in the proteolysis of the zona pellucida (ZP) of the oocyte, thus enabling the sperm to penetrate this extracellular matrix and to gain access to the oocyte plasma membrane (Klemm *et al.,* 1991). Using homologous recombination in our institute Adham *et al.,* (1997) have successfully produced male mice lacking acrosin in their sperm. However, these sperm can penetrate the ZP and fertilize oocytes. It was shown that acrosin-deficient sperm have a selective disadvantage with respect to wild type sperm. The acrosin-deficient sperm have a delay in ZP penetration when compared with wild-type sperm.

In mice carrying the autosomal recessive mutation "abnormal spermatozoon head shape" (*azh*) all spermatozoa display a highly abnormal head morphology that differs from the compact and hook-shaped head of the normal murine sperm (Hugenholtz *et al.,* 1984). Moreover, the *azh* mutation causes tail abnormalities often resulting in coiled sperm tails or in decapitation of the sperm head from the flagellum. Isolation and characterization of the murine *Hook1* gene and its predominant expression in testis has been reported (Mendoza *et al.,* 2002). It was found that two exons of the *Hook1* gene are deleted in the *azh* mutant mouse leading to a putative truncated protein that lacks both the conserved homodimerization domain and the putative organelle-binding domain. Disruption of *Hook1* function in the *azh* mutant mouse causes abnormal sperm head shape and fragile attachment of the flagellum to the sperm tail.

Impaired ciliary and flagellar functions resulting in male infertility and recurrent respiratory tract infections are found in patients suffering from Primary Ciliary Dyskinesia (PCD). In most cases, axonemal defects are present, i.e. PCD patients often

lack inner dynein arms in their sperm tails and cilia, supporting the hypothesis that mutations in dynein heavy chain genes may cause PCD. To elucidate the role of the mouse dynein heavy chain 7 (*MDHC7*) gene, which encodes a component of the inner dynein arm, the disruption of this dynein heavy chain isoform by homologous recombination has been reported (Neesen *et al.,* 2001). Due to the replacement of the ATP-binding site (P-loop) by the neomycin resistance gene, the *MDHC7* gene was inactivated. No functional gene product could be detected in MDHC7$^{-/-}$ animals. Loss of the dynein heavy chain results in male infertility and a reduction of ciliary beat frequency (CBF), but not in structural defects of tracheal cilia or sperm flagella.

The sexual dimorphic position of the gonads in mammals depends of two ligaments, the Cranial Suspensory Ligament (CSL) and the gubernaculum. During male embryogenesis, outgrowth of the gubernaculum and regression of the CSL results in transabdominal descent of the testes, whereas in the female, development of the CSL in conjunction with failure of the gubernaculum development holds the ovaries in a position lateral to the kidneys. Regression of the CSL and induction of gubernaculum development were suggested to be mediated by testosterone. *Insl3* a member of the insulin-like superfamily is specifically expressed in Leydig cell of the pre- and postnatal testis and in theca cells of the postnatal ovary. This sexual dimorphic pattern of *Insl3* expression led to the suggestion that the *Insl3* factor could play an important role in sexual differentiation and gonadal function. Key insights into the role of *Insl3* came from analysis of *Insl3* knock out mice (Zimmermann *et al.,* 1999). Knock out male mice showed cryptorchidism, which is a common disorder of sexual differentiation in human. Furthermore, the rat insulin II promoter-mouse-*Insl3* hybrid gene was designed to direct the overexpression of the *Insl3* gene in pancreatic β-cells during the pre- and postnatal development of male and female transgenic mice (Adham *et al.,* 2002). Expression of the transgenic allele rescued the cryptorchidism in male mutant. The overexpression of the *Insl3* in female mice causes the ovaries to descend into a position over the bladder. The results clearly demonstrate that *Insl3* play an important role in testicular descent during prenatal development attached to the abdominal wall via the well developed CSL.

1.1 Gene Trap Approach in Embryonic Stem Cells

Gene trapping in murine ES cells is attractive, because ES cells are accessible to genetic manipulation while retaining their pluripotency and capacity to form germ line chimeras. A promotorless reporter gene (βgal) fused in frame to a selection gene (neomycin) is integrated into the ES cell genome. The survival of the cells depends on the activation of the selection gene by the promoter of an endogenous gene, so that insertions into genes are selected over random insertions. At the same time, the marker gene is activated, so that the expression domains of the unknown (or trapped) gene can be traced by a simple staining method. The insertion of exogenous sequence facilitates cloning of the unknown trapped gene. Furthermore, the endogenous locus is disrupted by the insertion (GT vector), which can lead to a loss of function. Gene trap screens have been performed and reported by a number of researchers (Skarnes *et al.,* 1992, 1995; Wurst *et al.,* 1995; Forrester *et al.,* 1996; Chowdhury *et al.,* 1997; Holzschu *et al.,* 1997). Most researchers have emphasized finding restricted expression patterns, because this approach increases the number of trapped genes that have a specific role in the organ of interest. In cooperation with Professor Peter Gruss from the Max Planck Institute for Biophysical Chemistry, Göttingen, we have obtained several gene trap mouse lines, in which the reporter gene is found to be expressed in testis. Therefore, we hypothesized that the trapped genes in these mouse lines could be good candidates for relevant testis genes. The research group of Professor Gruss have generated many gene trap mouse lines by following ways:

- Special vectors have been constructed that allow trapping of genes in mouse embryonic stem cells (Figure 1.1). These gene trap vectors contain several functional units:
    - **Splice acceptor (SA) and splice donor sequences (SD) sites** which maintain the proper splicing between a gene trap vector and an endogenous trapped gene when integration occurs in introns
    - ***En-2* intron sequence**: it has been observed in a number of cases that linearized DNA introduced into ES cells by electroporation can lose sequences from either end or before integration into the genome.

Therefore this "buffer" sequence can be added to the 5' end of the vector to protect the functional units

- **ATG codon**: a reporter gene can be provided with its own efficient start of translation (Kozak consensus sequence)
- **IRES sequence**: eukaryotic RNA translation is usually dependent on 5' cap-mediated ribosome binding site. Addition of viral internal ribosomal entry site (IRES) sequences between SA site and coding sequence of the reporter gene in GT vector should allow initiation of translation of the reporter gene independent of upstream reading frame
- **Reporter gene (βgal)**: the reporter gene most commonly used to date in GT vectors is the *E.coli* β-galactosidase gene (Lac-Z). A promoterless Lac-Z gene can integrate randomly into the host genome, but it is activated only if a correct integration within a transcriptionally active endogenous gene has occurred. The β-galactosidase protein is quite stable and due to its enzymatic activity, easy to detect by X-Gal staining in cultured cells, mouse embryos, and adult tissues.
- **Selector gene (neo):** a neomycin phosphotransferase gene was used to select the ES cells containing gene trap insertion.

- Gene trap vectors were introduced into ES cells by electroporation
- Clones containing gene trap insertions were selected by G418 (neomycin) resistance
- Neomycin resistance clones were then picked into 96-well plates which are later replicated to generate frozen stocks for long term storage, cell lysates for RNA isolation and to assay reporter gene expression in the expression screens developed around *in vitro* differentiation assays.
- ES clones were aggregated with morula-stage embryos from NRMI mice to generate mouse chimeras.
- Chimeric mice were bred to NMRI mice and germ line transmitted F1 offspring were obtained.

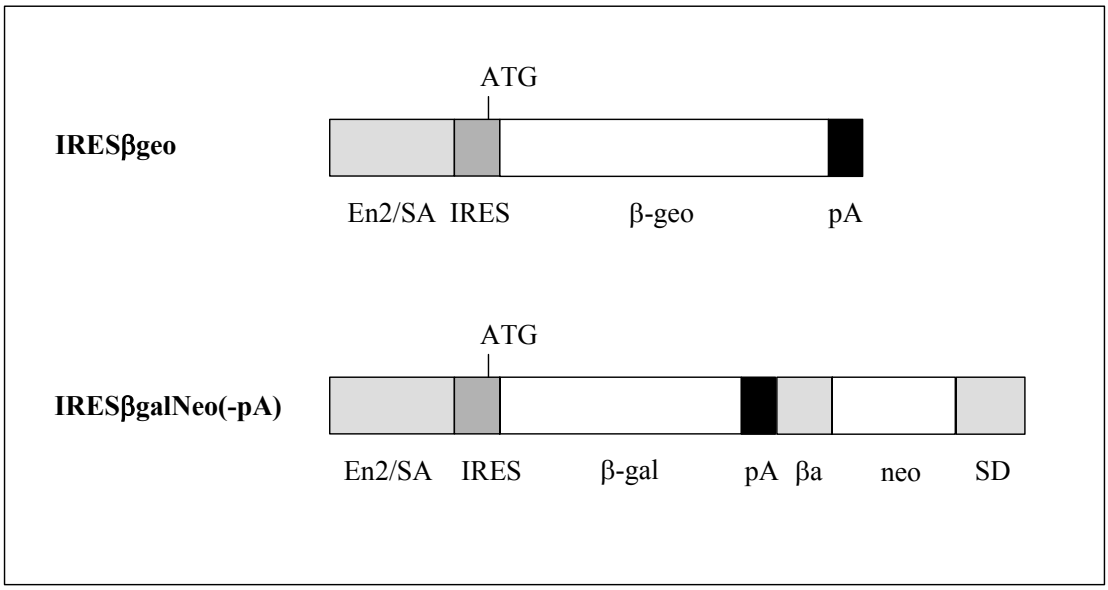


Figure 1.1. **Schematic representations of the gene trap vectors used in this study**. Abbreviations: SA, mouse *engrailed-2* splice acceptor site; SD, mouse *Pax-2* splice donor site; IRES, internal ribosomal entry site from the encephalomyocarditis virus; ATG, translation initiation codon; β-geo, a fusion gene between the β-*galactosidase* (βgal) and neomycin *phosphotransferase* (neo) genes; pA, SV40 polyadenylation signal; βa, human β-actin promoter. Both vectors were subcloned in the Bluescript KS+ plasmid and linearized with *Sca I* restriction enzyme before electroporation into ES cells. The IRESβgeo vector was used to generate gene trap mouse line 8 and the IRESβgalNeo(-pA) was used to generate the gene trap mouse line 16.

1.2 Gene Trap Mouse Lines

In my Diploma thesis, I have characterized several gene trap mouse lines concerning testicular expression using X-gal staining on sections of testis and other organs. Two of these lines (Line 8 and Line 16) were selected for further characterization because reporter gene was expressed strongly in testis.

For genotyping of the gene trap mice for GT vector integration, DNA was isolated from tail biopsies, as described previously (Sambrook *et al.,* 1989) PCR amplification of 508 bp of the Lac-Z gene was amplified by using sense and antisense primers (Figure 1.2).

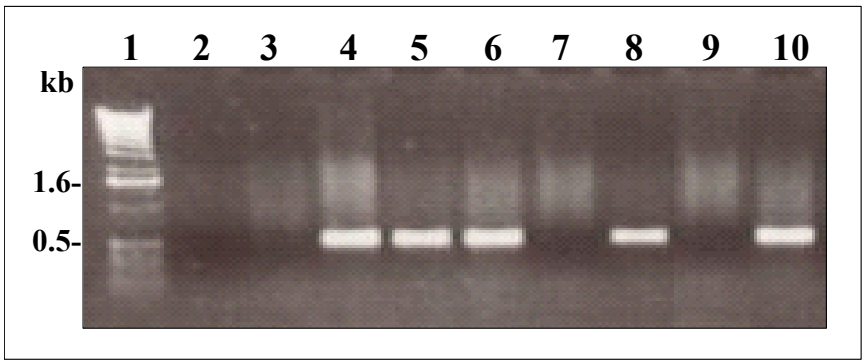


Figure 1.2. Lac-Z PCR for genotyping of offspring. (1) Marker for size of the DNA fragment. (2) Sample without DNA; (3 and 4) Negative and positive controls, DNA was isolated from wild type and heterozygous animals, respectively; (5, 6, 8 and 10) Lac-Z positive offspring, those either heterozygous or homozygous for the GT vector integration; (7 and 9) wild type offspring.

Lac-Z positive males were crossed with wild type females to produce the heterozygous male and female animals. Then, the heterozygous animals were bred to generate homozygous animals. Using Lac-Z PCR we were not able to distinguish between homozygous and heterozygous offspring which derived from heterozygous parents. Therefore, we performed "test breeding" for both gene trap lines. For the test breeding: 20 Lac-Z positive males were breeded with wild type females, separately. From these breedings, we expected that some males (homozygous for the trapped gene) would produce only Lac-Z positive offspring while other males (heterozygous for the trapped gene) would produce Lac-Z positive (heterozygous) and Lac-Z negative (wild type) offspring. All males were fertile and they produced Lac-Z negative and Lac-Z positive offspring. These results clearly demonstrated that all tested male mice from F1-breeding were heterozygous. Homozygous offspring could not be obtained. The same result was obtained when F2-females were breeded with wild type males. Because of these results we hypothesized those embryos homozygous for the trapped genes die before birth.

Objectives in this work

1. Identification of the trapped genes in Line 8 and Line 16
2. Structure and expression analysis of these genes
3. Determination of the time of embryonic lethality
4. Analysis of molecular basis of embryonic lethality

2. MATERIAL AND METHODS

2.1 Materials

2.1.1 Chemicals

Acrylamide/Bis-acrylamide 40% (w/v) (19:1)	Gibco/BRL, Karlsruhe
Acetic acid	Merck, Darmstadt
Agar	Fluka, Neu-Ulm
Agarose	Gibco/BRL, Karlsruhe
Ammonium acetate	Fluka, Neu Ulm
Ammonium persulfate	Sigma, Deisenhofen
Ampicillin	Sigma, Deisenhofen
Ampuwa	Fresenius, Bad Homburg
Bacto-tryptone	Difco, Detroit, USA
Bacto-Yeast-Extract	Difco, Detroit, USA
BCIP	Boehringer, Mannheim
Bisacrylamide	Serva, Heidelberg
Blocking reagent	Roshe, Penzberg
Bromophenol blue	Sigma, Deisenhofen
BSA (Factor V)	Biomol, Hamburg
Cell culture media (DMEM)	Gibco/BRL, Eggenstein
Coomasie G-250	Sigma, Deisenhofen
Chloroform	Baker, Deventer, Holland
DAPI	Vector, Burfingame
Dextran sulfate	Pharmacia, Freiburg
Dimethyl sulfoxide (DMSO)	Sigma, Deisenhofen
Dithiothreitol	Sigma, Deisenhofen
dNTPs	GibcoBRL, Karlsruhe
Ethanol	Baker, Deventer, NL
Ethidium bromide	ROTH, Karlshure
FCS	Gibco/BRL, Karlsruhe

Formaldehyde	Gibco/BRL, Karlsruhe
Ficoll 400	Pharmacia, Freiburg
Formamide	Fluka, Neu Ulm
Glutaraldehyde	Sigma, Deisenhofen
Glycerol	Gibco/BRL, Karlsruhe
Glycine	Biomol, Hamburg
Goat serum	Sigma, Deisenhofen
HCl	Merck, Darmstadt
H_2O_2	Merck, Darmstadt
HEPES	Merck, Darmstadt
IPTG	Biomol, Hamburg
Isopropanol	Merck, Darmstadt
KCl	Merck, Darmstadt
M2-medium	Sigma, Deisenhofen
M16-medium	Sigma, Deisenhofen
Methanol	Merck, Darmstadt
$MgCl_2$	Merck, Darmstadt
MOPS	Merck, Darmstadt
Methyl benzoat	Fulka, Neu Ulm
β-mercaptoethanol	Serva, Heidelberg
Mineral oil	Sigma, Deisenhofen
Na acetate	Merck, Darmstadt
Na citrate	Merck, Darmstadt
NaCl	Merck, Darmstadt
Na_2HPO_4	Merck, Darmstadt
NaH_2PO_4	Merck, Darmstadt
NaOH	Merck, Darmstadt
NBT	Boehringer, Mannheim
PBS	Gibco/BRL, Karlsruhe
Phosphoric acid	Merck, Darmstadt

Pepton	Gibco/BRL, Karlsruhe
Picric acid	Fulka, Neu Ulm
Phenol	Gibco/BRL, Eggenstein
Polyethylene glycol 6000	Serva, Heidelberg
Proteinase K	Pharmacia, Freiburg
[α^{32}P] dCTP	Amersham, Braunschweig
RNase A	Sigma, Diesenhofen
RNase Inhibitor	Roshe, Penzberg
RNAse away	Biomol, Hamburg
Salmon sperm DNA	Sigma, Deisenhofen
SDS	Serva, Heidelberg
Taq-DNA-Polymerase	Gibco/BRL, Eggenstein, FINNzymes, Finland, Amersham, Braunschweig
T4-DNA-Ligase	Gibco/BRL, Eggenstein
TEMED	Serva, Heidelberg
Triton X-100	Serva, Heidelberg
Tris	Sigma, Deisenhofen
Tween-20	Fluka, Deisenhofen
X-Gal	Biomol, Hamburg
Xylen cyanole	Bio-Rad, München

All those chemicals which are not mentioned above were obtained from Merck, Darmstadt or ROTH, Karlsruhe.

2.1.2 Kits

DYEnamic ET-Terminator mix	Amersham, Braunschweig
Gel extraction kit	Qiagen, Hilden
Plasmid mini preparation kit	Qiagen, Hilden
Plasmid midi preparation kit	Qiagen, Hilden
PCR Purification kit	Qiagen, Hilden
QIAquick Gel Extraction kit	Qiagen, Hilden
5`RACE System Version 2.0	GibcoBRL, Eggenstein
pGEM-T Easy cloning system	Promega, Mannheim
HexaLabel^{TM} DNA labeling kit	MBI Fermentas, St.Leon-Rot
GST Band Purification kit	Novagen, Darmstadt
Expand long tamplate PCR system	Beringer, Mannheim

2.1.3 Frequently used Buffers and Solutions

Solutions were prepared according to Sambrook *et al*., (1989) with deionised dH_2O, unless, otherwise stated.

Denaturation solution	1.5 M NaCl 0.5 M NaOH
50x Denhardt´s Solution	1 % Ficoll 1 % Polyvinylpyrrolidon 1 % BSA, pH 7.0 sterile filtrated solution was stored at –20 °C.
DEPC-dH_2O	0.1 % (v/v) Dimethyl-dicarbonate (DEPC) was solved in dH_2O, incubated 24 h at RT and autoclaved afterwards.
Glycerol loading buffer	10 mM tris/HCl (pH 7.5) 10 mM EDTA (pH 8) 0.025% bromophenol blue

	0.025% xylenecyanol 30% glycerol
5x TBE buffer	450 mM tris-base 450 mM boric acid 20 mM EDTA (pH 8)
10 X MOPS Buffer	41.8 g MOPS 16.6 ml 3 M sodiumacetate 20 ml 0.5 M EDTA
Neutralisation solution	1.5 M NaCl 1 M tris/HCl (pH 7.0)
10x PBS buffer	1.3 M NaCl 70 mM Na_2 HPO_4 30 mM NaH_2 HPO_4, (pH 7.4)
PBT buffer	0.1% tween-20 in 1x PBS
20x SSC	3 M NaCl 0.3 M sodium acetate, (pH 7.0)
5x TBE buffer	455 mM tris 445 mM boric acid 10 mM EDTA
10x TE buffer	100 mM tris/HCl, (pH 8.0) 1 mM EDTA

2.1.4 Laboratory Material

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

Whatman blotting paper (GB 002, GB 003 and GB 004)	Schleicher and Schüll, Dassel
Cell culture flask	Greiner, Nürtingen
Disposable filter Minisart NMI	Sartorius, Göttingen
Filter Paper 0858	Schleicher and Schüll, Dassel
HPTLC Aluminum Folio	Merck, Darmstadt
Hybond C	Amersham, Braunschweig
Hybond N	Amersham, Braunschweig
Petri dishes	Greiner, Nürtingen
Pipette tips	Eppendorf, Hamburg
Microcentrifuge tubes	Eppendorf, Hamburg
Transfection flasks	Lab-Tek/Nalge, Nunc, IL, USA
X-ray films	Amersham, Braunschweig
Superfrost Slides	Menzel, Gläser

2.1.5 Culture Medium

2.1.5.1 Bacterial Media

(LB medium)

All media for bacteria were prepared with dH_2O, autoclaved and stored at 4^oC. Plates were also stored at 4^oC. Antibiotics, IPTG and X-Gal were added to the cooled (~50^oC) sterile media.

Medium	Composition
LB-Medium	1% (w/v) peptone or bacto-tryptone 0.5% (w/v) yeast extract 0.5% (w/v) NaCl LB plates with 1.5% (w/v) Agar
LB-Amp-medium	LB medium with 100 µg/ml ampicillin
LB-Kan-medium	LB medium with 25 µg/ml kanamycin
LB-Amp-IPTG-X-Gal medium (Oja)	LB medium with 100 µg/ml ampicillin, 833 µM IPTG, 0.4% (w/v) X-Gal

Antibiotic stock solutions:

Ampicillin	100 mg/ml
Kanamycin	25 mg/ml
IPTG	0.2 g/ml
X-Gal	2% in DMF

All solutions were prepared with dH_2O (except X-Gal), sterile filtrated and stored at -20°C.

2.1.5.2. Yeast Media

Medium	Composition
YPD medium	20 g/L difco peptone 10 g/L yeast extract 20 g/L agar (for plates only)
SD medium	6.7 g yeast nitrogen base without amino acids 20 g agar (for plates only) 850 ml dH_2O 100 ml of the appropriate sterile 10x Dropout Solution. Allow medium cool to ~55^oC before adding additional adenine, X-gal, dextrose if necessary

2.1.5.3 Mammalian Cell Culture Media

M2 and M16 media were purchased from Sigma, (Deisenhofen) and were used for washing and cultivation of mouse preimplantation embryos.

Dulbecco's MEM (DMEM) medium containing:

10% fetal calf serum (FCS)
2 mM L-Glutamine
1% of penicillin (6 mg/ml)/streptomycin (5 mg/ml) solution

This medium was used to culture the NIH 3T3 cells. For long time storage of the cells in liquid nitrogen, the following freezing medium was used:

Freezing medium: 30% DMEM medium
50% FCS
20% DMSO

2.1.6 Sterilisation of Solutions and Equipments

All solutions which were 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 described above. Glass wares were sterilised overnight in an oven at 220°C. The solutions were prepared with deionised water or Ampuwa H_2O (pyrogene free H_2O for molecular biology work). The solutions for RNA preparations and RNA analyses were handled with dimethyl-dicarbonate (DEPC) to inactivate any RNase contamination.

2.1.7 Biological Materials

2.1.7.1 Vectors

The following vectors were used:

pBluesript SK (+/-)	(Stratagene)	for standard cloning
pBluesript KS (+/-)	(Stratagene)	for standard cloning
pGEM-T Easy	(Promega)	to clone PCR fragments
pEGFP-C1	(Clontech)	to generate a green fluorescent fusion protein
pET 41 (a-b)	(Novagen)	to generate a GST fusion protein, expressed in bacterial cells
pTriEx 1.1 Neo	(Novagen)	to generate a epitop tagged fusion protein, expressed in mammalian cells
pGADT7	(Clontech)	for yeast two hybrid assay
pGBKT7	(Clontech)	for yeast two hybrid assay

2.1.7.2 Bacterial Strains

The following *Escherichia coli* strains were used:

E. coli DH5α, for cloning	K-12 strain, F^- Φ80*lacZ*ΔM15 Δ(LacZYA-*arg*F) U169 deoR *rec*A1 *end*A1 *hsd*R17 (r_k^-, m_k^+) *pho*A supE44 λ^- *thi*-1 gyrA96 relA1

E. coli BL21, for cloning	B strain, F^- *omp*T *hsd*$S_B(r_B^- m_B^-)$ *gal Dcm*
E. coli BL21 (DE3), as an expression host	B strain, F^- *omp*T *hsd*$S_B(r_B^- m_B^-)$ *gal dcm* (DE3)

2.1.7.3. Yeast cells

The following yeast strains were used for YTH assay:

Y187 for yeast mating	MATα, *ura*3-52, his3-200, *ade* 2-101, *trp* 901, *leu* 2-3, 112, gal4Δ, *met*$^-$, gal80Δ, URA3:: $GAL1_{UAS}$-$GAL1_{TATA}$-lacZ, MEL1
AH109 for cloning	MATa, *trp*1-191, *leu*2-3, 112, *ura*3-52, *his*3-200, gal4Δ, gal80Δ, LYS2:: $GAL1_{UAS}$-$GAL1_{TATA}$-HIS3, MEL1, $GAL2_{UAS}$-$GAL2_{TATA}$- ADE2, URA3::$MEL1_{UAS}$-$MEL1_{TATA}$-lacZ

2.1.7.4 Mammalian Cells

NIH 3T3 cells (Immortalized swiss albino mouse embryo fibroblasts, Rockville, USA "NIH Swiss Mouse") were used for GFP experiments.
Mouse preimplantation embryonic cells from wild type and gene trap mouse lines were used for genotypical and phenotypical analyses.

2.1.7.5 Animals

Mouse strain C57BL/6J was ordered from Institute of Human Genetics, Göttingen. Mice were kept four to six per cage after weaning at the age of 21 days in a room with controlled light and darkness cycle (12 hrs light, 12 hrs darkness) and room temperature 21^oC. Animals have free access to standard mouse chow and tap water. The mice were specific pathogen-free and they were routinely screened for common mouse pathogens. All

procedures using mice were approved by the University of Göttingen Ethical Committee on Use and Care of Animals.

2.1.8 Synthetic Oligonucleotide Primers

All the primers used in this work were synthesized by the company Roth (Karlsruhe) or Eurogentec (Köln) and were adjusted to a stock concentration of 100 pmol/µl with Ampuwa H_2O.

Used for	Primer name	Primer sequence (5'-3')
genotype PCR	Gen F1	TGTTAGACGAGCTTTCTACCAGTG
	Gen R2	ATGTAACCCTAAAGGGGTTAGGAC
	Gen R1 nest.	CAATGTGCAAGTAGGGAGTGAAGC
	CSH-F4	GTTCAGATTCAAGACAGGAAGTTGC
	Trans-rev	CTTGCAGCGGCCGCCTCAGTAGTCAGCA
	Trans-forw	GACACCCCATGGTGAGCTTCCTGAGACGA
RT-PCR	Ex5-f	CCAAATAATCTTCGTGAGGAGCTAA
	Ex6-r	TCCCAC TGAGAATATCTTGTTTTAGC
	Ex9-f	CACAAGCAGAAACCAATTACCTGAA
	Ex10-r	TTGGAGTCAGTCCCAAACTATAGTC
	Ex13-f	GCAAGAAGATCAACCTCCTTTGAAAG
	Ex14-r	CAAGATCTTCAGGTTGTGTTGTG
	Ex16-r	GTGGGAGATTCTCTATCTCCACCTG
	Ex19-f	GCCATTGAGAACCCAGCATTGAATG
	Ex24-f	GCATCTTTGGCAGGGACAGATGAAT
	ORF-F1	TCGAGTGAAACTTGGGAAGCGTACGAG
	ORF-F2	TCCGAAATGCTGAGCTTCCTGAGACGAA
	cDNA-R1	TCCAACTGGTAAGCCTAGTAGGG
	cDNA-R2	GAAACAGCAATGAAAGGCTCTG
GFP expriment	14R-Apa	GATAATGGGCCCTCTTCAGGTTGTGTTGT
	15F-Sal	CATTAAGTCGACCTTAATGCTTCCGCTCA
vector specific primers	T7	TAATACGACTCACTATAGGG
	T3	ATTAACCCTTCACTAAAG
	Sp6	AGGTGACACTATAGAATAC
	pMal F	CGGTCGTCAGACTGTCGATGAAGCC
	pMal R	CCAGGGTTTGCCCAGTCACGACGTT
5'RACE PCR	GSP1	CCAAAAGACGGCAATATGGTGG
	GSP2	AACGCACACCGGCCTTATTC
	GSP3	GCTGCTCTGTCAGGTACCTGTTGG

2.1.9 DNA Libraries

The following DNA libraries were used:

Library	Vector	Host cell	Cloning site	Insert size
Mouse testis MACHMAKER cDNA library (Clontech)	pACT2	E. coli	*Xho I*-$(dT)_{15}$	0.5-3.0 Kb
Mouse-genomic DNA library (RZPD), 11 filters set	Lawrist 7	E. coli DH10B	*NotI/ BamHI*	20-45 Kb

2.1.10 Antibodies

The following antibodies were used:

Name	Company
mouse monoclonal antibody against α-tubulin (clone B-5-1-2)	Sigma-Aldrich Chemie GmbH, Munich
anti-rabbit IgG alkaline phosphatase conjugate	Sigma-Aldrich Chemie GmbH, Munich
anti-mouse IgG alkaline phosphatase conjugate	Sigma-Aldrich Chemie GmbH, Munich
anti-rabbit IgG (whole molecule) Cy3 conjugate	Sigma-Aldrich Chemie GmbH, Munich
anti-mouse IgG (whole molecule) FITC conjugate	Sigma-Aldrich Chemie GmbH, Munich
anti-6X Histidine Antibody from mouse	R&D systems GmbH, Wiesbaden
anti-mouse GM-130 (Golgi marker) IgG1	BD Transduction Laboratories, Heidelberg
Anti-mouse γ-adaptin	BD Transduction Laboratories, Heidelberg

2.2 Methods

2.2.1 Isolation of Nucleic Acids

2.2.1.1 Mini-Preparation of Plasmid DNA

This method was used to to produce small amounts of pure plasmid DNA. From an overnight bacterial culture 1.5 ml were transferred to an Eppendorf tube and centrifuged at 4,000x g for 10 min. The supernatant was discarded and the rest of the medium was eliminated by inversion on absorbent paper. The pellet was resuspended in 150 µl of buffer P1. For lysis 200 µl of P2 buffer was added and mixed by inverting the tube 5-6 times, the incubation was conducted by 30 sec at RT. The sample was neutralized with 200 µl of P3 buffer and mixed by inverting 4-5 times. After centrifugation for 15 min at 14,000x g, the supernatant was pipetted into a new Eppendorf tube and centrifuged again under the same conditions. The DNA in the supernatant was precipitated by adding 3 volumes of ice cold 100% ethanol and pelleted by centrifugation at 14,000x g for 15 min. The DNA pellet was washed with 500 µl of 70% ethanol and centrifuged again for 5 min. Then the pellet was dried and resuspended in 30 µl of Ampuwa H_2O.

Buffer P1 (Resuspention buffer)	50 mM	Tris/HCl, pH 8.0
	10 mM	EDTA
	100 µg/ml	RNase A
Buffer P2 (Lysis buffer)	200 mM	NaOH
	1% (w/v)	SDS
Buffer P3 (Neutralization buffer)	3 M	Potassium acetate, pH 5.5

2.2.1.2 Midi-Preparation of Plasmid DNA

High quality plasmid DNA was prepared with Anion-Exchange Resin columns from QIAGEN (Hilden) Plasmid Midi kit. For a midi preparation, 100 ml of an overnight culture (in LB medium plus antibiotic) was centrifuged at 4,000x g for 10 min. The bacterial pellet was resuspended in P1 buffer containing 100 μg/ml of RNase A. After lysis (P2 buffer) and neutralisation (P3 buffer), the suspension was recentrifuged at 13,000x g for 15 min.

The supernatant containing the DNA was applied to a previously equilibrated column (with QBT buffer), allowed to run by gravity and to bind the DNA to the anion exchange matrix. The DNA was then purified from metabolites in 2 washing steps with buffer QC and eluted with an elution buffer (QF) and isopropanol was added followed by centrifugation at 13,000x g, 4°C for 30 min to pellet DNA. The pellet was dried, resuspended in Ampuwa H_2O and stored at –20°C until use.

Buffer QBT (Equilibration buffer)	750 mM	NaCl
	50 mM	MOPS, pH 7.0
	15 %	isopropanol (v/v)
	0.15%	Triton X-100 (v/v)
Buffer QC (Wash buffer)	1.0 M	NaCl
	50 mM	MOPS, pH 7.0
	15 %	isopropanol (v/v)
Buffer QF (Elution buffer)	1.25 M	NaCl
	50 mM	Tris/HCl, pH 8.5
	15 %	isopropanol (v/v)

2.2.1.3 Isolation of Genomic DNA from Tissue Samples

(Laird et al., 1991)

1.5 to 2.0 cm of the mouse tail was incubated in 700 µl of lysis buffer I at 55°C overnight in a thermomixer. To the tissue lysate, and equal volume of phenol was added, mixed by inverting several times and centrifuged at 8000xg at room temperature for 5 min. After transferring the upper aqueous layer into a new tube, the same procedure was repeated, first with 1:1 ratio of phenol and chloroform and then with chloroform. Finally, the DNA was precipitated with 0.7 volume of isopropanol, washed with 70% ethanol, and dissolved in 100-200 µl of TE buffer or H_2O (Ampuwa) and incubated at 60°C for 15 min.

Lysis buffer I	100 mM Tris/HCl (pH 8.0)
	100 mM NaCl
	100 mM EDTA
	0.2 mg/ml Proteinase K
	0.5% SDS

2.2.1.4 Isolation of Genomic DNA from Cultured Blastocysts

Culturing of blastocysts is described in 2.2.21.3. After culturing, DNA was prepared by incubating the individual embryos with 20 µl of lysis buffer II for 4 h at 60^{o}C followed by incubation at 90^{o}C for 30 min. Three microliters of embryonic DNA was used for PCR.

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

2.2.1.5 Preparation of Genomic DNA from Preimplantation Embryos

Embryos were obtained from superovulated and plugged female mice (2.2.21.1) at the following stages: 1-cell (18-20 hrs p.c.), 2-cell (42-44 hrs p.c.), 4-cell (50-52 hrs p.c.), 8 cell (66-72 hrs p.c.) morula and blastocysts (90-92 hrs p.c.). Individual embryos were collected in a PCR tube (0.2 ml) containing 5 µl Ampuwa H_2O. Samples were repeatedly (2-3 cycles) frozen and thawed. Then samples were directly used for genotyping by using PCR approach.

2.2.1.6 Isolation of Total RNA from Tissue Samples

In order to avoid any RNase activity, all the equipment and solutions used for RNA isolation were previously treated with RNase AWAY and DEPC water. Special RNase free Eppendorf cups were used during the procedure. In this method, RNA is extracted in a denaturing solution of Guanidine-Isothiocyanate (GITC). Adult mice tissues (brain, liver, lung, kidney, ovary, testis, spleen, heart) were collected and used immediately for the RNA isolation. All of the collected tissues were separately prepared. 100 mg of each tissue was mixed with 1 ml of the GITC buffer solution (Total-RNA solution, Biomol) and then homogenized for 10-20 sec with an electric homogeniser (Ultraturrax T25, Schütt). Then 0.2 volume of cold chloroform was added and mixed. The samples were placed on ice for 5 min. After centrifugation of the suspension at 3.200x g for 10 min at 4^{o}C, the upper phase was taken and mixed with 1 volume of ice-cold isopropanol and placed on ice for 15 min. Then the RNA was precipitated by centrifugation at 13.000x g at 4°C and the pellet was washed with 1 ml of ice-cold 75% ethanol. The pellet was dried at room temperature (RT) and finally resuspended in 50-100 µl of DEPC-H_2O. To avoid protein contamination, the samples were incubated at 65°C for 0.5-2 min, cooled on ice and the denatured proteins were pelleted by centrifugation for 30 sec. The RNA aliquots were stored at –80°C. The RNA concentration was determined with a spectrophotometer.

2.2.1.7 Phenol-Chloroform Extraction

One volume of phenol was added to the sample, vortexed and centrifuged at 13,000x g for 2 min. The supernatant was transferred to a new reaction tube and 1 volume of chloroform

was added. After vortexing and centrifugating as before, the supernatant was transferred to a fresh reaction tube for ethanol precipitation.

2.2.1.8 Ethanol Precipitation of Nucleic Acids

To precipitate nucleic acids, the salt concentration of the sample was adjusted to 300 mM with NaCl or with 0.1 volume of 3 M NaAc (pH 4,8) and 2.5 volume of 100 % ethanol was added. The sample was placed at -80°C for 20 min and pelleted by centrifugation (13,000x g, 15 min, 4°C). The pellet was washed with 70 % ethanol, centrifuged, dried and solved in an appropriate volume of H_2O. The sample was then checked electrophoretically.

2.2.2 Enzymatic Modifications of DNA

2.2.2.1 Dephosphorylation of 5' Ends of DNA

The dephosphorylation of the 5' ends of DNA prevents the recircularisation of a previously restricted vector with compatible cohesive ends during a ligation experiment. This was achieved with the alkaline phosphatase (CIP from calf intestine), which removes the 5' phosphate from the free ends of DNA molecules. Therefore, the efficiency of the ligation of foreign DNA into a vector can be greatly enhanced. This treatment of the DNA was carried out as follows:

1 µg	vector DNA
1 µl	CIP, alkaline phosphatase (5 U/µl)
1 µl	10x CIP buffer
up to 20 µl	Ampuwa dH_2O

The reaction was performed at 37°C for 1 h and was terminated by heat inactivation at 68°C for 15 min and addition of 0.1 volumes of 10% (w/v) SDS and 0.1 volume of 10x TE buffer. After a phenol/chloroform extraction, the DNA was ethanol precipitated with 0.1 volume of 3 M NaAc (pH 4.8).

10 x CIP buffer:	500 mM	Tris/HCl, pH 9.0
	10 mM	$MgCl_2$
	1 mM	$ZnCl_2$
	10 mM	Spermidine

2.2.2.2 Cloning of DNA Fragments

Foreign DNA fragments were subcloned into vectors after digestion with restriction enzymes, electrophoresis separation (2.2.6.), isolation (2.2.7.) and finally ligation into a vector. The vector was restricted with one or two enzymes in the multicloning site (MCS). When one enzyme was used to restrict the vector, a dephosphorylation step of the 5´ends (2.2.2.1.) was necessary to prevent the religation of the vector. The ligation reaction was done as follows with an overnight incubation at 16°C:

0.5 µg	vector DNA
1-10x	molar ratio of foreign DNA fragment
1 µl	T4 DNA ligase enzyme (5U/µl)
1 µl	10x ligation buffer
up to 10 µl	Ampuwa H_2O

2.2.2.3 Cloning of PCR Fragments

PCR products (2.2.5.) were subcloned through the TA-ligation (TA: thymidine-adenine), if PCR was performed with a Taq polymerase that leaves a deoxyadenosine at the 3´end of the amplified fragment in the elongation phase of the PCR cycle. The PCR fragments were TA-ligated to the pGEM-T Easy vector (Promega). This vector contains an open insertion site inside the MCS that has a 3´overhanged thymidine where the PCR fragment ligates. The ligation reaction was done as described in 2.2.2.2.

2.2.3 Production of *E. coli* Competent Cells

By this procedure, the competence of *E. coli* cells to accept free DNA is artificially increased. This is achieved by modifying the cell wall with $CaCl_2$ and RbCl. 30 ml of LB medium was inoculated with a single bacterial colony (DH5α or BL21) and shaked overnight at 37°C. At the next day, 1 ml of this pre-culture was added to 100 ml of LB medium and incubated until an OD_{600}=0.5 was reached. After cooling down on ice for 10 min and centrifugation at 5,000x g, 4°C for 10 min (Sorvall RC5B) the cell pellet was resuspended carefully in 30 ml of TFB I (on ice) and was incubated 10 min on ice. After centrifugation at 5,000x g at 4°C for 10 min, the bacteria were resuspended in TFBII solution, frozen in 200 µl aliquots in liquid nitrogen and stored at –80°C until their use.

TFB I:	100 mM	RbCl
	50 mM	$MnCl_2$
	10 mM	$CaCl_2$
	30 mM	KAc, pH 5.8
	15 % (v/v)	Glycerin
TFBII:	10 mM	RbCl
	75 mM	$CaCl_2$
	10 mM	MOPS, pH 7.0
	15 % (v/v)	Glycerin

2.2.4 Transformation of Competent Bacterial Cells

200 µl of competent cells (*E. coli* DH5α or BL21) were thawed on ice (10 min), mixed with 5-10 µl of ligation reaction mixture and placed on ice for 20 min with occasional mixing. The transformation reaction mixture was then incubated at 42°C for 90 sec for a heat-shock and placed on ice for 2 more minutes. In order to accelerate the bacterial growth, 300 µl of LB medium was added to the reaction mixture, which was incubated at 37°C for 45 min under shaking. After incubation, 100-150 µl of the reaction mixture was spread on the proper selection plate and incubated overnight at 37°C.

2.2.5 Polymerase Chain Reaction

The polymerase chain reaction (PCR) 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 sequences that may be present at very low concentrations in very complex mixtures.

2.2.5.1 PCR for Amplification of DNA Fragments

The amplification cycles were performed in an automatic thermocycler. The PCR reaction contains in general the following substances.

10 ng	DNA
1 µl	forward primer (10pmol)
1 µl	reverse primer (10pmol)
1 µl	10mM dNTPs
5 µl	10x PCR buffer
1.5 µl	50mM $MgCl_2$
1 µl	*Taq* DNA Polymerase (5 U/µl)
Up to 50 µl	Ampuwa H_2O

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

	Cycles	Temperature	Time
Initial denaturation	1	94-97 ^{o}C	5-15 min
Denaturation		94-95 ^{o}C	30 sec-1 min
Annealing	25-35	46-75 ^{o}C	30 sec-2 min
Extension		72 ^{o}C	1 min/1000 bp
Final extension	1	72 ^{o}C	7-10 min
Cooling	1	4-8 ^{o}C	stop

2.2.5.2 One-Step RT-PCR

To obtain specific RT-PCR products, the QIAGEN OneStep RT-PCR kit was employed which contains optimized components that allow both reverse transcription and PCR amplification to take place in what is commonly referred to as a "one-step" reaction.

Master mix;	Per reaction
5 x Qiagen Onestep RT-PCR buffer	10 µl
dNTP mix (containing 10 mM of each dNTPs)	2 µl
Forward primer (10 pmol)	1 µl
Reverse primer (10 pmol)	1 µl
Qiagen OneStep RT-PCR Enzyme mix	2 µl
RNase inhibitor (20 units per 1 µl)	1 µl
RNase-free water	31 µl

2 µl (2 µg) of total RNA isolated from mouse tissues was added to 48 µl of prepared Master mix in a PCR tube, the sample was placed in the thermal cycler and the RT-PCR program run according to the user manual. After the amplification step, the sample was checked on an agarose gel. Thermal cycler conditions were:

Reverse transcription:	30 min	50°C
Initial PCR activation step:	15 min	95 °C
35 cycles		
Denaturation	30 sec	94 °C
Annealing	40 sec	56-60 °C (depending on primers)
Extension	1 min	72 °C

2.2.5.3 5'RACE PCR

Rapid Amplification of cDNA Ends (RACE) is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and unknown sequences at either the 3' or the 5'-end of the mRNA. 5'RACE method has been reviewed by both Frohman *et al.,* 1988. 5'RACE System; Version 2.0 (GibcoBRL, Karlsruhe) was used for this work. Method has following steps:

1. First strand cDNA

The First Strand cDNA was synthesized from total testicular RNA (2.2.13.1.) from a heterozygous gene trap animal, using a gene specific antisense primer (GSP-1). ~1µg of total RNA and 200 units SuperScript™ II RT were used and reaction was incubated for 50 min at 42^0C.

2. Purification of cDNA

Excess nucleotides and GSP1 were removed from the first strand product using GlassMax cartridge. Purification step was necessary for tailing of the first strand product.

3. TdT Tailing of cDNA

TdT tailing creates the abridged anchor primer (AAP) binding site on the 3'-end of the cDNA using the terminal deoxynucleotidil transferase. Following components were added to each tube:

DEPC-treated water	6.5 µl
5X tailing buffer	5 µl
2 mM dCTP	2,5 µl
GlassMax purified cDNA template	10 µl

Samples were incubated for 2 min at 94^{o}C, then chilled for 1 min on ice and 10 units of TdT were added to the samples, mixed gently, and further incubation was carried out at 37^{0}C for 10 min. TdT was inactivated at 65^{0}C for 10 min.

4. PCR of dC-tailed cDNA

Amplification of a targeted cDNA, synthesized with the 5'RACE System required priming with two oligonucleotides. The "Abridged Anchor Primer" (AAP), which was specific for the oligo-dC tail was added by TdT, served as the forward primer. The reverse primer (GSP2) should anneal to an internal (nested) site within the cDNA sequence (with respect to the primer used for first strand synthesis, GSP1).

5. Nested Amplification

Often a single PCR of 25-35 cycles will not generate enough specific products to be detected by ethidium bromide staining. A dilution (1:10) of the primary PCR product (above) was re-amplified using abridged amplification universal primer (AUAP) or universal amplification UAP and the nested GSP3 primer.

The products of the second PCR were analysed on agarose gel, then cloned (2.2.2.3.) and sequenced (2.2.8.)

2.2.6 Gel Electrophoresis for Separation of Nucleic Acids

2.2.6.1 DNA Horizontal Electrophoresis in Agarose Gels

DNA fragments were electrophoretically separated according to their molecular size in agarose gels (0.6%-2%). The agarose was (2-3 min) boiled in 0.5x TBE and after cooling

down (~ 60°C), 10 µl of ethidium bromide solution (5 mg/ml) per 100 ml gel was added and poured in a gel caster. The same buffer was used as electrophoresis buffer. Before loading the samples, 0.2-0.5 volumes of loading buffer was added and mixed. The samples were then loaded into the wells of the gel and electrophoresis was carried out at a steady voltage (100-150 V or 3-4 V/cm^2). Under this condition, the DNA moved towards the cathode. The DNA fragments were observed and photographed under UV light at a λ of 254 nm or 312 nm due to the intercalation of the fluorescent dye ethidium bromide into the double strand of DNA.

2.2.6.2 Vertical Agarose Gels

Restricted genomic DNA was separated in vertical chambers for gel electrophoresis (Glasgerätebau Ochs, Bovenden/Lenglern) with 1x E buffer as gel and running buffer. The lower anode chamber opening was sealed with a 2% agarose ground gel. After gelification, a 0.8-1.3% running agarose gel was poured. The DNA samples were mixed with 20% (v/v) loading buffer and run at 15 V overnight. Finally, the gel was photographed under UV light (254 nm) and transferred to a membrane (2.2.9.1.)

10x E Buffer:	300 mM	NaH_2PO_4
	267 mM	Na_2HPO_4
	50 mM	EDTA
		pH 7.3-7.4

2.2.6.3 RNA Electrophoresis in Denaturing Agarose Gels

For the separation of RNA by electrophoresis, denaturing vertical agarose gels were employed to avoid secondary structures formed in RNA due to hydrogen bonds. Before use, the chambers were treated with RNase ZAP (Ambion, Wiesbaden) to eliminate RNase, and finally rinsed with dH_2O. The gel was prepared by boiling 1.5 g of agarose with 87.5 ml of DEPC-H_2O and 12.5 ml of 10x E buffer, until complete solution. After cooling of the gel to approximately 50^{o}C, 25 ml of 37% formaldehyde was added and was immediately

poured into a vertical gel bed (RNase free). Before loading, the RNA samples were mixed with the following solutions:

	2 µl	10x E buffer
	3 µl	37 % formaldehyde
	8 µl	formamide
	20 µg	RNA sample
up to	50 µl	DEPC H_2O

The mixture was denatured at 65°C for 10 min then cooled on ice. 16 µl of stop mix that contained 0.2 µg of ethidium bromide was added and the samples were loaded into the wells of the gel. Electrophoresis was carried out at 6.5 V/cm^2 at 4^{o}C or at 4 V/cm^2 at RT with 1x E buffer until the blue dye of the loading buffer reached the front of the gel. The agarose gel was photographed under UV light (254 nm) and the RNA was transferred to a membrane (2.2.9.2.).

2.2.7 Isolation of DNA Fragments from Agarose Gels

For the isolation of DNA from agarose gels, the DNA band was cut out from the gel with a sterile scalpel and extracted with the QIAEX II kit (QIAGEN, Hilden). In this system, the agarose is solved in the presence of high NaJ concentrations. When the DNA is liberated, it binds to added glass particles. For this, 300 µl of buffer QX1 (containing NaJ) per 100 mg of agarose gel and 10 µl glass milk (QIAEX II) per 5 µg DNA were added. This step was followed by incubation at 55°C for 10 min with constant mixing and centrifugation at 14,000x g for 30 sec. The pellet was washed again with buffer QX1 and centrifuged like before. Then, the pellet was washed twice with the buffer PE and spinned, the pellet was air-dried at RT for 15 min. To elute DNA, 25 µl of Ampuwa water was added to the dry glass milk pellet and was incubated at 55°C under shaking for 10 min. The glass milk was pelleted by centrifugation at 13,000x g and the supernatant was transferred to a new Eppendorf tube: The centrifugation was repeated to eliminate glass milk rests.

2.2.8 DNA Sequencing

The sequencing method is modified derived from Sanger *et al*. (1977) using 4 different fluorescently marked ddNTPs (didesoxynucleosid-5´-triphosphate). The sequence reaction was:

	0.2- 1.5 µg	PCR product or plasmid DNA
	10 pmol	vector or gene specific primer
	4 µl	ET reaction mix (dNTPs, dideoxy dye terminators and *Taq* polymerase.
up to	20 µl	Ampuwa H_2O

Elongation and chain termination took place in a thermocycler (MWG). The following sequence PCR program was used:

Initial denaturation	98°C	5 min
25 cycles		
Denaturation	98°C	30 sec
Annealing	50°C	15 sec
Elongation	60°C	4 min

The reaction products were analyzed with automatic sequencing equipment.

2.2.9 Blotting Techniques

2.2.9.1 Southern Blotting

The Southern blot procedure refers to different sizes of denatured DNA molecules that are transferred from agarose gels onto a solid support such as nitrocellulose membranes (Hybond C, Amersham, Braunschweig) or nylon membranes (Hybond N, Amersham, Braunschweig) through the diffusion of salt solutions by capillarity transfer (Southern, 1975). These membranes are then hybridized with a labelled DNA probe.

Horizontal or vertical agarose gel electrophoresis (2.2.6.) was used for the separation of DNA fragments. After electrophoresis, the gel was shaken twice for 20 min in denaturating solution and then twice for 20 min in neutralization solution. Then, the gel was placed on 20x SSC soaked Whatman paper which ends are immersed in transfer solution and a recipient with 20x SSC. The membrane was placed directly on the gel and covered by a layer of 2x SSC soaked Whatman paper gel. The air bubbles were eliminated using a roller on the Whatman paper. More dry paper was placed on the wet Whatman paper and finally an object~200-500 g was placed on top. Then, the transfer was allowed to occur overnight at RT. The membrane was soaked in 2x SSC and the DNA was fixed onto it either by baking for 2 hrs at 80°C or by UV cross-linking (120 mJ; UV Stratalinker TM 1800, Stratagene, Heidelberg).

Alternatively, a Turboblotter^{TM} (Schleicher & Schuell, Dassel) was used to transfer the genomic DNA. 20 pieces of dry gel-sized Whatman paper (Schleicher & Schuell, GB004) were used, followed by 4 pieces of GB002 Whatman paper and the 2x SSC wetted membrane, followed by the gel, 3 pieces of 20x SSC wet GB002 and a GB004 paper bridge soaked in 20x SSC that was in contact with a channel filled with 20x SSC.

2.2.9.2 Northern Blotting

Northern blot hybridization is a procedure by which different sized RNA molecules are separated in a denaturing agarose gel (2.2.6.3.), transferred onto a solid support of nitrocellulose or nylon membrane and then subjected to hybridization with a labeled DNA or RNA probe.

The same procedure was used as described for the Southern blot analysis. However, in this case the gel was not denatured and the RNA was transferred directly onto the membrane using 20x SSPE as transfer buffer, prepared with DEPC treated H_2O. The membrane stripe with the RNA standard was cut out and shook gently with 5% acetic acid for 15 min and then placed 10 min in methylene blue solution (0.04% in 0.5 M NaAc, pH 5.2). The

exceeding methylene blue was eliminated by rinsing with H_2O until the blue dyed RNA bands became visible.

20x SSPE buffer	174 g	NaCl
	27.6 g	Na_2PO_4
	7.4 g	EDTA pH 7.4

2.2.9.3 Colony Transfer and Hybridization

The colony hybridization is a rapid and effective technique that detects recombinant sequences isolated directly from cells grown on plates and transferred to membranes. 88 mm ϕ nitrocellulose or nylon membranes (Optitran BA-S85, Schleicher & Schuell) were placed on the plates for 1-2 min to transfer the colonies to the filters, whereas reference position points were marked to identify later the positive colonies. The culture plate was incubated at 37°C, so the colonies could grow again. The marked membranes were placed on surfaces with the following solutions:

5 min	10% (w/v) SDS
3 min	Denaturation solution
10 min	Neutralisation solution
10 min	2xSSC

DNA was fixed by UV cross-linking. Then the membrane was ready for hybridization with a ^{32}P-labeled probe (2.2.10.). After hybridization the positive colonies were localized.

2.2.10 Radioactive Labeling of Nucleic Acids

To detect specific nucleic acid sequences in Northern and Southern blot filters, DNA probes were ^{32}P labeled with the Hexalabel™ DNA labeling kit (MBI Fermentas, St. Leon-

Rot). DNA (30-40 ng) was mixed with 10 µl of 5x reaction buffer and filled with Ampuwa H_2O to a final volume of 40 µl. This reaction was vortexed and denatured at 95°C for 10 min and then placed on ice. After cooling down, 3 µl of Mix C (without dCTP[^{32}P]) were added. The reaction was transferred to the isotope laboratory, where 1 µl of Klenow enzyme and 30-40 µCi α-[^{32}P]dCTP were added. The reaction was incubated at 37°C for 10 min after which 4 µl dNTPs were added and incubated for further 5 min. The reaction was stopped with 1 µl of 0.5 M EDTA, pH 8.0. The already labeled DNA was purified using MicroSpin^{TM} S-200 HR columns (Amersham, Braunschweig) by adding 150 µl Ampuwa H_2O and centrifugation at 3,000x g for 2 min. The liquid resulting from the centrifugation was denatured at 95°C for 10 min and placed on ice for 2 min. Then, the probe was added to the hybridization tube.

2.2.11 Hybridization Methods

2.2.11.1 Hybridization of 32P-Labeled DNA with Northern or Southern Membranes

The blotted nitrocellulose or nylon filters containing either DNA or RNA (2.2.9.) were rolled and placed into a hybridization tube (nucleic acid facing inwards), which was filled with 2x SSC and freed from air bubbles. Once the filters were placed, the 2x SSC was poured out and replaced with 15 ml of pre-warmed (65°C) hybridization solution and 500 µl of denatured (95°C, 10 min) salmon sperm DNA. The tubes were then pre-hybridized for 2-3 hrs in a rotation oven at 65°C. Then, the denatured radioactive probe (2.2.10.) was added to the tube and the hybridization was carried out overnight under the same conditions as the pre-hybridization.

At the next day, the hybridization solution was poured out and the filters were washed at 65°C with the washing solution 1 for 5-20 min, depending on the activity shown by the filters. The radioactive signals were checked with a hand monitor (β and γ detector, 122LB, Berthold, Bad Wildbad). A second wash with washing solution 2 was done for 5-10 min if the filter had an activity of >2 Bq/cm^2. Finally, the filters were wrapped into plastic

folio and exposed to an X-ray film in an autoradiography cassette, which was placed at –80°C.

Hybridization solution	5x	SSC
	5x	Denhardt´s solution
	10% (w/v)	Dextran sulfate
	0.5% (w/v)	SDS
Washing solution 1	2x	SSC
	0.1% (w/v)	SDS
Washing solution 2	0.2x	SSC
	0.1% (w/v)	SDS

2.2.11.2 Screening of Mouse Filter Libraries

The cosmid library (11 filters) was screened with ^{32}P-labeled gene specific probes to isolate the mouse genomic homologue of the gene in subject. The hybridization was done in a recipient appropriate for the size of the filters, where the filters were introduced with enough hybridization solution to cover them. The pre-hybridization was done with 1 ml of denatured (95°C, 5 min) salmon sperm DNA and 500 ml of hybridization solution pre-warmed to 65°C for 30 min. Then, the ^{32}P-labeled probe was added and the hybridization was done at this temperature overnight. At the next day the filters were rinsed shortly with the washing solution 3 and were sealed wet into plastic sheets. Then, the filters were exposed to an X-ray film in an autoradiography cassette and placed at –80 °C overnight.

1M Na_2HPO4 (pH 7.2)	316 ml	1 M NaH_2PO_4
	684 ml	1 M Na_2HPO_4

Hybridization buffer	0.5 M	Na Phosphate sol. pH 7.
	7%	SDS
	1 mM	EDTA
Washing solution 3	10 mM	Na Phosphate sol. pH 7.2
	0.1%	SDS

2.2.12 Histological techniques

2.2.12.1 Pre-treatment of Glass Slides

Glass slides were coated with organosilan so that the paraffin sections have a better adhesion on them. The slides were sterilized by boiling with a solution of HCl 0.1N and the resulting sterile slides were incubated in 1x Denhardt´s solution overnight. Later, they were fixed with ethanol/acetic acid (3:1) for 20 min. Then, rehydratation of the slides was performed in a decreasing series of ethanol (96%, 70%, 50%, and 30% each) for 5 min. The slides were then incubated overnight in a 1% organosilan solution at 70°C and they were washed several times with dH_2O. Next, they were baked at 100°C for several hours. In this condition, the slides can be stored until 6 months at 4°C. The slides were activated in buffered glutaraldehyde for 30 min for a better adhesive power. After the glutaraldehyde was rinsed for 5 min in dH_2O, the activation was stabilized in a 0.1 M Sodium-m-periodate solution for 15 min. After this, the slides were rinsed 3x for 5 min in 1x PBS and dried at 42°C. The activated slides can be stored at 4°C.

Organosilan	1% (v/v)	γ-aminopropyltrithoxysilane pH 3.45
Buffered glutaraldehyde	10% (v/v)	glutaraldehyde in 1x PBS, pH 7.0

2.2.12.2 Preparation of Paraffin Sections

The X-gal stained tissues (2.2.15.2.) were dehydrated for 1 h each in an increasing series of ethanol (50%, 70%, 90%, and 96%). For the paraffin embedding, the alcohol must be removed from the tissue, which was achieved with an overnight incubation in methyl benzoate. Washing twice for 10 min with Roticlear, the methyl benzoate was also removed. The embedding was done with a paraffin mixture with the commercial name of Paraplast^{TM}, which contains DMSO for a better infiltration in the tissue. The jars used in the procedure are temperated at 60° C. The Paraplast^{TM} was melt at 60°C in an oven for 1 h. The Roticlear was replaced with a 1:1 mix of Roticlear and Paraplast^{TM}, and the tissue was incubated at 60°C in this medium twice for 20 min. Then, the tissue was incubated 2x 20 min in Paraplast^{TM} alone. The tissue was transferred to the embedding mould, which was filled with liquid Paraplast^{TM} and laid at RT to solidify. After this, the paraffin block was cut and mounted onto the microtome to make sections of 2-10 μm. The sections were mounted on pre-treated slides with the help of a 70% ethanol drop at 42°C and laid at this temperature until the liquid evaporate. The paraffin sections were then stored at 4°C.

Tissue sections were incubated twice for 10 min in Roticlear to remove the paraffin. Then, sections were re-hydrated in a decreasing ethanol series (100%, 96%, 70%, 50%, and 30%) for 2 min each, finally washed in PBS or dH_2O and air-dried. Now the sections were ready for light microscopical analysis.

2.2.13 Protein and Biochemical Methods

2.2.13.1 Isolation of Total Protein from Mouse Tissue

PBS medium was added to the tissues and homogenization was done. After adding 1 volume of 2x SDS loading buffer, the samples were handled with ultrasound on ice 2x 10 min. Then, proteins were denatured 10 min in boiling water and cooled on ice. The samples were handled with ultrasound again and big tissue pieces were pelleted for 5 min at 13,000x g. 10 μl of the supernatant was separated in an SDS-PAA gel (2.2.13.3.) after mixing with 1 volume of 2x SDS loading buffer and the samples were stored at –20°C.

2.2.13.2 Determination of Protein Concentration

(Bradford, 1976)

To determine the protein concentration, Bio-Rad protein assay was employed which is a dye-binding assay based on the differential colour change of a dye in response to various concentrations of protein. The assay is based on the observation that the absorbance maximum for an acidic solution of Coomasie Blue G-250 shifts from 494 to 595 nm when binding to a protein occurs. The BSA stock solution of 1 mg/ml was diluted in order to obtain standard dilutions in range of 10 µg/ml to 100 µg/ml. The Bio-Rad's color reagent was diluted 1:5 with H_2O, and filtered through 0.45 µm filters. In a 96-well microtiter plate, 20 µl of each standard dilution and the samples to be measured were pipetted with 280 µl of the color reagent. The absorption of the colour reaction was measured at 595 nm in a microplate reader (Microplate Reader 450, Bio-Rad München).

2.2.13.3 SDS-Polyacrylamide (PAA) Gel Electrophoresis

PAA gels from 10-15% were used to separate protein mixtures according to their different molecular weights. Before electrophoretic separation, proteins were prepared. For this, tissues were homogenized in an appropriate buffer mechanically and/or by ultrasound handling (1 min). Supernatants were mixed with 1 volume of 2x SDS-loading buffer, denatured by boiling 10 min and centrifuged 2 min at full speed. Finally, samples were cooled, handled again with ultrasound and centrifuged for 5 min (13,000x g) to pellet the unsolved material. Normally, 10 µl of the sample were loaded into a gel well. To prepare SDS-PAA gels, the following components for the separating gel listed in the table were mixed together and poured between glass plates. The separating gel was covered with some dH_2O and let at least 20 min to polymerize. Then, the water was removed and the collecting gel was prepared according to the table, mixed and poured over the separation gel. The gel comb was laid avoiding air bubbles. After the polymerization, the gel was attached to the

electrophoresis chamber and covered with electrophoresis buffer. The electrophoresis was carried out for 1.5-5 h at 45 mA depends on the size of the gel.

SEPARATION GEL	**10%**	**12%**	**15%**
Acrylamide 29:1 (40%)	5 ml	6 ml	7 ml
Lower Tris	5 ml	5 ml	5 ml
dH_2O	10 ml	9 ml	8 ml
SDS 10%	100 µl	100 µl	100 µl
APS 10%	200 µl	200 µl	200 µl
TEMED	6 µl	6 µl	6 µl

COLLECTING GEL	**3%**
Acrylamide 29:1 (40%)	1.3 ml
Upper Tris	1.3 ml
dH_2O	7.4 ml
SDS 10%	50 µl
APS 10%	50 µl
TEMED	6 µl

Electrophoresis buffer:	1.44% (w/v)	Glycine
	0.3% (w/v)	Tris base
	0.1% (w/v)	SDS
Acrylamide 29:1 (40%):	38.6% (w/v)	Acrylamide
	1.4% (w/v)	Bis-acrylamide
Lower Tris:	1.5 M	Tris/HCl, pH 8.8
	0.4% (w/v)	SDS
Upper Tris:	0.5 M	Tris/HCl, pH 6.8
	0.4% (w/v)	SDS

2 x SDS Loading Buffer	4.8% (w/v)	SDS
	20% (v/v)	Glycerin
	10% (v/v)	β-Mercapto-ethanol
	0.1% (w/v)	Bromphenol blue
	100 mM	Tris/HCl, pH 7.8

2.2.13.4 Staining of PAA Gels

The PAA gels were stained with Coomassie Brilliant Blue 250. After the electrophoresis, the gel was stained overnight. To detect clearly the protein bands, the gel was incubated in a destaining solution overnight under rocking. Alternatively, the „Simply Blue™ SafeStain"(Invitrogen, Groningen) solution was used to dye the protein PAA gel. The PAA gel was rinsed 3x for 5 min with dH_2O. Then, it was covered with SimplyBlue™ solution and stained for 1 h at RT under gentle shaking. The staining solution was discarded and the gel was washed with dH_2O for 1 h. A second water wash was required for a clear gel background.

Staining solution:	30 % (v/v)	Methanol
	10% (v/v)	Acetic acid
	0.2% (w/v)	Coomassie Brilliant Blue R250
Destaining solution:	30% (v/v)	Methanol
	10% (v/v)	Acetic acid

2.2.13.5 Semi Dry Blot

Proteins were separated by SDS PAGE and transferred to a membrane. For this, 6 pieces of GB004 Whatman filter paper were cut at the size of the gel. For the blot, the graphite plates from the transfer equipment were cleaned with dH_2O. Two pieces of filter paper were

soaked in anode I buffer, 1 piece was soaked in anode II buffer and the paper pieces were piled on the transfer plate. The membrane was moistened with methanol and laid on the filter paper pile. The protein gel was taken out from the glass plates, the collecting gel was separated and eliminated; meanwhile the separating gel was carefully laid on the membrane. Finally, 3 pieces of filter paper were soaked in the cathode buffer and laid on the gel. Air bubbles were eliminated with a roller. After the cathode plate was put in place, the transfer was done for 1 h at 3.5 mA/cm^2. The membrane was subjected to immune detection.

Anode I buffer	0.3 M	Tris/HCl, pH 10.4
	20% (v/v)	Methanol
Anode II buffer	25 mM	Tris/HCl, pH 10.4
	20% (v/v)	Methanol
Cathode buffer	40 mM	ε-aminocaproic acid
	25 mM	Tris/HCl, pH 9.4
	20% (v/v)	Methanol

2.2.13.6 Immune Detection on Protein Filters

The unspecific binding sites present in a membrane were saturated by using blocking buffer for 30-60 min. Then, the filter was incubated in a plastic bag with 1 ml/10 cm^2 of antiserum (diluted 1:500 in washing buffer) for 1 h at RT or overnight at 4°C. Unbound antibodies were removed by using washing buffer 3x 20 min. The second antibody coupled with alkaline phosphatase was diluted 1:10000 in washing buffer and added to the filter for 1 h under swinging shaking. Later, unbound antibodies were removed with washing buffer 4 times for 10 min each. The filter was then 2 times washed with buffer 3 for 5 min and later incubated in a plastic bag with chromogenic solution, containing 45 µl of NBT solution and 35 µl of BCIP solution in 10 ml of Buffer 3. Then, the filter was washed shortly in dH_2O and dried on filter paper.

10x Washing stock	1.4 M	NaCl
	100 mM	Tris/HCl, pH 7.5
	0.5% (v/v)	Tween 20
Blocking buffer	5% (w/v)	non fat dry milk in 1x washing stock
Washing buffer	2% (w/v)	non fat dry milk in 1x washing stock
Buffer 3	100 mM	NaCl
	50 mM	$MgCl_2$
	100 mM	Tris/HCl pH 9.5
NBT solution	75 mg/ml	NBT
	70%	Dimethylformamide (DMF)
BCIP solution	50 mg/ml	BCIP in DMF

2.2.14 Generation of Polyclonal Antibody

2.2.14.1 Production of a GST-Tag Fusion Protein

A fusion protein with 220 aa GST-tag (at the N-terminal) and His-Tag (at the C-terminal) was produced using the pET system (Novagen, Darmstadt). The pET 41a (+) expression vector carry a strong bacteriophage T7 transcription signal; expression can be induced by providing source of T7 RNA polymerase in the host cell (BL21 DE3). T7 RNA polymerase is very selective and active when fully induced by adding IPTG at a right concentration.

One strategy was to use a modified PCR product that contains the *Bam HI* and *PstI* restriction sites at 5´and 3´ends respectively. This PCR fragment was then restricted with both enzymes, ligated into the *Bam HI* and *PstI* restricted vector and the resulting ligation reaction was transformed into BL21- DE3 competent cells. A positive clone was selected

and a single colony was pre-cultured in 3 ml of LB Medium with kanamycin overnight at 37°C under shaking (200 rpm). 2 ml of this pre-culture were used to inoculate 100 ml of LB Medium-kanamycin which was incubated at 37°C under shaking (200 rpm) until an OD_{550}= 5.5 was reached. 3 mM IPTG solution was added to the culture from 100 mM stock for induction of the fusion protein and incubation continued for another 3 h. After that, cells were harvested by centrifugation at 4,500x g, 4°C for 12 min. The supernatant was discarded, (carefully pipetting the remaining medium) and the pellet was resuspended in 20 mM Tris/HCl (pH 8.0) solution. For the analysis of the total protein content via SDS-PAGE, 10 µl were removed from this suspension. The suspension was sonicated under ice-cooling until lysis was complete. To remove insoluble components, the suspension was centrifuged at 14,000xg, 4°C for 15 min. The clear supernatant was carefully transferred to a clean tube. This extract was ready for affinity purification of the recombinant protein (GST-Bind Kits Novagen, Darmsdadt). 10 µl aliquots were separated in a SDS-PAA gel to check the protein size. After this, the protein extract was used to immunize two New Zealand rabbits and later on to isolate the *Gtl8* antibody present in the antiserum of these two rabbits. The purified antibody was used for immunohistological studies.

2.2.14.2 Purification of Monospecific Antibodies

The antibodies against the protein of interest present in the polyclonal serum have to be purified so that protein analysis can be performed. GST and His-Tag fusion proteins were used for this purpose. The recombinant protein was separated by SDS PAA gel and was transferred to a membrane (2.2.13.5.). After the transfer, a stripe from the edge of the filter was cut out and fusion protein was detected with anti-His antibody, to localize the fusion protein. Then, the part of the filter containing the fusion protein was cut out, blocked for 1h and incubated for 1 h with 200 µl of the polyclonal antiserum. The membrane was washed 3 times for 10 min with washing buffer and 2 times for 5 min with PBS. The bound monospecific antibodies were eluted from the membrane with 1-5 ml of elution buffer under vortexing for 15 min. After a washing step with PBS for 5 min, the membrane was blocked again for a minimum of 15 min and the elution procedure was repeated 3 times more. The eluted antibodies were transferred to a Centrisart tube (Sartorius AG, Göttingen)

and concentrated by centrifugation at 2,500x g. The supernatant in the inner tube was

removed and the outer tube was refilled with PBS and centrifuged again. These steps were repeated until the KSCN was completely removed from the antibody solution (3-4 times). The monospecific antibody solution was stored at 4°C.

Elution buffer:	3 M	KSCN
	0.1% (w/v)	BSA in 1x PBS

2.2.15 X-gal Staining

2.2.15.1 X-gal Staining of Preimplantation Embryos

A histochemical staining procedure for *E. coli* β-galactosidase activity in preimplantation embryos was used to detect Lac Z expression. Embryos were rinsed twice with PBS (pH 7.2) and then fixed for 10 min at 4^0C with 0.25% glutaraldehyde in PBS. After four times washing in PBS, the embryos were incubated in staining solution consisting of 0.04% 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal; Sigma Chemical) 1mM $MgCl_2$, 10 mM potassium ferricyanide and 10 mM potassium ferrocyanide in PBS. Incubation was carried out for 12h at 37^0C in air. Stained embryos could be analysed and photographed using light microscopy after the replacement of embryos on the glass slide.

2.2.15.2 X-gal Staining of Adult Mouse Tissues

A histochemical staining procedure for *E. coli* β-galactosidase activity in adult mouse tissue was used to detect Lac Z expression. Tissues were rinsed in cold PBS and fixed in FixA solution for 30 min on ice with shaking. Than the tissues were washed again in PBS and fixed in FixB solution for 1h on ice shaking. After final washing with PBS, tissues were incubated overnight in X-gal staining solution at 30^0C under dark condition.

FixA	1% v/v formaldehyde 0.2% v/v glutaraldehyde 0.02% v/v NP-40 in PBS
FixB	1% v/v formaldehyde 0.2% v/v glutaraldehyde 0.2% v/v NP-40 0.1% w/v sodium deoxycholate in PBS
X-gal staining solution	1 mg/ml X-gal 5 mM potassium ferricyanide 5 mM potassium ferrocyanide 2 mM $MgCl_2$ in PBS

2.2.16 Immunofluorescence Staining of Mouse Preimplantation Embryos

(Moreno et al, 2002)

For immunocytochemistry the zona pellucida of mouse oocyte**s** or preimplantation embryos was removed by a short incubation in acid Tyrode's medium, and the oocyte**s** were gently attached to SuperFrost slide. Fixation was carried out by adding 2% formaldehyde, followed by a 1- to 2-h incubation at 4^0C. The samples were then permeabilized for 60 min in PBS containing 1% Triton X-100 and nonspecific reactions were blocked by further incubation in PBS containing 2 mg/ml BSA and 100 mM glycine. An appropriate primary

antibody was solubilized in this blocking solution and incubated with the slides for 1–2 h. After extensive washing in PBS containing 0.1% Triton X-100, the samples were sequentially labelled with indocarbocyanine (Cy3) or fluorescence isothiocyanate (FITC)-conjugated secondary antibodies. The samples were then washed 4 times with PBS containing 0.1% Triton X-100; the nuclei were counterstained with DAPI (Vector). Immunostaining of the sections was examined using a fluorescence-equipped microscope (BX60; Olympus).

2.2.17 Immunofluorescence Staining of Mouse Fibroblast Cells

The cells were grown over night on the coverslips in DMEM medium (2.1.5.3.) and fixed in ice cold methanol/aceton (ratio 1:1) solution for 10 min at RT. Samples were rinsed in PBS and an initial blocking step was performed with the blocking solution (PBS containing 5% goat serum, 3% BSA and 1x Roti block solution) for 1 h. An appropriate primary antibody solution in PBS was applied for overnight in a humidified chamber at 4°C. After four washes with PBS/0.2% Tween-20, the coverslips were incubated with FITC- or Cy3-conjugated goat anti-rabbit or goat anti-mouse IgG (Sigma) for 1 hr at RT. One drop of mounting medium with DAPI was dispensed onto the slides after washing with PBS/0.2% Tween-20. Fluorescent cells were visualised with Olympus BX60 microscope using 20X or 60X neofluor lens, photographed using digital camera and analysed using analysis 3.0 soft imaging system.

2.2.18 Fluorescence in Situ Hybridization (FISH)

Chromosomes were prepared from mouse peripheral blood lymphocytes. A mouse genomic cosmid clone containing the mouse *gtl8* gene was labelled by standard nick translation with biotin-16-dNTP (Boehringer Mannheim) and used as a probe for hybridization. Hybridized probes were detected by FITC-conjugated avidin (Lichter et al., 1988). Chromosomes were counterstained with 4,6-diamidino-2-phenylindole dihydrochloride (DAPI; Vector). Images of emitted light were captured separately by using the DAPI and FITC filter set and subsequently merged and aligned.

2.2.19 Generation of a Green Fluorescent Fusion Protein (GFP)

The green fluorescent protein (GFP) is a very useful tool to perform sub-cellular localisation of proteins and to observe their expression, due to the green fluorescent light that emits at 507 nm after being excited at 488 nm. A fusion protein with the enhanced green fluorescent protein (EGFP) was produced by cloning in frame the entire or truncated coding regions of the protein of interest in C-terminal position with respect to the EGFP sequence. The pEGFP-C1 vector (Clontech, Heidelberg) carries a strong promoter (pCMV) and the SV40 poly A, which direct proper processing of the 3´end of the EGFP mRNA. The sequence of interest was cloned in the MCS using the restriction sites of *SalI* and *ApaI*. This ligation reaction (2.2.2.2.) was transformed in DH5α competent cells (2.2.4.). Then, a positive clone was selected and a single colony was cultured to obtain enough quantities of DNA (2.2.1.2.). After checking the sequence subcloned in the expression vector by restriction analysis and sequencing (2.2.8.), the DNA was used for transfection.

2.2.20 Transfection of NIH 3T3 Cells with the *GFP-fusion* Construct

Approximately 4 x 10^5 fibroblast cells (NIH 3T3) were plated in a cell chamber (with slide) (Lab-Tek, Wiesbaden) and cultured overnight in 5 ml DMEM medium containing 10% FCS and penicillin/streptomycin at 37°C and 5% CO_2. 1 µg of DNA (*Gfp-fusion*) was diluted with the DNA-condensation buffer (buffer EC of the Effection Transfection kit, Qiagen, Hilden) to a total volume of 150 µl, 8 µl of enhancer was also added to DNA and mixed by vortexing for 1 sec. The mixture was incubated at RT for 5 min. 25 µl of Effectin Transfection Reagent was added to the DNA-Enhancer mixture, mixed by pipetting and incubated for 10 min at RT to allow complex formation. 1 ml of cell medium was added to the reaction tube, mixed and immediately added into the cell chamber. Cells were incubated for overnight at 37°C, 5% CO_2 to allow for gene expression. The glass slide of the culture chamber was rinsed with PBS and then fixed with methanol for 5 min. One drop of mounting medium with DAPI was dispensed onto the slide. Fluorescent cells were visualised with Olympus BX60 microscope using a 60X neofluor lens, photographed using digital camera and analysed using software 3.0, Soft Imaging System.

2.2.21 Yeast Two Hybrid Assay

The Gal4-based MATCHMAKER two-hybrid system III of Clontech laboratories, Inc. was followed for the yeast two hybrid assays. Plasmid vectors, pGBKT7 and pGADT7 encoding the Gal4 binding domain (Gal4-BD) and Gal4 activating domain (Gal4-AD) respectively, were used to express hybrid proteins. To screen the proteins that interact with *Gtl8* in yeast two-hybrid system, a pretransformed mouse testis library in the Gal4-AD vector was screened using the full-length *Gtl8* (80 kD) or the FERM domain of *Ggtl8* cloned into the Gal4-BD vector as a bait. All positive yeast clones were selected for β-galactosidase activity on SD/-Ade/-His/-Leu/-Trp/X-α-gal plate. DNA from positive clones were isolated and sequenced. Sequence informations were blasted to identify the putative interaction partners, using public databases including NCBI.

2.2.22 Techniques for Recovery and Culture of Preimplantation Embryos

2.2.22.1 Superovulation

Seven to eight week old female mice were superovulated by intraperitoneal injections of 5 IU of pregnant mare's serum gonadotropin (PMSG, Sigma St Louis, MO, USA) followed 44-48 hrs later by 5 IU of human chorionic gonadotrophin (HCG; Sigma). After the second injection, females were housed overnight with males and were checked by a vaginal plug the following morning. The E0.5 was considered to be 12:00 noon at the day of vaginal plug.

2.2.22.2 Recovery of Preimplantation Embryos

1-cell to 8-cell stage embryos were flushed from the oviduct of superovulated females. Plugged female mice (1.5-2.5 days pc) 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 M2 medium. Under dissection microscope, the 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

flashed with 0.05 ml of M2 medium (2.1.5.2.). Embryos were collected with a pipette and washed through several M2 drops. To remove the zona pellucida, embryos were treated in one drop of Tyrode's acid solution and transferred into the drop of M2-medium as soon as their zona pellucida was dissolved. 2 cell and 4 cell stage embryos were incubated in the drop of trypsin (Trypsin/EDTA, PAN Biotech) in order to remove the polar bodies. The collected embryos were washed five times in a drop of PBS and each single embryo was transferred into a PCR cup (0.2 ml) containing 5 µl dH_2O. Those samples were used for genotyping. In order to decrease the possible contamination of maternal cells, 2 cell stage embryos were cultivated in M16 medium at 37^0C and 5% CO_2 until they developed to expected developmental stages as a 4-cell to blastocyst.

2.2.22.3 In vitro Culture of Blastocyst Stage Embryos

Embryos were flushed out from the uteri of plugged females at day 3.5 and placed in gelatinized 96-well dish. Embryos were cultured in ES cell medium without the addition of leukaemia inhibitory factor. The dish was kept in a humid incubator at 37^oC, 5% CO_2 and embryonic outgrowth was scored daily. Every second day, 5 µl of fresh medium was added to the wells. DNA was prepared on the fourth day of the culture (2.2.1.4).

2.2.23 Computer Analysis

For the analysis of the nucleotide sequences, programs like BLAST, BLAST2, MEGABLAST and other programs from National Center for Biotechnology Information (NCBI) were used (www.ncbi.nlm.nih.gov). Information about mouse alleles, phenotypes and strains were used from Jackson Laboratory (www.informatics.jax.org). For proteins studies ExPASy tools (www.expasy.ch) were used. Mouse genome sequence and other analysis on mouse genes, transcript and putative proteins were downloaded from Celera discovery system (www.celera.com).

3. RESULTS

The gene trap approach in embryonic stem (ES) cells offers three features in one experimental approach: 1) analysis of the expression patterns of trapped genes by using a simple staining method, 2) rapid cloning of the trapped genes, and 3) generation of mutant mouse lines.

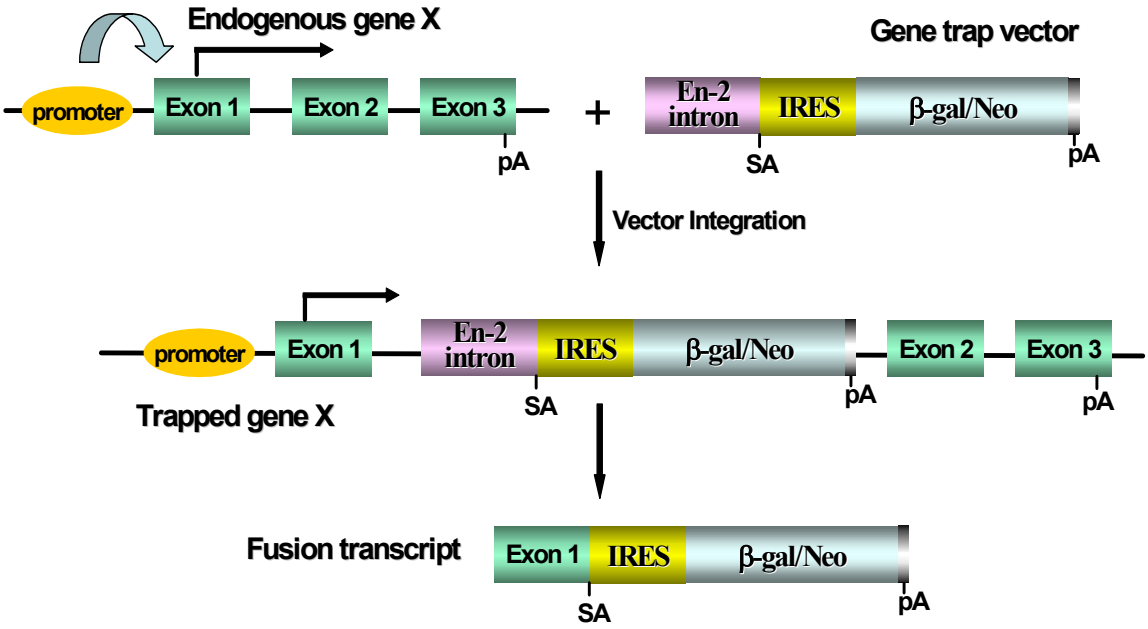


Figure 3.1: Schematic view of gene trap event in embryonic stem cells

Generally a gene trap (GT) vector contains the splice acceptor (SA) sequence from the mouse En-2 gene joined in frame with promoterless β-galactosidase (β-gal) gene for reporter and neomycin phosphotransferase (neo) gene as a selector, followed by the SV40 polyadenylation signal (pA) at the 3'-end. In this vector, the internal ribosomal entry site (IRES) from the encephalomyocarditis virus located between SA sequence and the reporter sequence allows the cap-independent translation of the reporter gene from fusion transcript. In most cases, the gene trap vector integrates into an intron of an endogenous gene, the reporter (in this case β-gal) will be expressed under the transcriptional control of the

"trapped" gene promoter. Thus, upon transcriptional activation of the trapped gene, a fusion transcript will be generated between the upstream exons and the reporter gene. In addition to faithfully depicting the expression pattern of the trapped gene, the fusion transcript serves to disrupt gene function by generating a truncated mRNA. Therefore, the function of trapped gene can be studied via phenotypical analysis.

To identify the genes involved in male germ cell development in mice, several gene trap mouse lines were obtained from the research group of Professor Peter Gruss of the Max Planck Institute for Biophysical Chemistry, Göttingen. The scheme for generating gene trap mouse is shown in Fig 3.2.

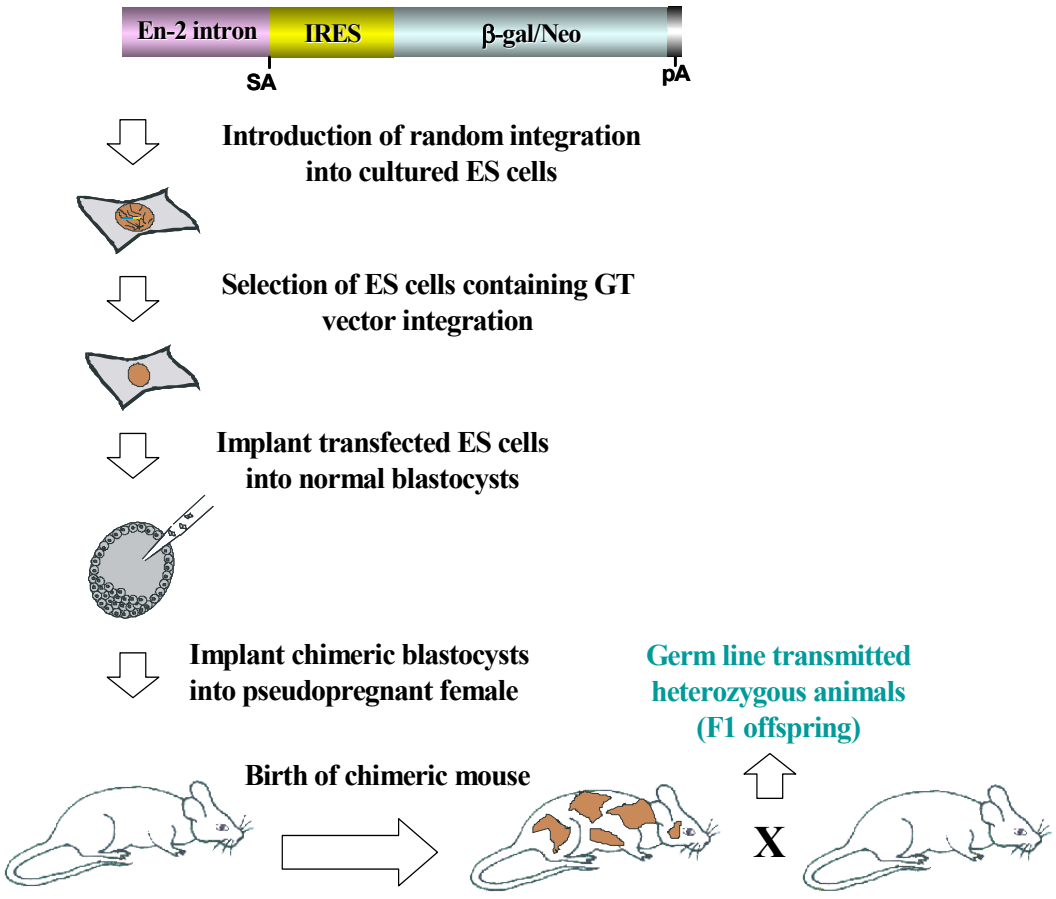


Figure 3.2: Generation of gene trap mouse line

Initially, the reporter gene expression was tested in testis tissue of heterozygous (F1) animals in order to find out whether the trapped genes are involved in the male germ cell development. Two gene trap mouse lines, namely line 8 and line 16 were selected for further investigation because of the β-galactosidase expression pattern was detected strongly in the testis. However, further analysis in these lines revealed that homozygous animal die before birth. Therefore, trapped genes in these mouse lines have been suggested to play crucial role for embryonic development.

3.1 Gene Trap Line 8

3.1.1 Isolation and Characterization of Murine *Gtl8* cDNA

3.1.1.1 Reporter Gene Expression

The in vivo expression pattern of the trapped gene in adult mice, β-galactosidase activity was examined in sections of several organs dissected from heterozygous gene trap line 8 animals. Lac-Z expression was ubiquitous in all organs tested, including testis (Fig 3.3), kidney, brain, ovary, heart (data not shown). The Lac-Z expression was strongly detected in spermatogonia and spermatids in testis section (Fig. 3.3).

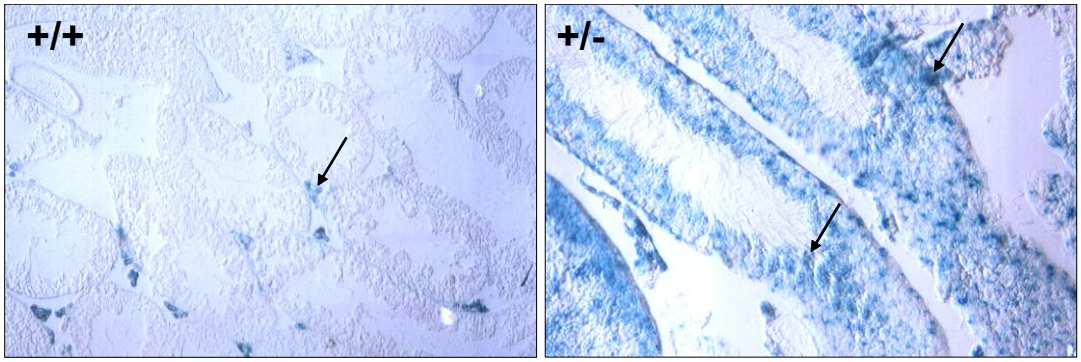


Figure 3.3: X-gal staining of adult mouse testis. Light microscopic image of testis tissue section from wild type and heterozygous animals after staining with X-gal and paraffin embedding. β-galactosidase activity was detected strongly in spermatogonia and spermatids (arrows) in testis. Unspecific staining was detected in Leydig cells of testis from wild type animal (arrow).

3.1.1.2 Identification of Trapped Gene in Gene Trap Mouse Line 8

The fusion transcript between endogenous gene and GT vector (Fig. 3.1) can be cloned by RACE (Rapid Amplification of cDNA Ends) and sequenced to determine the identity of the trapped gene. For this purpose, the primers SH1, SH2, and SH3 (sequences shown in 2.2.5.3) were designed on IRES sequence (Fig. 3.1) of GT vector to amplify upstream

nucleotides. Total RNA was extracted from adult testis tissue of heterozygous animal and used for 5'RACE PCR. A 400 bp fragment was amplified and cloned into pGEM-T Easy vector. Sequencing of the 5'RACE product revealed that fusion of the IRES-β-gal/neo occurred with a 288 bp unknown sequence (Fig 3.4). The 288 bp sequences were aligned with previously reported EST and nucleotide sequences in the database (BLAST, NCBI). Two overlapping mouse EST clones Soares-mammary-gland-NbMMG (gi|2691321) and B spermatogonia Mus musculus cDNA clone (gi|12595467), were exhibited 100% homology to the 288 bp sequence. The total size of these two cDNAs was 1.2 kb. Predicted amino acid sequence of the 1.2 kb cDNA fragment showed only part of the coding region and no putative ATG (start codon), 5'UTR and 3'UTR. Therefore, the isolation of additional cDNA sequences was performed.

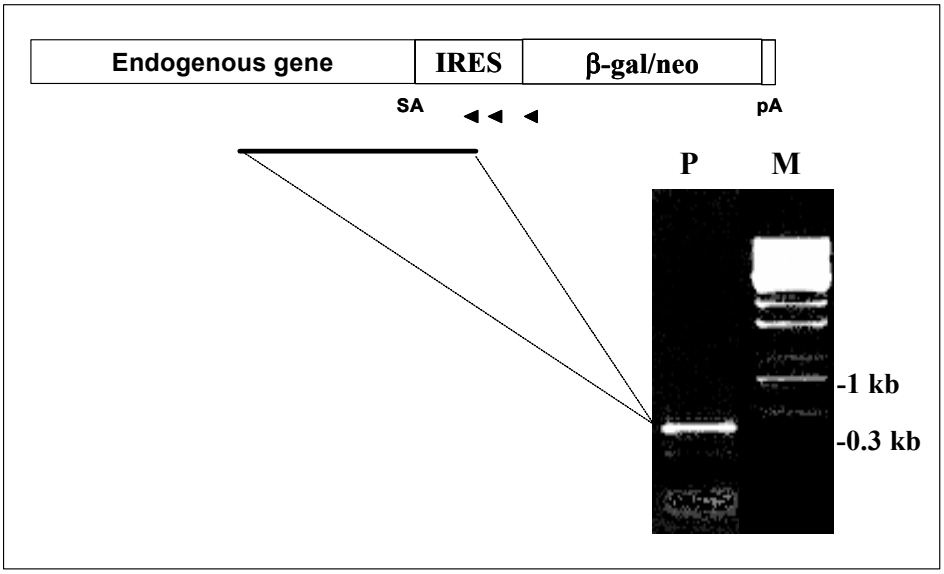


Figure 3.4: 5'RACE PCR analysis of Gene trap Mouse Line 8. Result of 5'RACE PCR experiment using three nested primers (arrowheads) designed on IRES and β-gal sequences of GT vector and UAP, AUAP (described in 2.2.5.3) primers from 5'RACE kit. Testicular RNA from heterozygous animal was used as a template for this PCR. About 400 bp fragment (**P**) was amplified. **M** is the marker for determination of the DNA fragment size.

3.1.1.3 Isolation and Cloning of Complete cDNA

In accordance with the name of the gene trap line, the trapped gene and its cDNA was called *Gtl8*. To obtain complete *Gtl8* cDNA, the mouse testis MATCHMAKER cDNA library (Clontech, Heidelberg) was used. Actually this library was designed for Yeast Two Hybrid assay. But the library contains pooled cDNAs from 8-12 weeks old mouse testes. PCR experiments were performed with MATCHMAKER testis cDNA library using *Gtl8* gene specific (F and R primers shown in Fig. 3.5) and library specific primers. Combinations of primers amplified 4 different cDNA fragments with length of 2.5 kb, 1.6 kb, 1.2 kb and two similar 0.5 kb bands. All fragments were subcloned into pGEM-T Easy vector and sequenced.

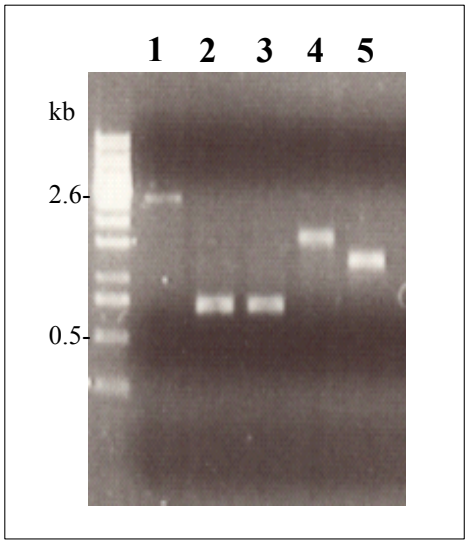


Figure 3.5: Cloning of complete *Gtl8* cDNA. Results of PCR experiment using mouse testis cDNA library. *Gtl8* sequence specific primers (F and R) and primers from Match Maker library were combined and amplified different parts of *Gtl8* cDNA fragments. All fragments were sub-cloned into pGEM-T Easy cloning vector and sequenced subsequently. The scheme for localization of primers and amplified fragments used in this study are shown in Figure 3.6.

The composed cDNA contained a 5'-untranslated region (5'-UTR) of 266 bp, an open reading frame (ORF) of 2196 bp and a 3'-UTR of 998 bp containing putative polyadenylation signal at nucleotide position 3380-3385 (AACAAA) (Fig. 3.7). Complete cDNA sequence was aligned with mouse genomic database (BLAST, NCBI) and the 25 exons were predicted.

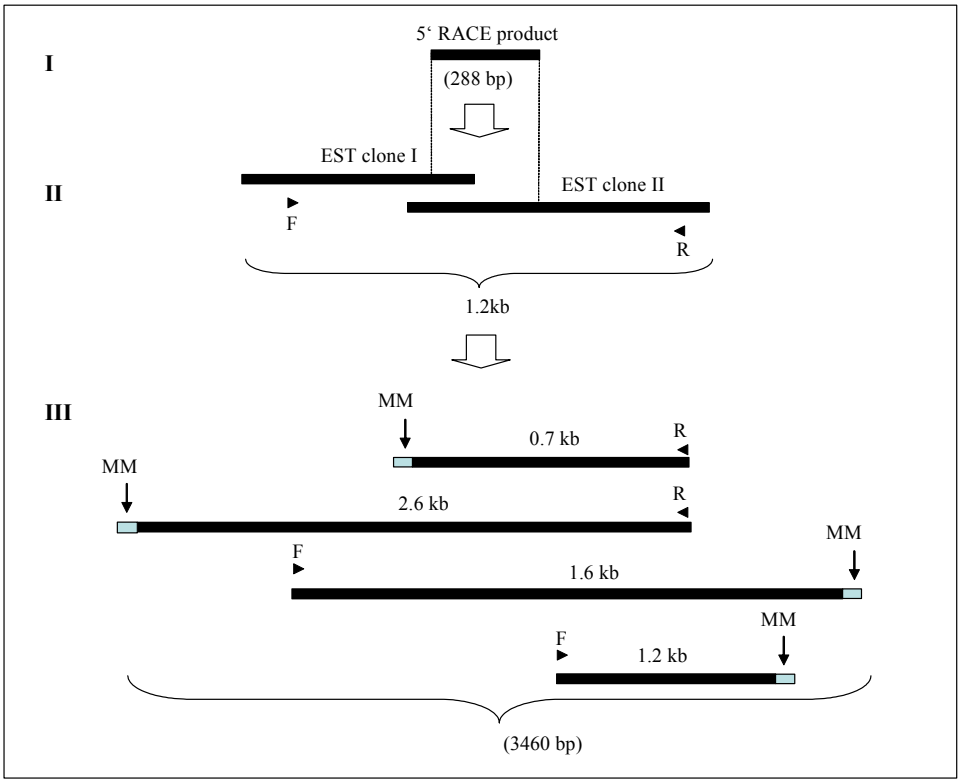
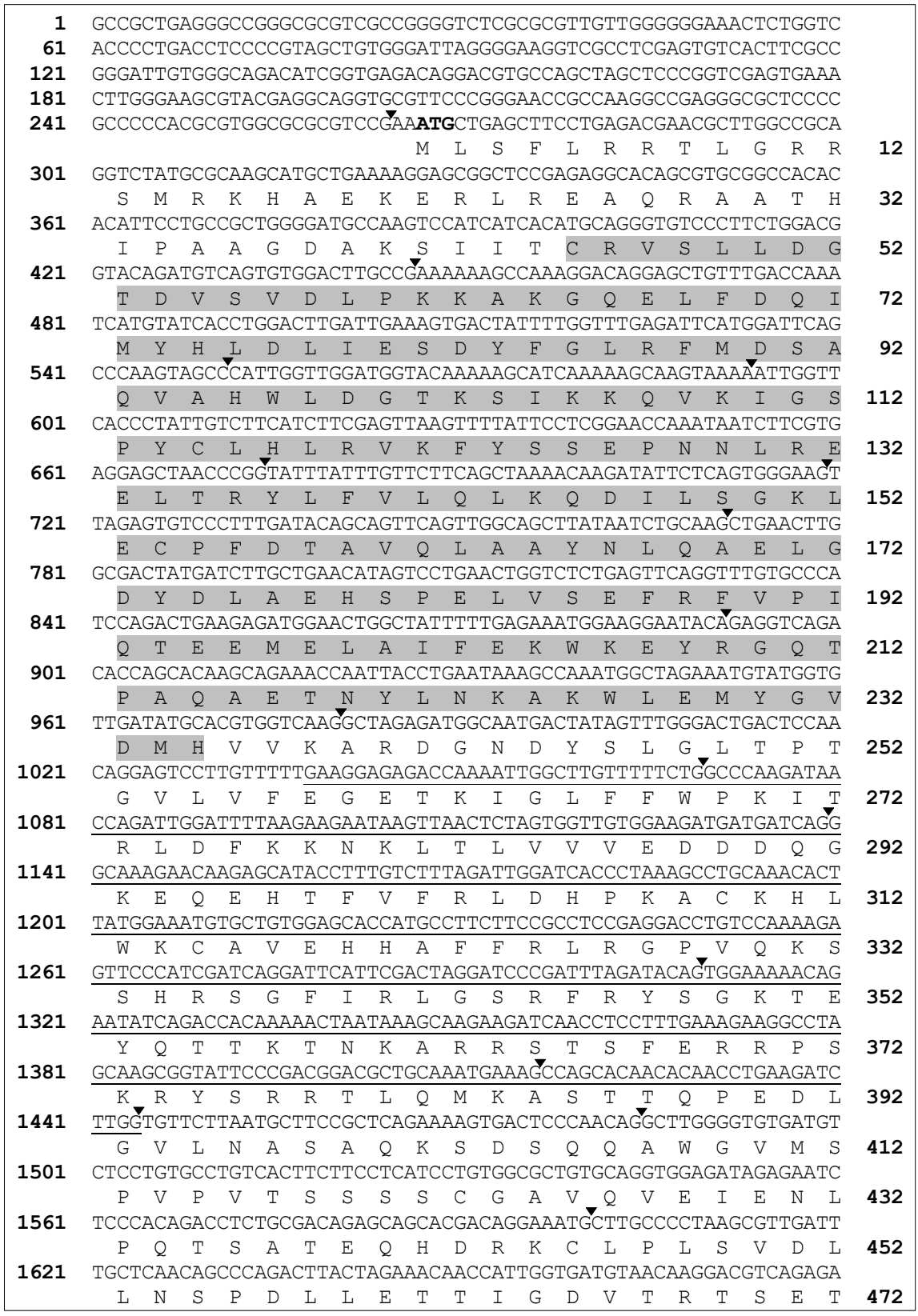


Figure 3.6: Schematic diagram of identification and cloning of the *Gtl8* cDNA. I. 288 bp sequence of the *Gtl8* cDNA was initially identified using 5'RACE PCR. II. Homologous EST clones were identified from database. III. PCR amplified fragments using combinations of *Gtl8* and vector specific primers. MM stands for vector specific primers.


1 GCCGCTGAGGGCCGGGCGCGTCGCCGGGGTCTCGCGCGTTGTTGGGGGGAAACTCTGGTC
61 ACCCCTGACCTCCCCGTAGCTGTGGGATTAGGGGAAGGTCGCCTCGAGTGTCACTTCGCC
121 GGGATTGTGGGCAGACATCGGTGAGACAGGACGTGCCAGCTAGCTCCCGGTCGAGTGAAA
181 CTTGGGAAGCGTACGAGGCAGGTGCGTTCCCGGGAACCGCCAAGGCCGAGGGCGCTCCCC
241 GCCCCCACGCGTGGCGCGCGTCCGAAATGCTGAGCTTCCTGAGACGAACGCTTGGCCGCA
M L S F L R R T L G R R 12
301 GGTCTATGCGCAAGCATGCTGAAAAGGAGCGGCTCCGAGAGGCACAGCGTGCGGCCACAC
S M R K H A E K E R L R E A Q R A A T H 32
361 ACATTCCTGCCGCTGGGGATGCCAAGTCCATCATCACATGCAGGGTGTCCCTTCTGGACG
I P A A G D A K S I I T C R V S L L D G 52
421 GTACAGATGTCAGTGTGGACTTGCCGAAAAAAGCCAAAGGACAGGAGCTGTTTGACCAAA
T D V S V D L P K K A K G Q E L F D Q I 72
481 TCATGTATCACCTGGACTTGATTGAAAGTGACTATTTTGGTTTGAGATTCATGGATTCAG
M Y H L D L I E S D Y F G L R F M D S A 92
541 CCCAAGTAGCCCATTGGTTGGATGGTACAAAAAGCATCAAAAAGCAAGTAAAAATTGGTT
Q V A H W L D G T K S I K K Q V K I G S 112
601 CACCCTATTGTCTTCATCTTCGAGTTAAGTTTTATTCCTCGGAACCAAATAATCTTCGTG
P Y C L H L R V K F Y S S E P N N L R E 132
661 AGGAGCTAACCCGGTATTTATTTGTTCTTCAGCTAAAACAAGATATTCTCAGTGGGAAGT
E L T R Y L F V L Q L K Q D I L S G K L 152
721 TAGAGTGTCCCTTTGATACAGCAGTTCAGTTGGCAGCTTATAATCTGCAAGCTGAACTTG
E C P F D T A V Q L A A Y N L Q A E L G 172
781 GCGACTATGATCTTGCTGAACATAGTCCTGAACTGGTCTCTGAGTTCAGGTTTGTGCCCA
D Y D L A E H S P E L V S E F R F V P I 192
841 TCCAGACTGAAGAGATGGAACTGGCTATTTTTGAGAAATGGAAGGAATACAGAGGTCAGA
Q T E E M E L A I F E K W K E Y R G Q T 212
901 CACCAGCACAAGCAGAAACCAATTACCTGAATAAAGCCAAATGGCTAGAAATGTATGGTG
P A Q A E T N Y L N K A K W L E M Y G V 232
961 TTGATATGCACGTGGTCAAGGCTAGAGATGGCAATGACTATAGTTTGGGACTGACTCCAA
D M H V V K A R D G N D Y S L G L T P T 252
1021 CAGGAGTCCTTGTTTTTGAAGGAGAGACCAAAATTGGCTTGTTTTTCTGGCCCAAGATAA
G V L V F E G E T K I G L F F W P K I T 272
1081 CCAGATTGGATTTTAAGAAGAATAAGTTAACTCTAGTGGTTGTGGAAGATGATGATCAGG
R L D F K K N K L T L V V V E D D D Q G 292
1141 GCAAAGAACAAGAGCATACCTTTGTCTTTAGATTGGATCACCCTAAAGCCTGCAAACACT
K E Q E H T F V F R L D H P K A C K H L 312
1201 TATGGAAATGTGCTGTGGAGCACCATGCCTTCTTCCGCCTCCGAGGACCTGTCCAAAAGA
W K C A V E H H A F F R L R G P V Q K S 332
1261 GTTCCCATCGATCAGGATTCATTCGACTAGGATCCCGATTTAGATACAGTGGAAAAACAG
S H R S G F I R L G S R F R Y S G K T E 352
1321 AATATCAGACCACAAAAACTAATAAAGCAAGAAGATCAACCTCCTTTGAAAGAAGGCCTA
Y Q T T K T N K A R R S T S F E R R P S 372
1381 GCAAGCGGTATTCCCGACGGACGCTGCAAATGAAAGCCAGCACAACACAACCTGAAGATC
K R Y S R R T L Q M K A S T T Q P E D L 392
1441 TTGGTGTTCTTAATGCTTCCGCTCAGAAAAGTGACTCCCAACAGGCTTGGGGTGTGATGT
G V L N A S A Q K S D S Q Q A W G V M S 412
1501 CTCCTGTGCCTGTCACTTCTTCCTCATCCTGTGGCGCTGTGCAGGTGGAGATAGAGAATC
P V P V T S S S S C G A V Q V E I E N L 432
1561 TCCCACAGACCTCTGCGACAGAGCAGCACGACAGGAAATGCTTGCCCCTAAGCGTTGATT
P Q T S A T E Q H D R K C L P L S V D L 452
1621 TGCTCAACAGCCCAGACTTACTAGAAACAACCATTGGTGATGTAACAAGGACGTCAGAGA
L N S P D L L E T T I G D V T R T S E T 472

```
1681  CTTCGGCACCATTCCCAGCACCAGATACCATTAATGTTGCCACCAGGTCAAATGAATTAG
        S  A  P  F  P  A  P  D  T  I  N  V  A  T  R  S  N  E  L  E   492
1741  AGGAATTCAAAGCTGAATGTGAGACATTAAAGGATGACACAGAGAAGCTTAAACAGCTTG
        E  F  K  A  E  C  E  T  L  K  D  D  T  E  K  L  K  Q  L  E   512
1801  AAACAGAGCAGACTATCTTGCCATCTCTTCGACCTACCATTGACATTAATGTCAACAGCC
        T  E  Q  T  I  L  P  S  L  R  P  T  I  D  I  N  V  N  S  Q   532
1861  AGGAAGAAGTGGTAAAGTTGACTGAGAAATGTCTTAATAATGCCATTGAGAACCCAGCAT
        E  E  V  V  K  L  T  E  K  C  L  N  N  A  I  E  N  P  A  L   552
1921  TGAATGCAGTAAAGGTTCCTCCTGACTTCAAGAGTAATATTTTGAAAGCTCAAGTAGAAG
        N  A  V  K  V  P  P  D  F  K  S  N  I  L  K  A  Q  V  E  A   572
1981  CTGTGCATAAGGTTACAAGAGAAGATAGTTTATTAACTCATAAAAATGCCAGCGTTCAGG
        V  H  K  V  T  R  E  D  S  L  L  T  H  K  N  A  S  V  Q  D   592
2041  ATGCTGCCACAAACAGTACTGCATTCAATGAAAATGATGTGCCCGTCTGTAAAGATTCCC
        A  A  T  N  S  T  A  F  N  E  N  D  V  P  V  C  K  D  S  L   612
2101  TCACTCCGGTGCATGGCACAGCTGCAGACAGTGCCTCTGTACTAAAGGATGCTACTGATG
        T  P  V  H  G  T  A  A  D  S  A  S  V  L  K  D  A  T  D  E   632
2161  AACTGGACGCCTTACTTCTTTCTCTGACAGAGAATCTGATGGACCACACAGTAACACCTC
        L  D  A  L  L  L  S  L  T  E  N  L  M  D  H  T  V  T  P  Q   652
2221  AGGTGTCTTCACCATCCATGATCACACCCCGGTGGATTATTCCACAGAGTGCTACCATCT
        V  S  S  P  S  M  I  T  P  R  W  I  I  P  Q  S  A  T  I  S   672
2281  CCAATGGGCTTGCAGGATATGGAGCATCTTTGGCAGGGACAGATGAATGTAGTCAGAAAG
        N  G  L  A  G  Y  G  A  S  L  A  G  T  D  E  C  S  Q  K  D   692
2341  ATGGATTCTCACTGATTTCCCCCCCAGCACCATTCTTGGTAGATGCTGTGACCAGCTCTG
        G  F  S  L  I  S  P  P  A  P  F  L  V  D  A  V  T  S  S  A   712
2401  CCCCTCCCTTGCCTGAAGACTCAACCTTGAAGCAGAAGTGCTTGCTGACTACTGAGCTCT
        P  P  L  P  E  D  S  T  L  K  Q  K  C  L  L  T  T  E  L      731
2461  GAGGGCTGCTGCTCCCCACCTGTAACTGGATTGCCCCACACCGTCCCCAGTCAGCCCTGG
2521  ATCCATGTGTCCTTTACTTGGGGAAAACCTTCTACCAGGATTTGACTTCAACTCCATAAT
2581  GAGAGTAGCTACATTTTCCGTCCAACTGGAAGTCTACCCCATATAACTGCTTAACTGAAG
2641  ATACCTCCTCCTCCTCCTGTAACTGTATTTGCATTGCCAATTCCTGAAGCAATGAAGGTA
2701  ACTAGTTTCCTGCAAACTGATGGCTCCACCTAGCGGTTGCAACTGTTTATAGCTTAGAGC
2761  AAGGCTTGCTGTCCAGAGCCTTTCATTGCTGTTTCCCTGGGTTGTCAGGGGAGCACACCT
2821  GAACTGAGAGCCTGCTGGAGAAATGCGTGGCACTTTCCTGTGCATTCCCTGGAAATGAGA
2881  ATTAGAAAATGGATGGAAACACTATGTGGCTAGATCCCAGGCCAGCTCCTGCACAGTTGC
2941  TAGGATTCTTGGAGAAGCTGAGAACATAGAGGGGTGAGGCTGTCACCCAGACTTTGAATT
3001  CTGAGCAGTAATGACCTTACATCAAAGGTAAATGTTATCTTTGGATAACTAGAGAGAGGT
3061  TCAGTCTCTAGGCGTGAACTCCTCAGACCACTTTCCTATTGGAAGCACCTAAACCCCTCG
3121  CCTGTTTTTAACTTTGCAGCCTGGGAGGGAGCAAGTGTCAGCTCAGAGACCCGAGGCCTG
3181  TGCTCTTCTTCCTGAGCCAGTGGGGAATTCGGCCCTACTAGGCTTACCAGTTGGAAAGTC
3241  AGGACCTGTTTAGGCATAGCCTCCAGTTCCTTTCGGTGTACACTGCGTGAGCTTACAGCT
3301  TCCAAGCTAAAAGGAAAAGAAAAATACTTGGCAGAAATGGGGAATGCCTGTGAGAGGGCT
3361  TGGGTTCTTGAATGGGTATAACAAATGAAGTCAGTTTACATAGGGGTTTTGTATTCATTT
3421  TAAAGTCCTCTAGAATCCCCCATTCCTTTAGCCTGTACCC(a)n
```

Figure 3.7: Nucleotide sequence and translation of *Gtl8* cDNA. cDNA consists of a 5'-UTR of 266 bp, an open reading frame (ORF) of 2196 bp which encodes for 731 predicted amino acids and a 3'-UTR of 998 bp. The start (ATG) and the stop (TGA) codons are in bold; the atypical polyadenylation signal is in open box. Initially identified nucleotides by 5'RACE PCR are underlined. Conserved FERM domain (in gray) and predicted coiled coil region were detected (in dark gray). Arrowheads indicate exon-intron boundaries.

3.1.2 Expression Analysis of the Murine *Gtl8* Gene

3.1.2.1 Northern Blot Analysis of the *Gtl8* Gene in Adult Mouse Tissues

Northern blot analyses were performed to elucidate the tissues in which the *Gtl8* gene is expressed. By using a ~500 bp (corresponding 1038-1504th nucleotides) *Gtl8* cDNA fragment as a probe, a hybridization signal of 3.6 kb was obtained in RNA from kidney, testis, and liver (Fig. 3.8). No signal was detected in brain, heart and spleen, may be it was due to less RNA on the filter. The integrity of RNA was confirmed by hybridization of the blot with a cDNA probe of the ubiquitously expressed human elongation factor.

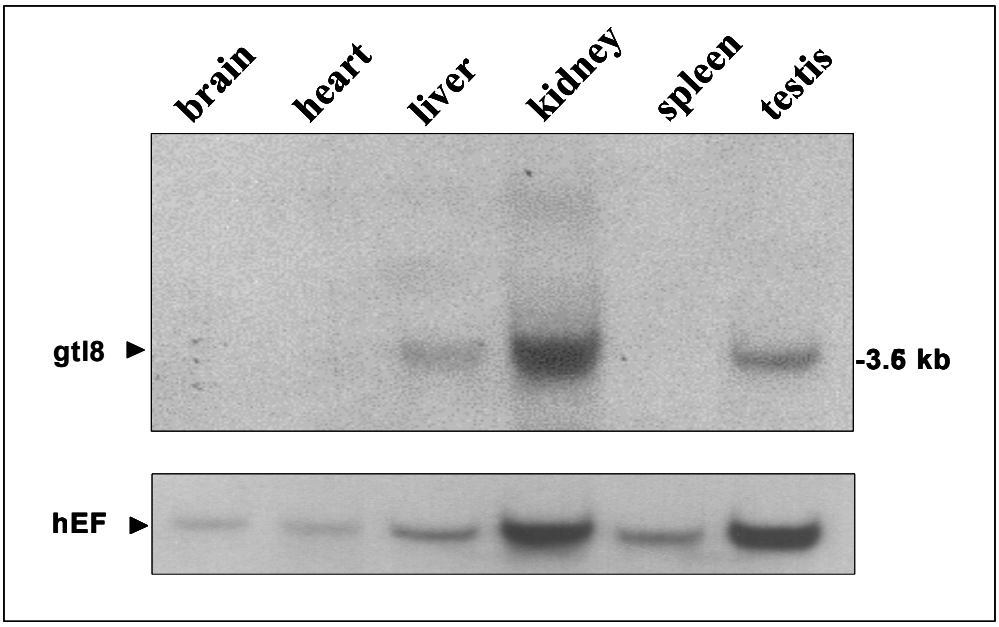


Figure 3.8: Northern blot analysis. Result of a Northern blot experiment using RNA from different adult tissues and a *Gtl8* cDNA fragment as a probe. *Gtl8* transcript 3.6 kb in size is detectable in RNA of liver, kidney and testis. Hybridization with the hEF-probe demonstrated the integrity of the loaded RNA.

3.1.2.2 RT-PCR Analysis with RNA from Adult Mouse Tissues

RNAs from different adult tissues were studied for *Gtl8* expression by RT-PCR. A 493 bp *Gtl8* fragment corresponding to the nucleotides of exon 11-16 of the murine cDNA could

be amplified from all samples examined, with exception of muscle RNA (Fig. 3.9 A). Alignment of this 493 bp cDNA with mouse EST in database, revealed no homology to other known sequences. Integrity of the RNA used for RT-PCR was proven by amplification of a fragment of the GAPDH cDNA (Fig. 3.9 B). The specificity of the PCR product was verified by hybridization using the murine *Gtl8* cDNA as a probe. (Fig. 3.9 C).

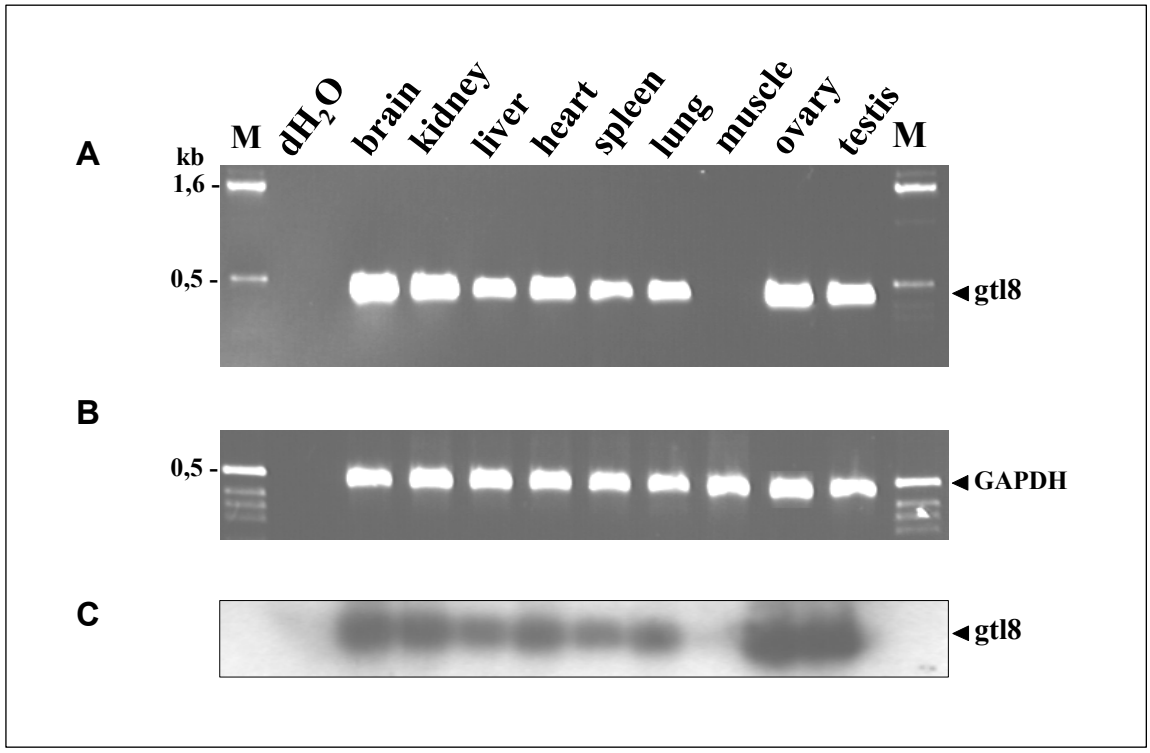


Figure 3.9: Expression analysis of *Gtl8* gene by RT-PCR. (**A**) *Gtl8* transcripts (493 bp amplification product) were observed in most of the tissues analysed using RT-PCR approach with two primers located in exons 11 and 16 of *Gtl8* cDNA. dH_2O: Sample without RNA was used as a negative control. (**B**) Quality and quantity of the RNA probes were proven by amplification of a 458 bp fragment of the murine glyceraldehyde-3-phosphate dehydrogenase transcript (GAPDH). (**C**) The specificity of the PCR product was verified by hybridization using the murine *Gtl8* cDNA as a probe. **M** is the marker for determination of the DNA fragment size

3.1.2.3 Expression of *Gtl8* during Early Embryogenesis

The expression of the *Gtl8/βgeo* mRNA was monitored using βgal detection in heterozygous mouse embryos. First, embryos were obtained from crosses between heterozygous female and wild type males (♀ +/- x ♂ +/+) (Fig. 3.10 A-E). Positive staining for βgal was detected from the unfertilized egg (Fig.3.10 A´). It persisted in morula (data not shown) and was present in the inner cell mass and trophectoderm of blastocyst (Fig. 3.10 E). A circular structure (blue point) was detected in X-gal stained embryos including unfertilized egg, two-cell and four-cell stage embryos (arrowheads in Fig. 3.10 A´, B´, C´). Interestingly the blue point was detected only in one blastomer of the 2 and 4-cell embryos. This staining was not detected in wild type embryos (Fig 3.10 A-B). βgal staining was observed in all embryos derived from ♀ +/- x ♂ +/+ breeding (Fig 3.10 D and E). This indicated the maternal *Gtl8* gene is active before fertilization and maternal storage persists at least until blastocyst stage. All 4 cell stage embryos (Fig 3.10 D) derived from heterozygous female mice were showed strong staining, but intensity of the blue staining were reduced in half of the blastocyst stage embryos (Fig 3.10 E arrows).

Second experiment was performed using the embryos derived from crosses between wild type female and heterozygous male (♀ +/+ x ♂ +/-). This experiment addressed the question during which stage of the embryonic development the embryonal *Gtl8* expression starts. No β-gal activity was observed in fertilized egg, 2-cell, 4-cell and morula stages embryos (data not shown). β-galactosidase expression was first detectable in the inner cell mass and the trophectoderm of heterozygous blastocysts (Fig. 3.10 G).

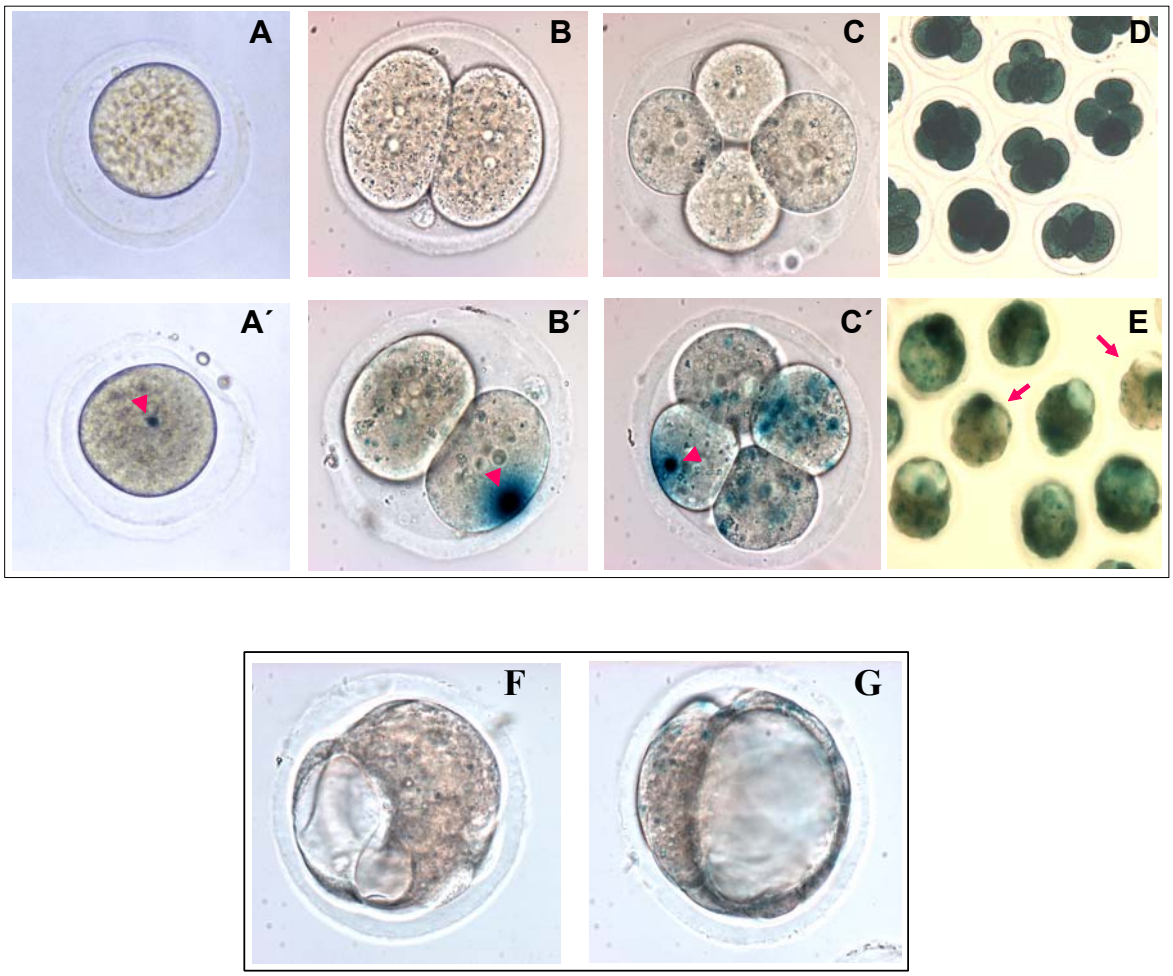


Figure 3.10: Expression of *Gtl8* in mouse pre-implantation embryo. (**A-C**) X-gal staining of unfertilized oocyte, 2-cell and 4-cell stage control embryos (from wild type crosses). (**A´-C´, E and D**) staining of maternally derived preimplantation embryos (from heterozygous female x wild type male crosses) with X-gal: (**A'**) unfertilized oocyte; (**B'**) 2-cell stage (E1.5); (**C'**) 4-cell stage (E1.5-2.0). Arrowheads indicate the circular structure (blue point). (**D**) 4-cell stage embryos shown in lower magnification. (**E**) Blastocyst stage (E3.5) embryos. Reduction of the maternal storage was observed in some blastocysts (arrows) if compared with four cell stage embryos (shown in **D**). (**F and G**): Staining of paternally derived blastocyst stage embryos (from wild type female x heterozygous male) with X-gal. (**G**) β-galactosidase expression was first detectable in the inner cell mass and the trophectoderm of blastocyst. (**F**) Wild type embryo was used as a negative control for the X-gal staining.

3.1.3 Genomic Analysis

3.1.3.1 Isolation of Cosmid Clones with Mouse Genomic DNA

To determine the genomic organization of the murine *Gtl8* gene, a cosmid library from the RZPD was screened, and two clones were isolated. The identity of these clones was proven by hybridization using the *Gtl8* cDNA (Fig. 3.11). Several fragments of the cosmid clones were cloned and sequenced to characterize the exon-intron structure of the *Gtl8* gene. Additional information was obtained from ncbi database (www.ncbi.nlm.nih.gov/genome/seq/MmBlast.html) and compared with the *Gtl8* cDNA.

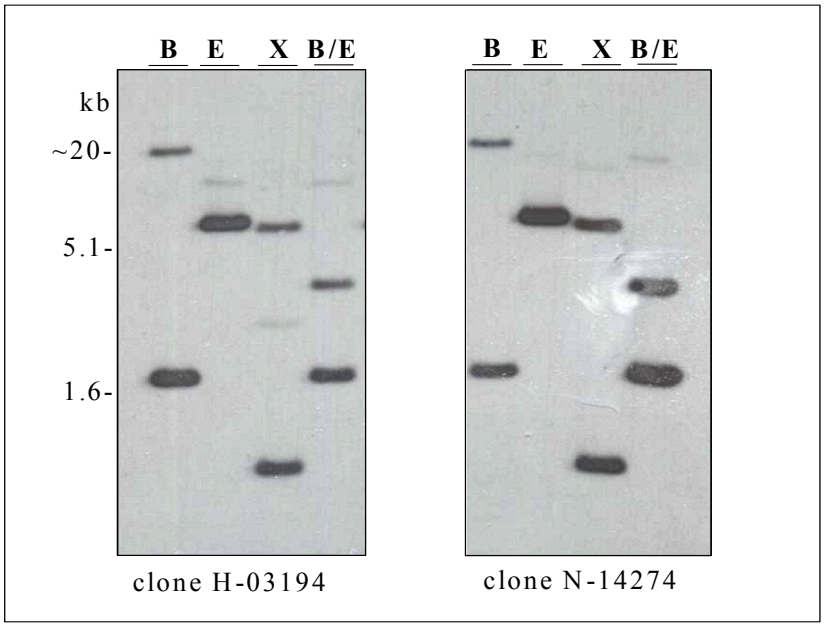


Figure 3.11: Southern blot analysis of mouse cosmid clone. Two independent cosmid clones (H-03194 and N-14274) were tested for their contents. Cosmid DNA was digested with Bam HI (B), EcoR I (E), Xba I (X) and Bam HI/EcoR I sites.

Using this approach, 25 exons of the murine *Gtl8* gene were confirmed in the genomic sequence, ranging from 28 to 1065 bp in size (Fig. 3.12). Additionally the exon-intron oundaries were determined precisely, using sequencing and comparison analysis (Fig. 3.13).

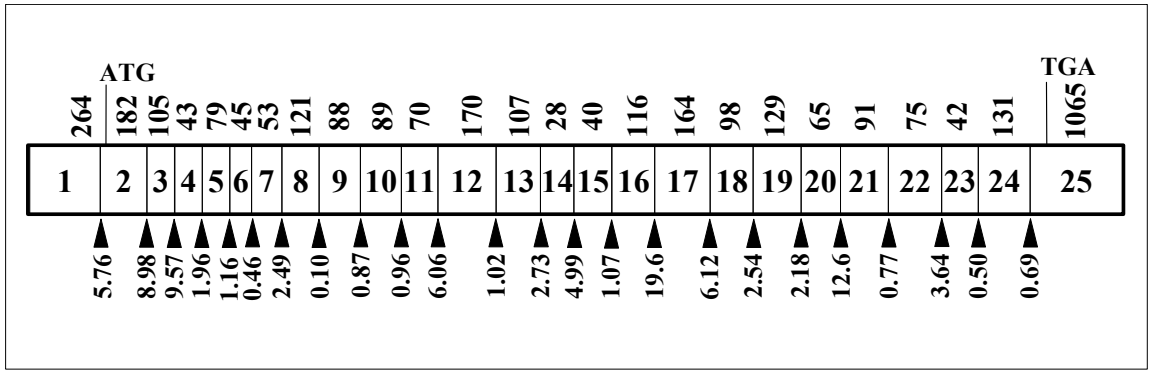


Fig 3.12: Genomic organization of the murine *Gtl8* gene. Schematic representations of the exon-intron structure of the murine *Gtl8* gene. Numbers directly above the exons and directly beneath the introns refer to their lengths (in bp and kb respectively). Translational start and stop sites are indicated.

To determine the copy number of the *Gtl8* gene in mouse genome, a Southern blot analysis was performed using the *Gtl8* cDNA as a probe. Genomic DNA and DNA of the isolated cosmid clones were hybridized and similar hybridization patterns were obtained, indicating that the *Gtl8* gene is a single copy gene in murine genome (Fig. 3.16)

Upstream intron	EXON	Downstream intron
...GTGGCCGCTG	... Exon 1 ... CGCGCGTCCG	/ **gt**aatgatgggttctcctgtg...
...caactcactggttttgtactggc**ag**	/ AAATGCTGAG ... Exon 2 ... GGACTTGCCG	/ **gt**aagtaggactttgctagctg...
...tcttattaatttatcttcttgctgtttc**ag**	/ AAAAAAGCCA ... Exon 3 ... CCAAGTAGCC	/ **gt**gagtaatcttttggaagttc...
...actcaacttttgtttttctgtttac**ag**	/ CATTGGTTGGA ... Exon 4 ... GCAAGTAAAA	/ **gt**aagtattgggaaatcatttg...
...taaagctcttttcttttttttaatggc**ag**	/ ATTGGTTCACC ... Exon 5 ... GCTAACCCGG	/ **gt**aagacagaggtgtgtcagga...
...tagctaatttgtgccttttatttttcctt**ag**	/ TATTTATTTGT ... Exon 6 ... CAGTGGGAAG	/ **gt**gagtattaaattatttagtaca....
...cggaaaaatttatattttgatctc**ag**	/ TTAGAGTGTC ... Exon 7 ... AATCTGCAAG	/ **gt**acgtgattcttgtgcttgatt...
...cataactgtgttaacttgtatttgt**ag**	/ CTGAACTTGG ... Exon 8 ... AAGGAATACA	/ **gt**agctcgttatttcatcacatt...
...cttataattactatatttaaatttt**cag**	/ GAGGTCAGAC ... Exon 9 ... CGTGGTCAAG	/ **gt**aagccttgtgttgtggtacctg...
...acactttattttctctccttacatt**ag**	/ GCTAGAGATG ... Exon 10 ... TGTTTTTCTG	/ **gt**aagcagtgtcttgtttttaatg...
...cctaatctgttgatcttatattattac**ag**	/ GCCCAAGATA ... Exon 11 ... TGATGATCAG	/ **gt**aggaacttactgtattacttgct...
...ttaataaacatcgcttcgtctttc**ag**	/ GGCAAAGAAC ... Exon 12 ... TTAGATACAG	/ **gt**tcagttccccaagttttcacat...
...ctgttgacttgcttcacttatttt**ag**	/ TGGAAAAACA ... Exon 13 ... CAAATGAAAG	/ **gt**gaagtgcagccctgctcc...
...atgttctctcttctcttctcttctgtc**ag**	/ CCAGCACAAC ... Exon 14 ... AAGATCTTGG	/ **gt**aagtactgccttgtaatgtag...
... tttagtaggttttcttctcttctttac**ag**	/ TGTTCTTAAT ... Exon 15 ... CTCCCAACAG	/ **gt**aagatagcagtaagtttcta...
...tttgactttccttattcttctgtccaa**ag**	/ GCTTGGGGTG ... Exon16 ... ACAGGAAATG	/ **gt**ttgttaacttcctaaaccaaaa...
...tgtaagtaaactttctgttacttttac**ag**	/ CTTGCCCCTA ... Exon 17 ... GAATGTGAGA	/ **gt**aagcacctgctctgtaatttatg...
...ctctctctctccttgctttttttttt**cag**	/ CATTAAAGGA ... Exon 18 ... CAACAGCCAG	/ **gt**attgggatctcttctgtatcta...
...aaagtgattatatcttttaacattc**ag**	/ GAAGAAGTGG ... Exon 19 ... TGTGCATAAG	/ **gt**aagatttctttaatagacat...
...tgtttcattatcatgatttaattttctt**ag**	/ GTTACAAGAG ... Exon 20 ... CCACAAACAG	/ **gt**acaatcttgggtaaactgaa...
...atccaggcctgcttttatttccttcc**ag**	/ TACTGCATTC ... Exon 21 ... TGTACTAAAG	/ **gt**aagaagattgctttccatttttca...
...tggactcttggtgccctgtcttctt**tag**	/ GATGCTACTG ... Exon 22 ... AACACCTCAG	/ **gt**aagtgtgtcttgaaacagcaa...
...aagtaagtggctatgatttttttttt**ag**	/ GTGTCTTCAC ... Exon 23 ... GATTATTCCA	/ **gt**aagttcatactctttatgattctt...
...ctgactgttattccttcctcttccac**ag**	/ CAGAGTGCTA ... Exon 24 ... CTGTGACCAG	/ **gt**gaggaattacttcttagttgca...
...cttactaggcctctttcttccatac**ag**	/ CTCTGCCCCC ... Exon 25 ... GCCTGTACCC	

Figure 3.13: The exon-intron boundaries of the mouse *Gtl8* gene.

3.1.3.2 Chromosomal Localization of the Murine *Gtl8* gene

The chromosomal localization of the mouse *Gtl8* gene was elucidated by fluorescence *in situ* hybridization using the labelled *Gtl8* cosmid clone as a probe. The *Gtl8* gene was mapped to mouse chromosome 1 region E2-F (Fig 3.14).

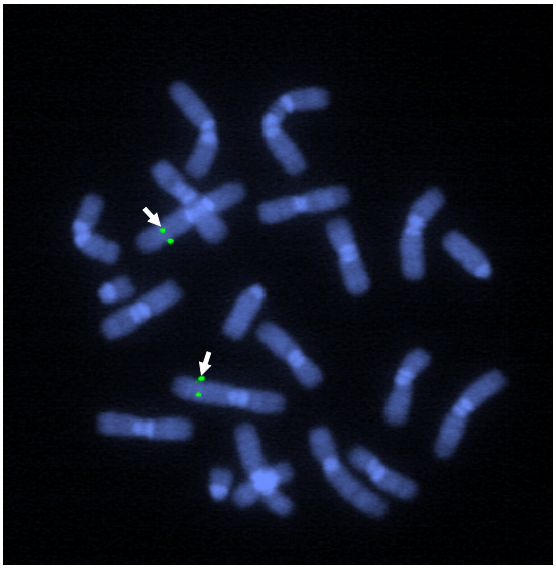


Figure 3.14: Chromosomal localization of the *Gtl8* gene. Arrow shows the position of hybridization signals on both chromosomes 1 region E2-F on metaphase chromosome of the mouse WMP-1 cell line. Metaphase chromosomes were counterstained with DAPI.

3.1.3.3 Mouse *Gtl8* Homology in Human.

From the alignment of mouse *Gtl8* complete cDNA with human genomic DNA in NCBI database revealed the *hypothetical* human gene localized on human chromosome 2q14 which is to synthenic mouse chromosome 1. The putative human gene contains 24 exons, 2196 bp ORF encoding 732 amino acids. The two genes are 83% identical on the amino acid level (Fig. 3.15)

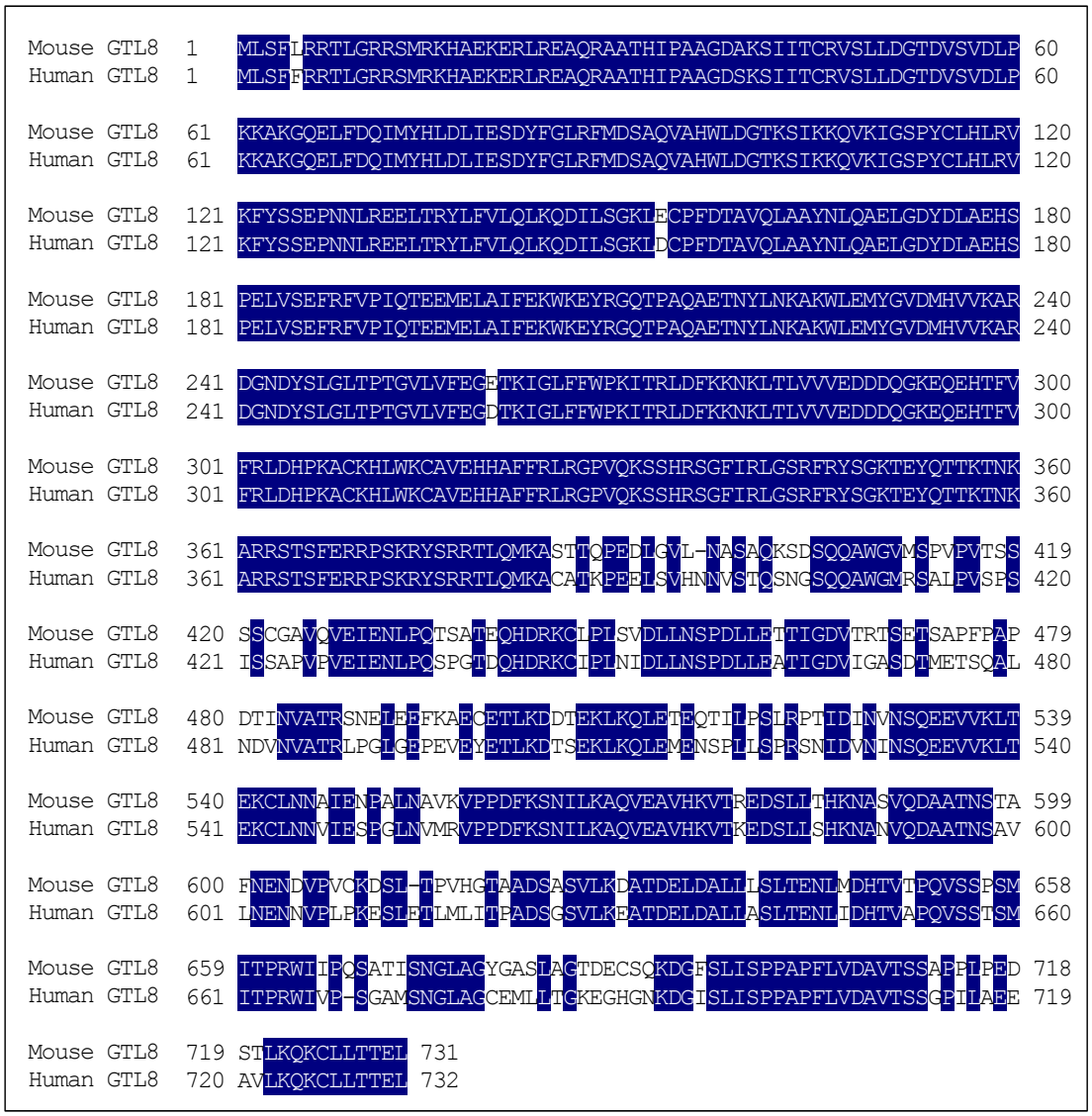


Figure 3.15: Alignment of the mouse and human *Gtl8* amino acid sequences. Black boxes indicate identical amino acids.

3.1.4 Phenotypical Analysis

3.1.4.1 Cloning and Characterization of the *Gtl8* Mutated Allele

In order to verify whether the *Gtl8* gene was disrupted by the integration of the GT vector, the genomic DNA from heterozygous and wild type animals was digested with different

restriction enzymes. The DNA was loaded on the agarose gel and blotted onto the nitrocellulose membrane for a Southern blot analysis. The blot was hybridized with *Gtl8* cDNA probe (corresponding to exons 12-14). A 2.0 and a ~20 kb band was observed in Bam HI digested heterozygous and wild type genomic DNA. (Fig. 3.16 A). An additional 8 kb band was observed only in the heterozygous DNA. This result indicated that the ~20 kb band of the wild type allele was affected by the integration and the additional 8 kb band was derived from the mutant allele due to the Bam HI restriction site of the GT vector (Fig 3.17). Restriction and hybridization analysis of the cosmid DNA showed similar results (Fig 3.16 B) as wild type genomic DNA.

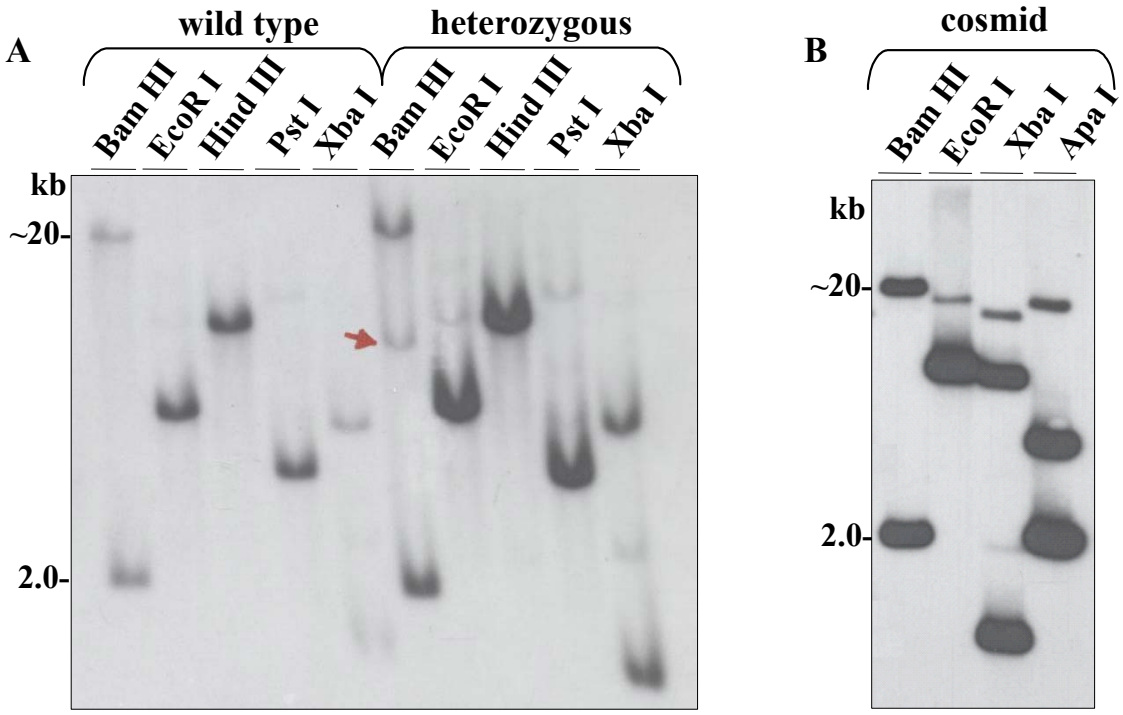


Figure 3.16: Detection of the mutant *Gtl8* allele. A. Southern blot analysis of genomic DNA from wild type and heterozygous animal. Arrow indicates the additional 8 kb band derived mutant allele in DNA of heterozygous animal. **B**. Cosmid clone was also tested for the content of the relevant information which allowed identifying and cloning the flanking region of the integration site on the wild type situation.

3.1.4.2 Gene Trap Vector Integration Site in *Gtl8* Locus

The GT vector integration site in *Gtl8* locus was precisely determined. Analysis of the 5' RACE PCR product and the complete *Gtl8* cDNA sequence indicated that the GT vector insertion has occurred in intron 14 of the *Gtl8* gene. The *Gtl8* genomic organization and cosmid DNA analysis allowed us to assume the size and sequence information of intron 14 in the wild type situation. Total length of the intron 14 was 4991 bp. The mutant DNA between the intron 14 and *En 2-intron* sequence of the GT vector was amplified by PCR using genomic DNA from the mutant mice as template. Sequence comparisons of the mutant and wild type PCR products revealed sequence identity up to the 3345 nucleotides of intron 14 (Fig. 3.17).

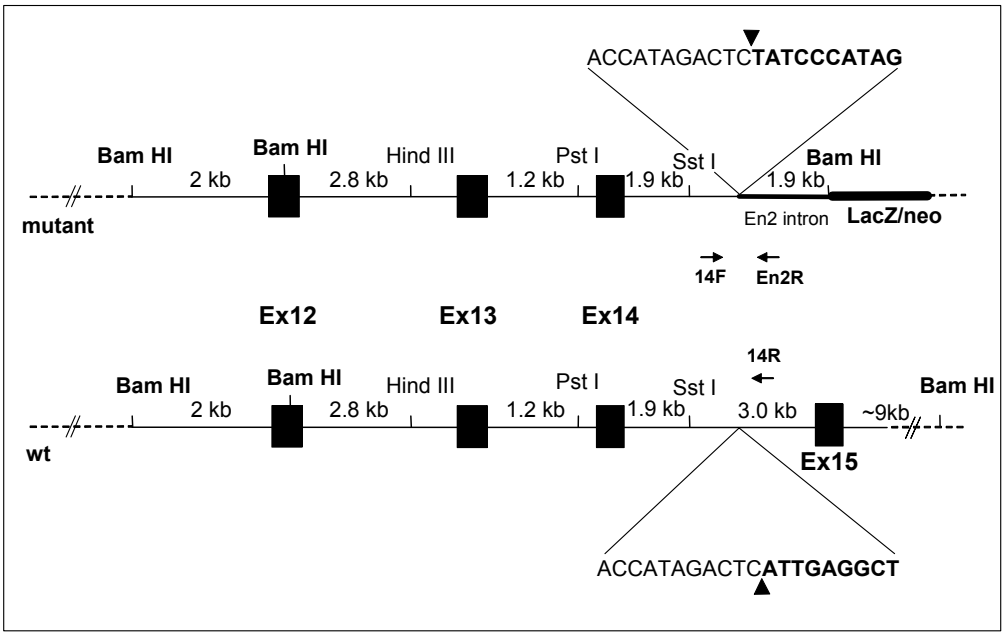


Figure 3.17: Characterisation of the gene-trapped genomic locus. Genomic structure of the mutated *Gtl8* locus and its wild type (wt) counterpart. Black triangle indicates the GT vector insertion site. Sequences downstream of the integration site corresponding to *En2-intron* and wild type loci are shown in bold. Primers are indicated as arrows. Restriction sites and length of the fragments are also shown. Bam HI restriction sites in bold, were used for the digestion of the genomic DNA (in Fig 3.16).

3.1.4.3 Embryonic Lethal Phenotype of *Gtl8* Homozygous Mutants

Mice heterozygous for the Gene Trap mutation in *Gtl8* gene were grossly phenotypically normal and fertile. Crosses between heterozygous mice failed to yield any homozygous mutant animals among the offspring (n= 46 in Table 3.1). From this experiment could be concluded that the insertion of the GT vector into *Gtl8* gene resulted in a fusion transcript (after exon 14) and thus caused a recessive loss-of-function mutation that was embryonic lethal. To determine the developmental stage at which the loss of *Gtl8* causes death, pre and postnatal embryos were isolated from heterozygous crosses. Neonates and post-implantation embryos were genotyped by Southern blot analysis. 2.0 and ~20 kb bands were detected in wild type and an additional 8 kb band was observed in heterozygous animals (Fig 3.18)

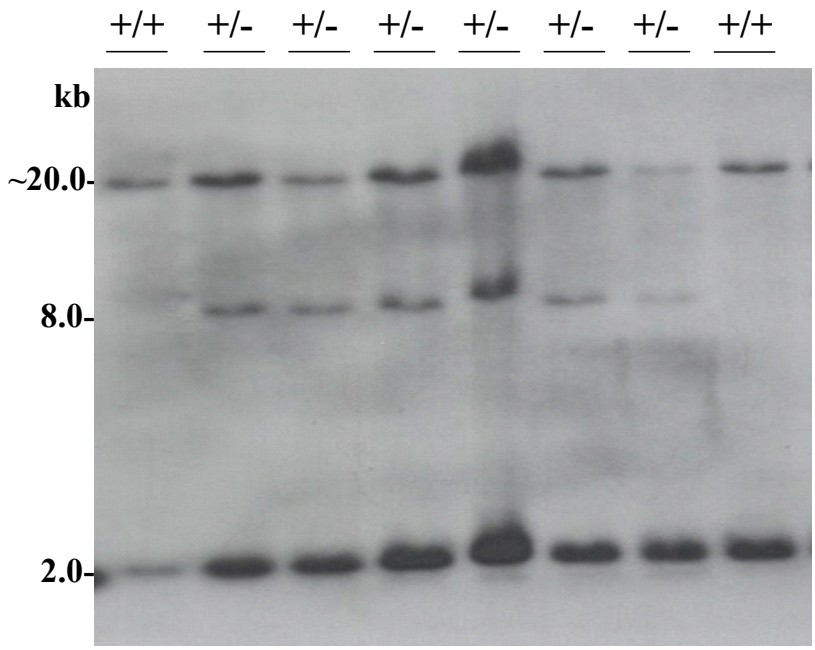


Figure 3.18: Southern blot analysis of post-implantation embryos and neonates from heterozygous crosses. Bam HI-digested genomic DNA reveals a gene trap-specific extra 8 kb fragment, that is absent in wild type DNA when hybridized with cDNA probe. cDNA probe corresponded to the sequences of exons 12 to 14.

Determination of the precise localization of the GT vector integration site in the *Gtl8* locus gave the opportunity to design specific primers on both the endogenous gene and the GT vector sequences and to amplify genomic DNA fragments for genotyping. Pre-implantation embryos were collected and cultured from two cell stage (E1.5) to avoid any maternal cell contaminations. Embryos were individually photographed and transferred to PCR cups and genotyped by two round of PCR using combinations of *Gtl8* and GT vector specific primers (Fig. 3.19). From individual 2 and 4 cell stage embryos, zona pellucida was dissolved and polar bodies were removed, before genotyping.

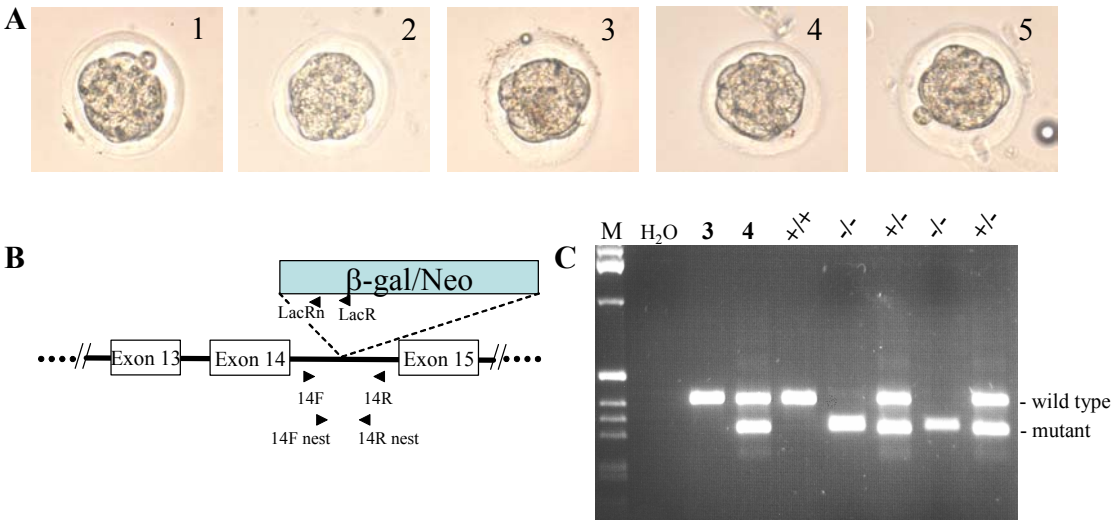


Figure 3.19: Genotyping of pre-implantation embryos. (**A**) Embryos were individually photographed (morula stage embryos). (**B**) Genomic PCR strategy for genotyping embryos; a part of *Gtl8* genomic region for GT integration is illustrated. In the first PCR amplification two *Gtl8* gene specific primers (14F and 14R) were combined with one GT vector specific primer (LacR). Nested primers (14 F nest, 14R nest and LacRn) were used for the second round of PCR amplification. The wild type allele yields a 404 bp fragment and the mutant allele results in a 268 bp fragment. (**C**) PCR based genotyping result: (lane 1), 1 kb DNA marker; (lane 2), water was used as a negative control. PCR products (lanes 3 and 4) using primer combinations on genomic DNA from adult animals. (lanes 5-9), PCR products obtained using the described combination of primers on wild type (+/+), heterozygous (+/-) and homozygous (-/-) mutant morula stage embryos.

Results of genotypic distribution of embryos and offspring, derived from heterozygous intercrosses, are shown in the following table (Table 3.1). Homozygous embryos could be detected until blastocyst stage, indicating that homozygous embryos are dying during pre-implantation development. Embryos in age of E4.5, E5.5 and E6.5 were very difficult to isolate from the heterozygous female mice without any contamination of the maternal cells. Therefore the blastocysts (E3.5) were obtained from timed heterozygous intercrosses were cultured individually in 96 well plates for 4 days. The blastocyst outgrowth (Table 3.1) was monitored daily, and DNA of embryos was extracted on the last day of culture and genotyped by PCR assay (Fig 3.19).

	Genotypes				Total
Age (days)	Wild type	Heterozygous	Homozygous	unknown	
1,5 (2-cell)	4	12	3	2	**21**
2,0 (4-cell)	8	9	6	2	**25**
2,5 (8-cell)	6	4	4	3	**17**
3,0 (morula)	7	19	6	4	**38**
3,5 (blastocyst)	14	39	2	0	**55**
outgrowth	8	5	0	3	**16**
8,5 dpc	7	17	0	0	**24**
10.5 dpc	9	16	0	0	**25**
12,5 dpc	3	11	0	0	**14**
13,5 dpc	5	12	0	0	**17**
3 weeks old	13	33	0	0	**46**

Table 3.1: Genotyping of embryos and offspring derived from heterozygous intercrosses.

3.1.5 Gtl8 Protein analyses

3.1.5.1 Putative Amino Acid Sequence of the Gtl8 Protein

The putative Gtl8 protein consists of 731 amino acid (a.a.) showing a molecular weight of 81635 Da and a predicted pI of 5.98.

Using different protein analysis using computer programs contained in the Expasy tools (http://www.expasy.ch/tools) two motifs were found in Gtl8 protein: 1. a coiled coil repeat was identified started at position 483 and ended at position 516 (Fig 3.8 and 3.21).
2. FERM domain at the N- terminal part of the molecule in amino acids position 45-235 (Fig. 3.7 and 3.20).

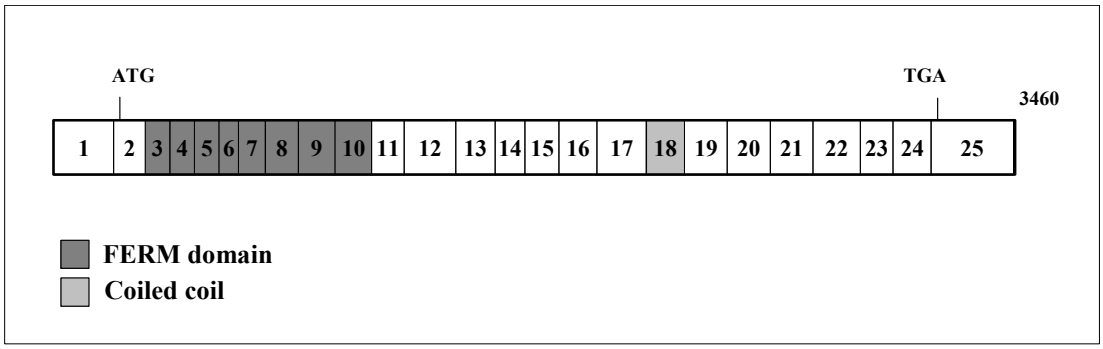


Figure 3.20: Putative protein domains present in the Gtl8 protein. The Gtl8 protein contains two putative domains: a FERM domain, and a Coiled coil domain.

FERM domains (Four.1 protein, Ezrin, Radixin, Moesin) are widespread protein modules of ~300 amino acids in length that are involved in localizing proteins to the plasma membrane. They are found in member of cytoskeletal-associated proteins that associate with various proteins at the N-terminus of the majority of FERM-containing proteins. Therefore Gtl8 protein was compared with other FERM domain containing proteins. Significant sequence homology was detected between the predicted amino acid sequences

of the Gtl8 and those of the Protein 4.1 family proteins. The comparison of the Gtl8 protein with the Protein 4.1 family proteins revealed 74, 51, 47% identity of amino acids to the NH_2-terminal regions of the EHM2, Protein 4.1G, and NBL4 respectively (Fig. 3.21). These three proteins were highest homology with the Gtl8 protein among more than 30 other mouse and human proteins identified to date. The region identity between the Gtl8 and the Protein 4.1 family proteins was restricted to the FERM domain. The remaining COOH-terminal amino acids of the Gtl8 protein showed no significant homology to other known proteins.

```
GTL8   CRVSLLDG-----TDVSVDLPKKAKGQELFDQIMYHLDLIESDYFGLRFMDSAQV
EHM2   CRVFLLDG-----TEVSVDLPKHAKGQDLFDQIVYHLDLVETDYFGLQFLDSAQV
4.1G   ----LLDG-----TEYSCDLEKHAKGQVLFDKVCEHLNLLEKDYFGLLFQESPEQ
NBL4   CEVLLLDESKLTLTTQQQGIKKSTKGSVVLDHVFHHVNLVEIDYFGLRYCDRSHQ
       ! ! ***!     *  ! !! *!!**! !!*    *! * * ***** ! !!

GTL8   AHWLDGTKSIKKQ---VKIGSPYCLHLRVKFYSSEPNNLREELTRYLFVLQLKQD
EHM2   THWLDHAKPIKKQ---MKVGPAYALHFRVKYYSSEPNNLREEFTRYLFVLQLRHD
4.1G   KNWLDPAKEIKRQ---L-RNLPWLFTFNVKFYPPDPSQLTEDITRYFLCLQLRQD
NBL4   TYWLDPAKTLAEHKELINTGPPYTLYFGIKFYAEDPCKLKEEITRYQFFLQVKQD
         *** !* !! !      ! !! ! ! !*!*   *  ! *! *** * **! !*

GTL8   ILSGKLECPFDTAVQLAAYNLQAELGDYDLAEHSPELVSEFRFVPIQTEEMELAI
EHM2   ILSGKLKCPYETAVELAALCLQAELGECELPEHTPELVSEFRFIPNQTEAMEFDI
4.1G   IASGRLPCSFVTHALLGSYTLQAELGDYDPEEHGSIDLSEFQFAPTQTKELEEKV
NBL4   VLQGRLPCPVNTAAQLGAYAIQSELGDYDPYKHTAGYVSEYRFVPDQKEELEEAI
       !!!* * *!  *!  * !  !*!***!!!  !*    !**! * * *!!! *  !

GTL8   FEKWKEYRGQTPAQAETNYLNKAKWLEMYGVDMH
EHM2   FQRWKEYRGKSPAQAELSYLNKAKWLEMYGVDMH
4.1G   AELHKTHRGLSPAQADSQFLENAKRLSMYGVDLH
NBL4   ERIHKTLMGQIPSEAELNYLRTAKSLEMYGVDLH
           *  !*  *!!*!  !*  ** *!***** *
```

Figure 3.21: Alignments of amino acid sequence corresponding to the FERM domains of the Gtl8 protein and other proteins of Protein 4.1 family. The deduced amino acid sequence of the Gtl8 protein is compared with those of the mouse EHM2, human Protein 4.1G, and human NBL4. Positions of identical amino acids among four proteins are indicated by asterisks (*) and those at least three proteins are indicated by (!) symbol.

3.1.5.2 Production of a Gtl8-GST Fusion Protein

In order to continue expression analysis of Gtl8, a fusion protein with GST was generated, and purified GST-Gtl8 fusion protein was used to produce an antibody against Gtl8. To achieve this, BamH I and Pst I restricted 837 bp fragment comprising 1290-2127 bp of the *Gtl8* cDNA was cloned into pET 41(a) (Novagen, Darmstadt) vector. The resulting clones were cultured, purified and tested for the expression of the fusion protein by Western blot after the IPTG induction. After this, the clone with the best protein expression was used to purify the fusion protein. Two rabbits were immunized with ~80-100 μg of the purified fusion protein, which was mixed 1:1 to Freund's Complete Adjuvant. After 14 days a second immunization was performed with a 1:1 mix of the fusion protein with Freund's Incomplete Adjuvant. This step was repeated after 14 days, and final bleeding was done two weeks after the last immunization. Later on anti-Gtl8 antibody present in the antiserum of these two rabbits was isolated. The purified antibody (α-Gtl8) was tested using immunological methods.

3.1.5.3 Western Blot Analysis

To test of the purified antibody, Western blot analysis was carried out. For this, a PAA gel was run with the GST-Gtl8 and MBP-Gtl8 proteins. Producing the MBP-Gtl8 fusion protein, the same cDNA fragment which was used for GST-Gtl8 fusion protein was inserted into pMALc2x expression vector. Gtl8 fusion with MBP (Maltose Banding Protein, New England Biolabs) protein was necessary, in order to control the antibody purification. It is possible that when rabbits injected with GST-Gtl8 fusion protein, antibodies against GST tag could also be produced and could be purified together with targeted antibody. Therefore MBP-Gtl8 was a real control for α-Gtl8 antibody in antibody purification product. The 66 kDa GST-Gtl8 and the ~75 kDa MBP-Gtl8 fusion proteins were detected in Western blot using the purified α-Gtl8 antibodies (Fig 3.22 A).

A western blot experiment was performed with some tissue protein extracts from wild type mouse, such as: testis, liver, lung using α-Gtl8 and α-tubulin antibodies (Fig 3.22 B). The

expected ~80 kb band corresponding to Gtl8 protein was detected in protein from testes, but was not observed in liver and lung. The filter was also probed with anti-α-tubulin antibodies to control the quantity and integrity of the protein samples.

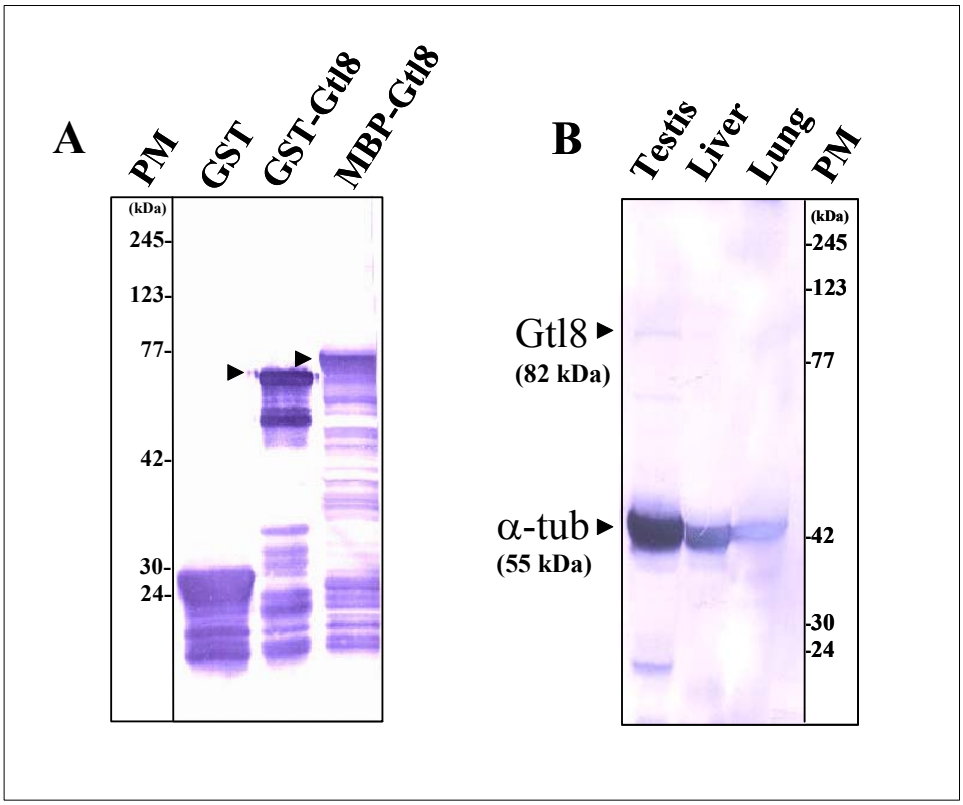


Figure 3.22. Western blot analysis using α-*Gtl8* antibody. (**A**) The 66 kDa GST-Gtl8 and the 75 kDa MBP-Gtl8 fusion proteins were recognized by purified α-Gtl8 antibodies (arrowheads). (**B**) The 82 kDa Gtl8 protein was detected using α-Gtl8 antibodies in testicular extract. Anti α-tubulin antibodies used as a control for protein integrity.

3.1.5.4 Immunocytological and Immunohistological Analysis of Gtl8

In order to elucidate the subcellular distribution of Gtl8 protein in cultured cells, immunofluorescence analyses were performed using α-Gtl8 and mouse α-actin antibodies. RT-PCR was carried out on RNA from different cell lines using *Gtl8* gene specific primers (data not shown), to choose a cell line where the *Gtl8* gene was expressed. According to this, the NIH 3T3 cell line was chosen.

First, mouse α-actin antibodies together with FITC-labelled secondary antibodies (green fluorescence) were used to visualize the cytoskeleton network of the cells. Actin filament was become visible displaying a network structure (Fig. 3.23 C). Second, Gtl8 specific antibodies detectable by Cy3-coupled antibodies (red fluorescence), was used to localize Gtl8 at the subcellular level. Unfortunately, results did not show any specific staining. Very weak red fluorescence staining was observed in nuclear and peri-nulear region (Fig. 3.23 B) but this staining was also detected in control experiment, when primary (α-Gtl8) antibody were omitted (data not shown).

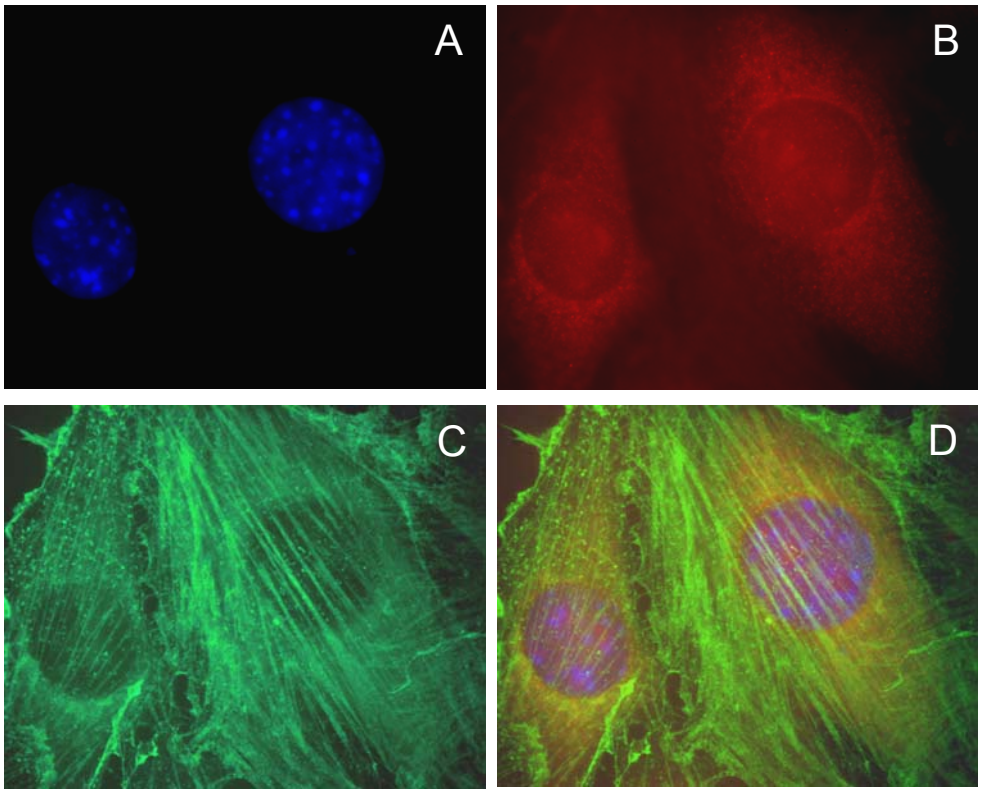


Figure 3.23: Immunofluorescence study using NIH 3T3 cells: (**A)** DAPI was used to stain the nucleus (blue colour). (**B**) Anti-Gtl8 antibodies detected by Cy-3 coupled secondary antibodies (red colour). (**C**) Actin antibodies detected by FITC-conjugated secondary antibodies visualized the cytoskeleton network of NIH 3T3 cells (green colour). (**D**) Overlay of the three images.

Reporter gene (βgal) expression was detected strongly in testis from heterozygous animal (3.1.1.1), indicating active expression of *Gtl8* in testis. Therefore, immunohistological analysis was performed on testis section using α-Gtl8 antibodies. But the α-Gtl8 antibodies did not result in any specific staining (Fig 3.24 B). In control, anti tubulin antibodies were used to visualize the tubulin network of the testicular cells at different stages of spermatogenesis in seminiferous tubules of wild type mice (Fig. 3.24 C).

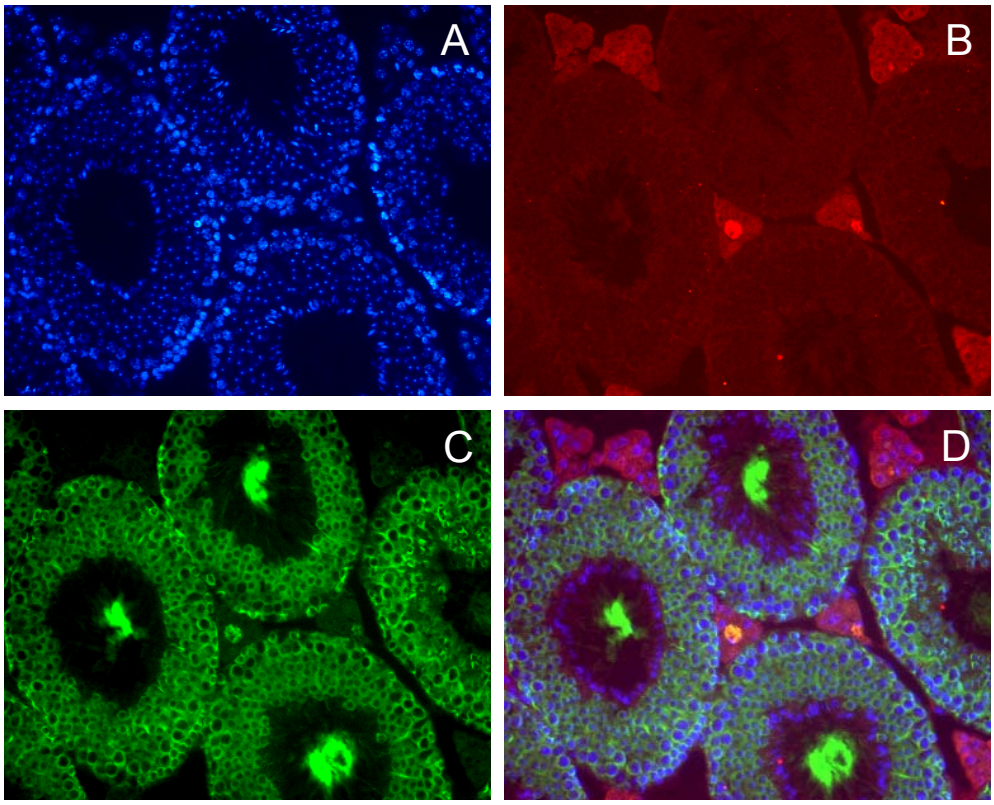


Figure 3.24: Immunohistochemical staining on wild type testis. Fluorescence micrographs of seminiferous tubules stained with DAPI (**A**), anti-Gtl8 (**B**) and anti-tubulin antibodies (**C**). No specific staining was observed with anti-Gtl8 antibodies. (**D**) Overlay images.

3.1.5.5 Gtl8-GFP Fusion Protein Analysis

To characterize the intracellular localization of the Gtl8 protein, three different *Gtl8* cDNA fragments corresponding to: 1.) whole protein coding sequence, 2.) N-terminal part including conserved FERM domain containing region and 3.) C-terminal part were inserted in frame with 3' end of an EGFP reporter gene and transfected into NIH 3T3 cells. Transient-transfection products were analysed by microscopy with fluorescence equipment 24-48 hrs after the transfection. In cells that were transfected separately with whole ORF or FERM domain containing constructs, the fluorescence signal was strongly detected in the plasma membrane, especially in cell to cell contact regions (Fig. 3.25 A-B'). However, when cells were transfected with C-terminal part, such localization was not observed. The fluorescence signals were observed in cytoplasm but not in the nucleus (Fig. 3.25 C and C'). In control experiments, cells were transfected with pEGFP vector alone. The fluorescence signal appeared evenly distributed in the whole cytoplasm and nucleus (Fig. 3.25 D and D'). Cell nuclei were counterstained with DAPI.

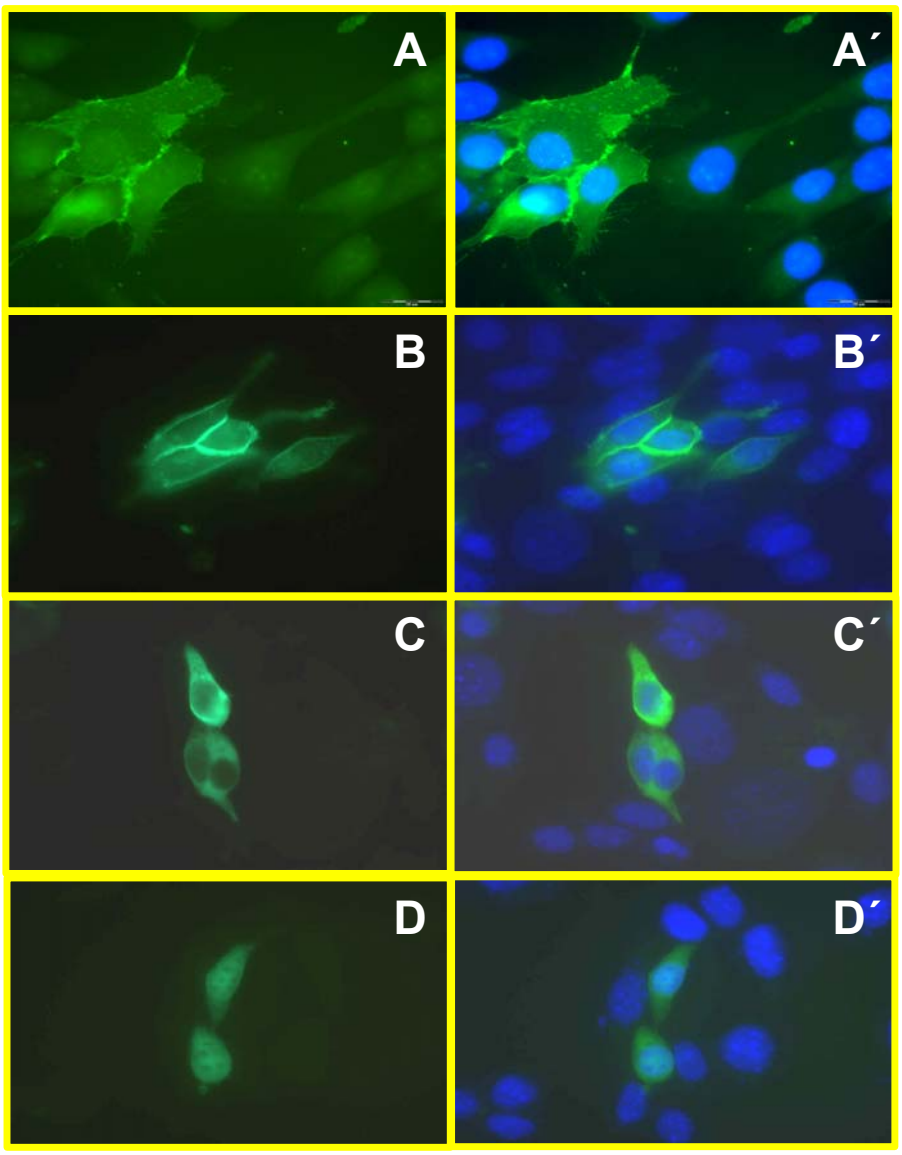


Figure 3.25: Sub-cellular distribution of EGFP-Gtl8 fusion proteins. NIH 3T3 cells were transfected with GFP-Gtl8 construct which contains the whole ORF of *Gtl8*. Fluorescence signals were observed in the plasma membrane especially in cell to cell contact region (**A**). Similar distribution was observed when cells were transfected with the Gtl8-GFP construct, containing the FERM domain (**B**). EGFP fused with C-terminal Gtl8 truncated protein was distributed throughout in cytoplasm but not in nucleus (**C**). In control experiments, cells were transfected with EGFP alone. The fluorescence signal appeared evenly distributed in the whole cytoplasm and nucleus (**D**) Cell nucleus were stained with DAPI, and overlay pictures shown are in (**A´-D´**)

3.2 Gene Trap Line 16

3.2.1 Isolation and Characterization of the Murine *Arfgef2* Gene

3.2.1.1 Identification and Cloning of the Murine *Arfgef2* Gene

Gene trap line 16 (Gtl16) was studied previously by some other students. Reporter gene (βgal) expression pattern was detected in testis of heterozygous animal using X-gal staining method. Therefore testis RNA was extracted from Gtl16 heterozygous mice. The 5' RACE PCR with testicular RNA was performed to identify the gene which is trapped in the gene trap mouse line 16. A ~900bp 5' RACE PCR product was amplified and the cDNA was subcloned into pGEM-T Easy vector and sequenced (Fig 3.27). The experimental process of the 5'RACE was same as described for *Gtl8* (3.1.1.2). The homology search in NCBI database revealed several overlapping EST clones. Using sequence informations of these EST clones, different primers were designed to amplify a whole cDNA of the trapped gene. Four additional fragments were isolated by RT-PCR on testicular RNA using the designed primers on EST clones (Fig 3.27 B). All RT-PCR products were subcloned and sequenced.

A 5,803 bp composite transcript was assembled using the five clones (Fig 3.26 B). Searches of the GenBank/EMBL nucleotide databases indicated that this sequence has not been reported. The composite cDNA contained an open reading frame of 5394 bp, with the first ATG at base 13 and stop codon (TAG) at base 5406. Using this methionine as a translation start site, a peptide of 1798 amino acids (202 kDa) was predicted. There was one consensus polyadenylation signal (AATAAA) in the 3' untranslated region of cDNA sequence (Fig 3.27)

Homology search in DDBJ/GenBank/EMBL nucleotide databases revealed nucleotide sequence of human (GenBank Accession No. NM_006420.1) ADP-ribosylation factor guanine nucleotide-exchange factor 2 (*ARFGEF2*) gene which has 93% similarity with our identified trapped gene. The gene encodes the brefeldin A (BFA)-inhibited guanine nucleotide factor 2, namely BIG2 protein. Therefore it could be the mouse ADP-

ribosylation factor guanine nucleotide-exchange factor 2 (*Arfgef2*) which encodes BIG2 in mouse.

3.2.1.2 Characterization of the Murine *Arfgef2* Genomic Structure

To expedite the characterization of the genomic structure of the mouse *Arfgef2* gene, genomic clones containing the mouse *Arfgef2* cDNA sequence were searched in GenBank nucleotide sequence database. One genomic clone was found, namely Mus musculus chromosome 2 genomic contig (NT_039210.2). On the basis of this information, 40 exons and exon-intron boundaries of the mouse *Arfgef2* gene were defined as shown Fig. 3.26.A. and Fig. 3.27.

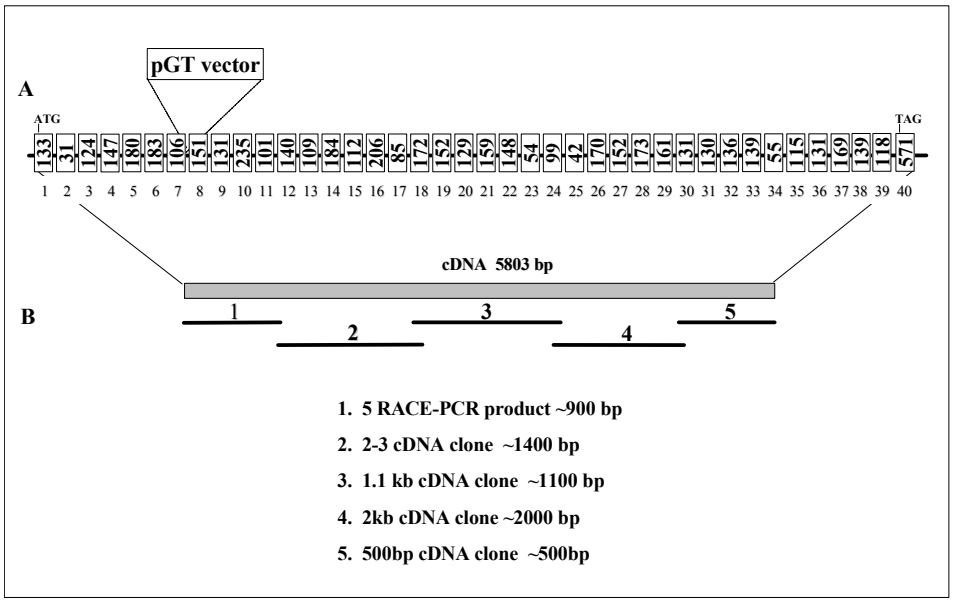
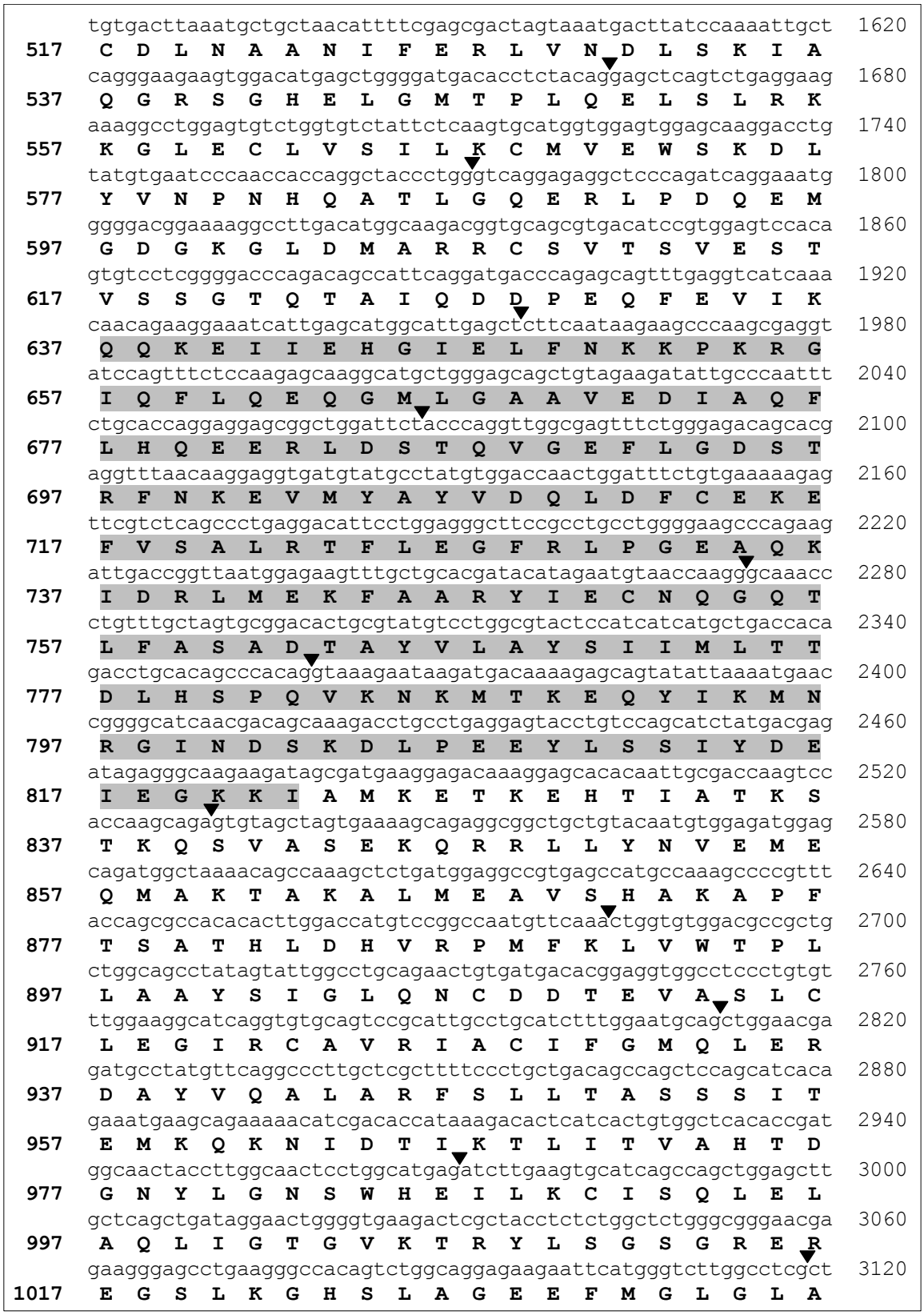
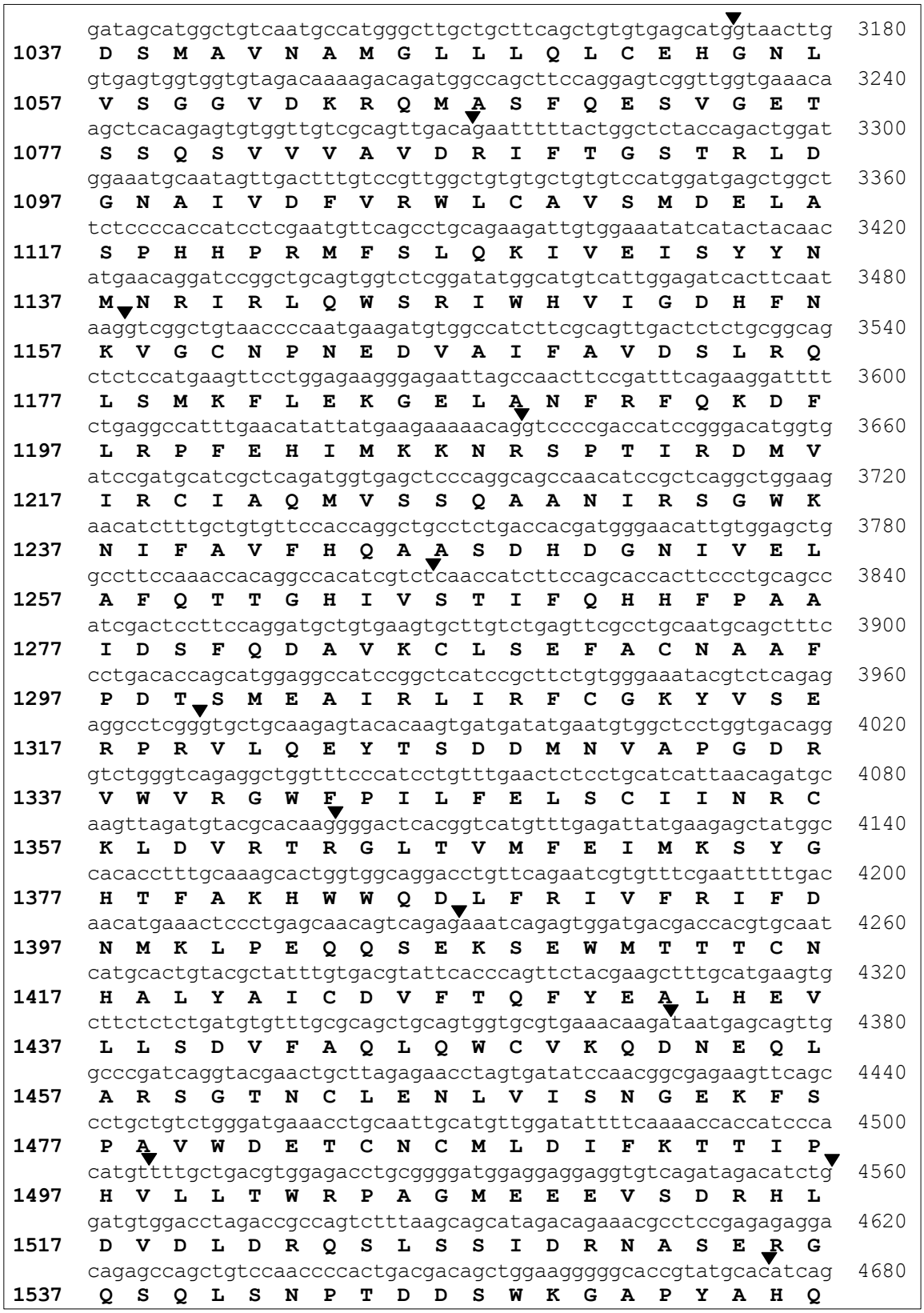


Figure 3.26: Schematic representation of cDNA cloning and genomic structure of the mouse *Arfgef2* gene. (**A**) *Arfgef2* gene consists of 40 exons. GT vector insertion occurred in intron 7 of *Arfgef2* gene. Flanking regions of the GT vector integration were cloned by PCR using the genomic DNA from heterozygous animal. (**B**) Isolated cDNA fragments (1-5) by using 5'RACE and RT-PCR techniques. All fragments were cloned and sequenced using cloning vector and gene specific primers. Complete *Arfgef2* cDNA is comprised of 5,803 nucleotides.


tgtgacttaaatgctgctaacattttcgagcgactagtaaatgacttatccaaaattgct 1620
517 C D L N A A N I F E R L V N D L S K I A
cagggaagaagtggacatgagctggggatgacacctctacaggagctcagtctgaggaag 1680
537 Q G R S G H E L G M T P L Q E L S L R K
aaaggcctggagtgtctggtgtctattctcaagtgcatggtggagtggagcaaggacctg 1740
557 K G L E C L V S I L K C M V E W S K D L
tatgtgaatcccaaccaccaggctaccctgggtcaggagaggctcccagatcaggaaatg 1800
577 Y V N P N H Q A T L G Q E R L P D Q E M
ggggacggaaaaggccttgacatggcaagacggtgcagcgtgacatccgtggagtccaca 1860
597 G D G K G L D M A R R C S V T S V E S T
gtgtcctcggggacccagacagccattcaggatgacccagagcagtttgaggtcatcaaa 1920
617 V S S G T Q T A I Q D D P E Q F E V I K
caacagaaggaaatcattgagcatggcattgagctcttcaataagaagcccaagcgaggt 1980
637 Q Q K E I I E H G I E L F N K K P K R G
atccagtttctccaagagcaaggcatgctgggagcagctgtagaagatattgcccaattt 2040
657 I Q F L Q E Q G M L G A A V E D I A Q F
ctgcaccaggaggagcggctggattctacccaggttggcgagtttctgggagacagcacg 2100
677 L H Q E E R L D S T Q V G E F L G D S T
aggtttaacaaggaggtgatgtatgcctatgtggaccaactggatttctgtgaaaaagag 2160
697 R F N K E V M Y A Y V D Q L D F C E K E
ttcgtctcagccctgaggacattcctggagggcttccgcctgcctggggaagcccagaag 2220
717 F V S A L R T F L E G F R L P G E A Q K
attgaccggttaatggagaagtttgctgcacgatacatagaatgtaaccaagggcaaacc 2280
737 I D R L M E K F A A R Y I E C N Q G Q T
ctgtttgctagtgcggacactgcgtatgtcctggcgtactccatcatcatgctgaccaca 2340
757 L F A S A D T A Y V L A Y S I I M L T T
gacctgcacagcccacaggtaaagaataagatgacaaaagagcagtatattaaaatgaac 2400
777 D L H S P Q V K N K M T K E Q Y I K M N
cggggcatcaacgacagcaaagacctgcctgaggagtacctgtccagcatctatgacgag 2460
797 R G I N D S K D L P E E Y L S S I Y D E
atagagggcaagaagatagcgatgaaggagacaaaggagcacacaattgcgaccaagtcc 2520
817 I E G K K I A M K E T K E H T I A T K S
accaagcagagtgtagctagtgaaaagcagaggcggctgctgtacaatgtggagatggag 2580
837 T K Q S V A S E K Q R R L L Y N V E M E
cagatggctaaaacagccaaagctctgatggaggccgtgagccatgccaaagccccgttt 2640
857 Q M A K T A K A L M E A V S H A K A P F
accagcgccacacacttggaccatgtccggccaatgttcaaactggtgtggacgccgctg 2700
877 T S A T H L D H V R P M F K L V W T P L
ctggcagcctatagtattggcctgcagaactgtgatgacacggaggtggcctccctgtgt 2760
897 L A A Y S I G L Q N C D D T E V A S L C
ttggaaggcatcaggtgtgcagtccgcattgcctgcatctttggaatgcagctggaacga 2820
917 L E G I R C A V R I A C I F G M Q L E R
gatgcctatgttcaggcccttgctcgcttttccctgctgacagccagctccagcatcaca 2880
937 D A Y V Q A L A R F S L L T A S S S I T
gaaatgaagcagaaaaacatcgacaccataaagacactcatcactgtggctcacaccgat 2940
957 E M K Q K N I D T I K T L I T V A H T D
ggcaactaccttggcaactcctggcatgagatcttgaagtgcatcagccagctggagctt 3000
977 G N Y L G N S W H E I L K C I S Q L E L
gctcagctgataggaactggggtgaagactcgctacctctctggctctgggcgggaacga 3060
997 A Q L I G T G V K T R Y L S G S G R E R
gaagggagcctgaagggccacagtctggcaggagaagaattcatgggtcttggcctcgct 3120
1017 E G S L K G H S L A G E E F M G L G L A


gatagcatggctgtcaatgccatgggcttgctgcttcagctgtgtgagcatggtaacttg 3180
1037 D S M A V N A M G L L L Q L C E H G N L
gtgagtggtggtgtagacaaaagacagatggccagcttccaggagtcggttggtgaaaca 3240
1057 V S G G V D K R Q M A S F Q E S V G E T
agctcacagagtgtggttgtcgcagttgacagaatttttactggctctaccagactggat 3300
1077 S S Q S V V V A V D R I F T G S T R L D
ggaaatgcaatagttgactttgtccgttggctgtgtgctgtgtccatggatgagctggct 3360
1097 G N A I V D F V R W L C A V S M D E L A
tctccccaccatcctcgaatgttcagcctgcagaagattgtggaaatatcatactacaac 3420
1117 S P H H P R M F S L Q K I V E I S Y Y N
atgaacaggatccggctgcagtggtctcggatatggcatgtcattggagatcacttcaat 3480
1137 M N R I R L Q W S R I W H V I G D H F N
aaggtcggctgtaaccccaatgaagatgtggccatcttcgcagttgactctctgcggcag 3540
1157 K V G C N P N E D V A I F A V D S L R Q
ctctccatgaagttcctggagaagggagaattagccaacttccgatttcagaaggatttt 3600
1177 L S M K F L E K G E L A N F R F Q K D F
ctgaggccatttgaacatattatgaagaaaaacaggtccccgaccatccgggacatggtg 3660
1197 L R P F E H I M K K N R S P T I R D M V
atccgatgcatcgctcagatggtgagctcccaggcagccaacatccgctcaggctggaag 3720
1217 I R C I A Q M V S S Q A A N I R S G W K
aacatctttgctgtgttccaccaggctgcctctgaccacgatgggaacattgtggagctg 3780
1237 N I F A V F H Q A A S D H D G N I V E L
gccttccaaaccacaggccacatcgtctcaaccatcttccagcaccacttccctgcagcc 3840
1257 A F Q T T G H I V S T I F Q H H F P A A
atcgactccttccaggatgctgtgaagtgcttgtctgagttcgcctgcaatgcagctttc 3900
1277 I D S F Q D A V K C L S E F A C N A A F
cctgacaccagcatggaggccatccggctcatccgcttctgtgggaaatacgtctcagag 3960
1297 P D T S M E A I R L I R F C G K Y V S E
aggcctcgggtgctgcaagagtacacaagtgatgatatgaatgtggctcctggtgacagg 4020
1317 R P R V L Q E Y T S D D M N V A P G D R
gtctgggtcagaggctggtttcccatcctgtttgaactctcctgcatcattaacagatgc 4080
1337 V W V R G W F P I L F E L S C I I N R C
aagttagatgtacgcacaagggactcacggtcatgtttgagattatgaagagctatggc 4140
1357 K L D V R T R G L T V M F E I M K S Y G
cacacctttgcaaagcactggtggcaggacctgttcagaatcgtgtttcgaatttttgac 4200
1377 H T F A K H W W Q D L F R I V F R I F D
aacatgaaactccctgagcaacagtcagagaaatcagagtggatgacgaccacgtgcaat 4260
1397 N M K L P E Q Q S E K S E W M T T T C N
catgcactgtacgctatttgtgacgtattcacccagttctacgaagctttgcatgaagtg 4320
1417 H A L Y A I C D V F T Q F Y E A L H E V
cttctctctgatgtgtttgcgcagctgcagtggtgcgtgaaacaagataatgagcagttg 4380
1437 L L S D V F A Q L Q W C V K Q D N E Q L
gcccgatcaggtacgaactgcttagagaacctagtgatatccaacggcgagaagttcagc 4440
1457 A R S G T N C L E N L V I S N G E K F S
cctgctgtctgggatgaaacctgcaattgcatgttggatattttcaaaaccaccatccca 4500
1477 P A V W D E T C N C M L D I F K T T I P
catgttttgctgacgtggagacctgcggggatggaggaggaggtgtcagatagacatctg 4560
1497 H V L L T W R P A G M E E E V S D R H L
gatgtggacctagaccgccagtctttaagcagcatagacagaaacgcctccgagagagga 4620
1517 D V D L D R Q S L S S I D R N A S E R G
cagagccagctgtccaaccccactgacgacagctggaaggggcaccgtatgcacatcag 4680
1537 Q S Q L S N P T D D S W K G A P Y A H Q

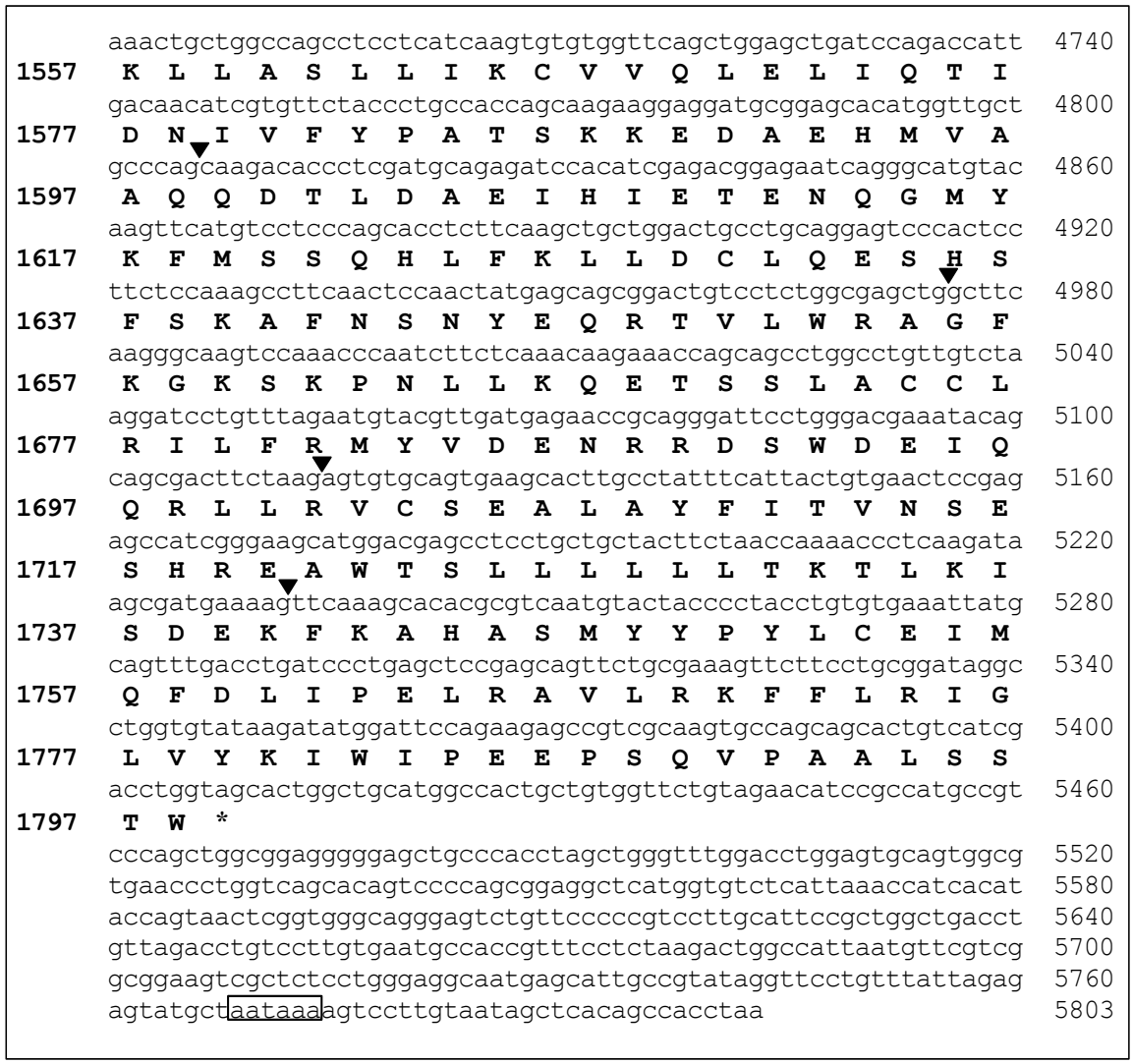
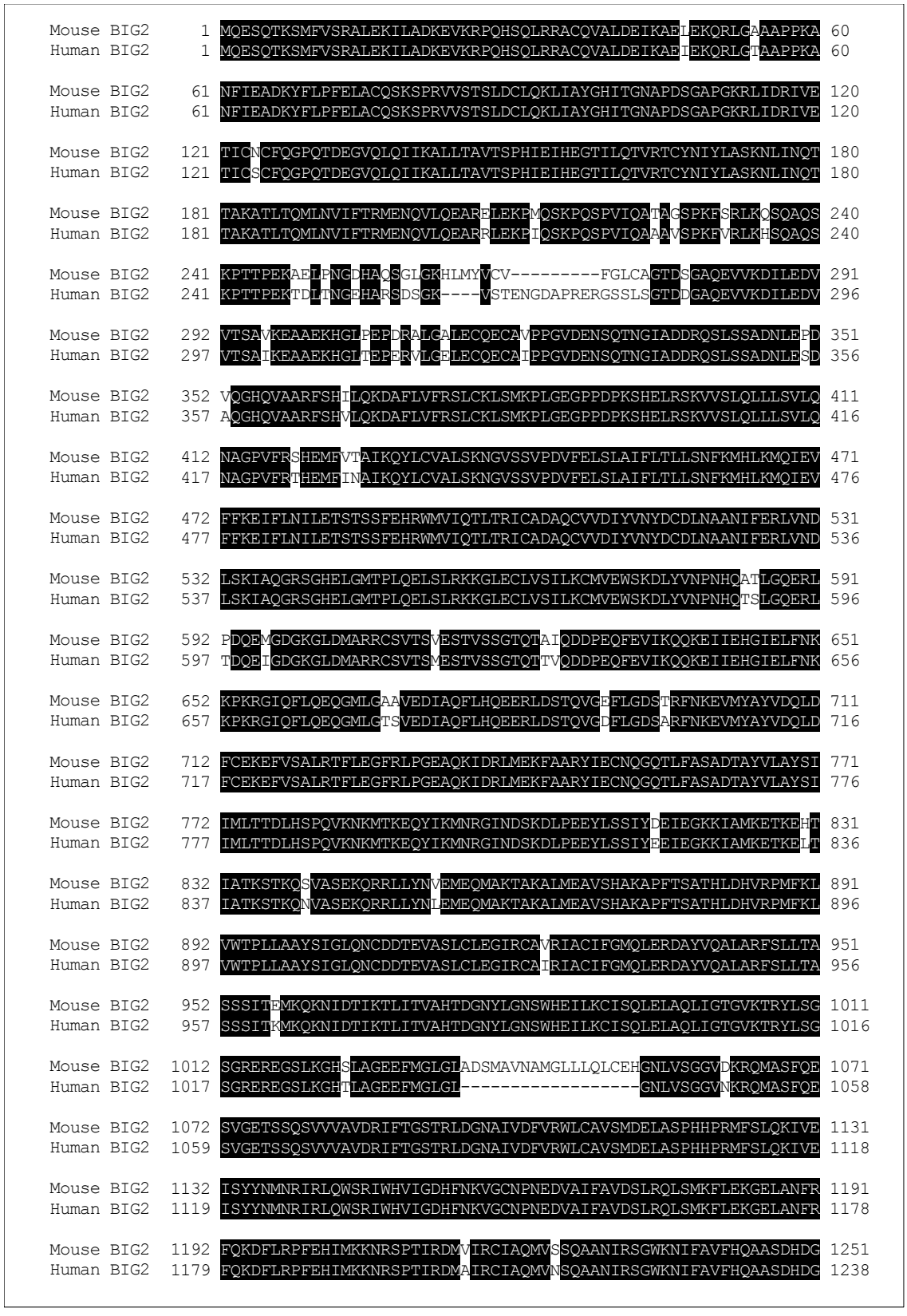


Figure 3.27: Nucleotide and predicted amino acid sequences of mouse *Arfgef2* gene. The *Arfgef2* cDNA fragment of ~900 bp, which was previously identified by 5' RACE technique, is underlined. The Sec7 domain is shaded light gray. The stop codon 5407-5409 is indicated by an asterisk. Open box indicates consensus polyadenylation signal (AATAAA). The black triangles refer to exon-intron boundaries deduced from the genomic sequence of the mouse chromosome 2 contig (NT_039210.2).

Comparison of the putative human and mouse BIG2 proteins revealed 93% identity in amino acid level (Fig 3.28).


Mouse BIG2 1 MQESQTKSMFVSRALEKILADKEVKRPQHSQLRRACQVALDEIKAELEKQRLGAAAPPKA 60
Human BIG2 1 MQESQTKSMFVSRALEKILADKEVKRPQHSQLRRACQVALDEIKAEIEKQRLGTAAPPKA 60
Mouse BIG2 61 NFIEADKYFLPFELACQSKSPRVVSTSLDCLQKLIAYGHITGNAPDSGAPGKRLIDRIVE 120
Human BIG2 61 NFIEADKYFLPFELACQSKSPRVVSTSLDCLQKLIAYGHITGNAPDSGAPGKRLIDRIVE 120
Mouse BIG2 121 TICNCFQGPQTDEGVQLQIIKALLTAVTSPHIEIHEGTILQTVRTCYNIYLASKNLINQT 180
Human BIG2 121 TICSCFQGPQTDEGVQLQIIKALLTAVTSPHIEIHEGTILQTVRTCYNIYLASKNLINQT 180
Mouse BIG2 181 TAKATLTQMLNVIFTRMENQVLQEARELEKPMQSKPQSPVIQATAGSPKFSRLKQSQAQS 240
Human BIG2 181 TAKATLTQMLNVIFTRMENQVLQEARRLEKPIQSKPQSPVIQAAAVSPKFVRLKHSQAQS 240
Mouse BIG2 241 KPTTPEKAELPNGDHAQSGLGKHLMYVCV---------FGLCAGTDSGAQEVVKDILEDV 291
Human BIG2 241 KPTTPEKTDLTNGEHARSDSGK----VSTENGDAPRERGSSLSGTDDGAQEVVKDILEDV 296
Mouse BIG2 292 VTSAVKEAAEKHGLPEPDRALGALECQECAVPPGVDENSQTNGIADDRQSLSSADNLEPD 351
Human BIG2 297 VTSAIKEAAEKHGLTEPERVLGPLECQECAIPPGVDENSQTNGIADDRQSLSSADNLESD 356
Mouse BIG2 352 VQGHQVAARFSHILQKDAFLVFRSLCKLSMKPLGEGPPDPKSHELRSKVVSLQLLLSVLQ 411
Human BIG2 357 AQGHQVAARFSEVLQKDAFLVFRSLCKLSMKPLGEGPPDPKSHELRSKVVSLQLLLSVLQ 416
Mouse BIG2 412 NAGPVFRSHEMFVTAIKQYLCVALSKNGVSSVPDVFELSLAIFLTLLSNFKMHLKMQIEV 471
Human BIG2 417 NAGPVFRTHEMFINAIKQYLCVALSKNGVSSVPDVFELSLAIFLTLLSNFKMHLKMQIEV 476
Mouse BIG2 472 FFKEIFLNILETSTSSFEHRWMVIQTLTRICADAQCVVDIYVNYDCDLNAANIFERLVND 531
Human BIG2 477 FFKEIFLNILETSTSSFEHRWMVIQTLTRICADAQCVVDIYVNYDCDLNAANIFERLVND 536
Mouse BIG2 532 LSKIAQGRSGHELGMTPLQELSLRKKGLECLVSILKCMVEWSKDLYVNPNHQATLGQERL 591
Human BIG2 537 LSKIAQGRSGHELGMTPLQELSLRKKGLECLVSILKCMVEWSKDLYVNPNHQTSLGQERL 596
Mouse BIG2 592 PDQEMGDGKGLDMARRCSVTSVESTVSSGTQTAIQDDPEQFEVIKQQKEIIEHGIELFNK 651
Human BIG2 597 TDQEIGDGKGLDMARRCSVTSMESTVSSGTQTTVQDDPEQFEVIKQQKEIIEHGIELFNK 656
Mouse BIG2 652 KPKRGIQFLQEQGMLGAAVEDIAQFLHQEERLDSTQVGEFLGDSTRFNKEVMYAYVDQLD 711
Human BIG2 657 KPKRGIQFLQEQGMLGTSVEDIAQFLHQEERLDSTQVGDFLGDSARFNKEVMYAYVDQLD 716
Mouse BIG2 712 FCEKEFVSALRTFLEGFRLPGEAQKIDRLMEKFAARYIECNQGQTLFASADTAYVLAYSI 771
Human BIG2 717 FCEKEFVSALRTFLEGFRLPGEAQKIDRLMEKFAARYIECNQGQTLFASADTAYVLAYSI 776
Mouse BIG2 772 IMLTTDLHSPQVKNKMTKEQYIKMNRGINDSKDLPEEYLSSIYDEIEGKKIAMKETKEHT 831
Human BIG2 777 IMLTTDLHSPQVKNKMTKEQYIKMNRGINDSKDLPEEYLSSIYEEIEGKKIAMKETKELT 836
Mouse BIG2 832 IATKSTKQSVASEKQRRLLYNVEMEQMAKTAKALMEAVSHAKAPFTSATHLDHVRPMFKL 891
Human BIG2 837 IATKSTKQNVASEKQRRLLYNLEMEQMAKTAKALMEAVSHAKAPFTSATHLDHVRPMFKL 896
Mouse BIG2 892 VWTPLLAAYSIGLQNCDDTEVASLCLEGIRCAVRIACIFGMQLERDAYVQALARFSLLTA 951
Human BIG2 897 VWTPLLAAYSIGLQNCDDTEVASLCLEGIRCAIRIACIFGMQLERDAYVQALARFSLLTA 956
Mouse BIG2 952 SSSITEMKQKNIDTIKTLITVAHTDGNYLGNSWHEILKCISQLELAQLIGTGVKTRYLSG 1011
Human BIG2 957 SSSITKMKQKNIDTIKTLITVAHTDGNYLGNSWHEILKCISQLELAQLIGTGVKTRYLSG 1016
Mouse BIG2 1012 SGREREGSLKGHSLAGEEFMGLGLADSMAVNAMGLLLQLCEHGNLVSGGVDKRQMASFQE 1071
Human BIG2 1017 SGREREGSLKGHTLAGEEFMGLGL----------GNLVSGGVNKRQMASFQE 1058
Mouse BIG2 1072 SVGETSSQSVVVAVDRIFTGSTRLDGNAIVDFVRWLCAVSMDELASPHHPRMFSLQKIVE 1131
Human BIG2 1059 SVGETSSQSVVVAVDRIFTGSTRLDGNAIVDFVRWLCAVSMDELASPHHPRMFSLQKIVE 1118
Mouse BIG2 1132 ISYYNMNRIRLQWSRIWHVIGDHFNKVGCNPNEDVAIFAVDSLRQLSMKFLEKGELANFR 1191
Human BIG2 1119 ISYYNMNRIRLQWSRIWHVIGDHFNKVGCNPNEDVAIFAVDSLRQLSMKFLEKGELANFR 1178
Mouse BIG2 1192 FQKDFLRPFEHIMKKNRSPTIRDMVIRCIAQMVSSQAANIRSGWKNIFAVFHQAASDHDG 1251
Human BIG2 1179 FQKDFLRPFEHIMKKNRSPTIRDMAIRCIAQMVNSQAANIRSGWKNIFAVFHQAASDHDG 1238

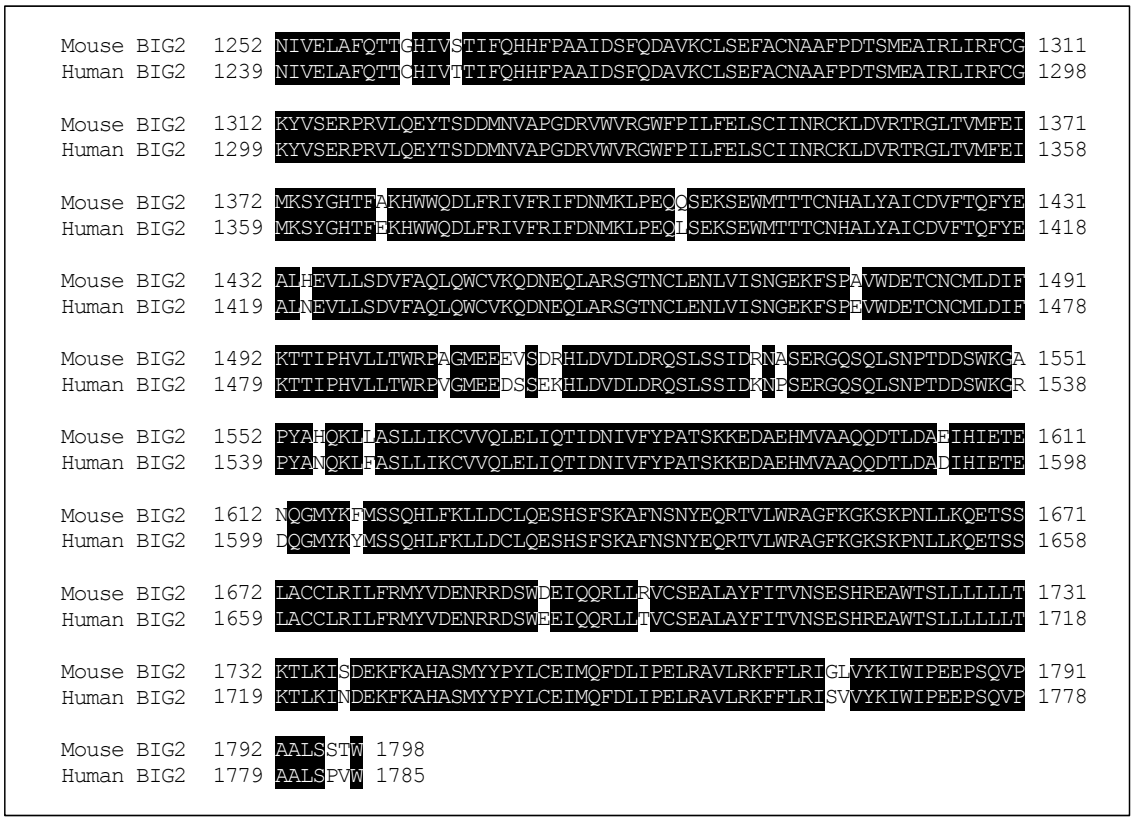


Figure 3.28: Alignment of the putative mouse and human *Arfgef2* protein sequences. Black boxes indicate identical amino acids.

3.2.2 Expression Analysis of the Murine *Arfgef2* Gene

3.2.2.1 RT-PCR Analysis

In order to determine in which tissues *Arfgef2* is expressed, RT-PCR analysis were carried out using total RNA from different mouse tissue. An *Arfgef2* transcript could be amplified from RNA samples of testis, heart, kidney, lung, liver, spleen, brain, ovary, muscle, ES cells (Fig. 3.29 A). The integrity of the RNAs used for RT-PCR was proven by amplification of the GAPDH cDNA fragment. Additionally, *Arfgef2* expression pattern was detected in RNA of pooled oocytes indicating the presence of the *Arfgef2* transcript in very early embryonic stage (Fig. 3.29 B).

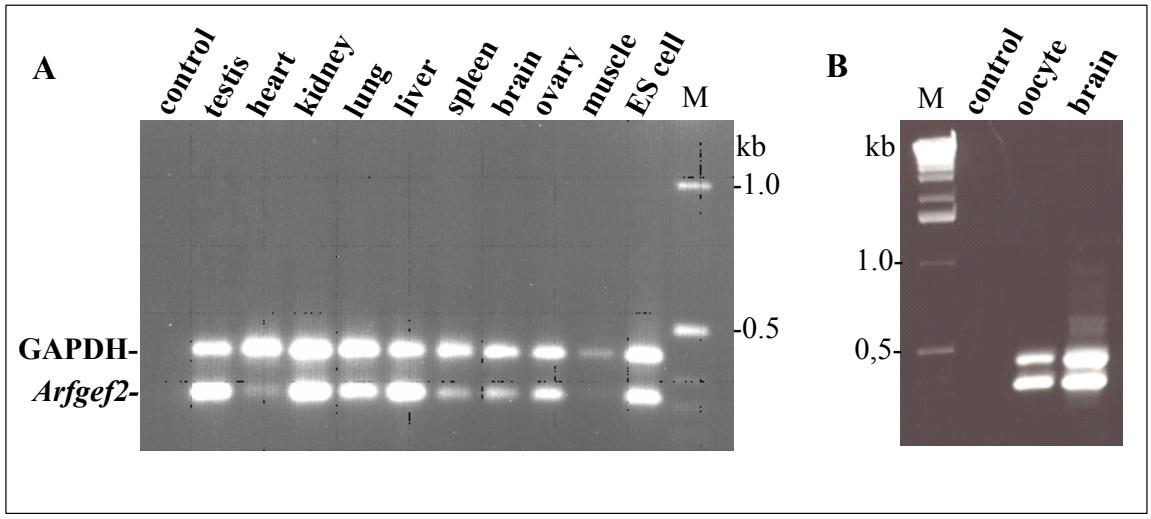


Figure 3.29: Expression analysis of *Arfgef2*. (**A**) RT-PCR was performed using RNAs from tissues of adult mouse and ES cells. A 325 bp RT-PCR product was obtained in all analysed tissues and ES cells. The GAPDH amplification was used as a control to prove the quality and quantity of the RNA. (**B**) *Arfgef2* expression was also detected in unfertilised oocytes. RNA was extracted from pooled oocytes (about 40). RNA from brain was used as a positive control. M=marker 1kb DNA ladder

3.2.2.2 Expression of *Arfgef2* during Embryogenesis

The expression of *Arfgef2*/*β-geo* mRNA was monitored using βgal detection in heterozygous mouse embryos. Transcription of the trapped allele was assessed in zygotic embryos obtained from crosses between wild-type males and heterozygous females. Positive staining for βgal was found to start in blastocyst stage (E3.5 Fig. 3.30) and was present in both, the inner cell mass and trophectoderm of blastocysts.

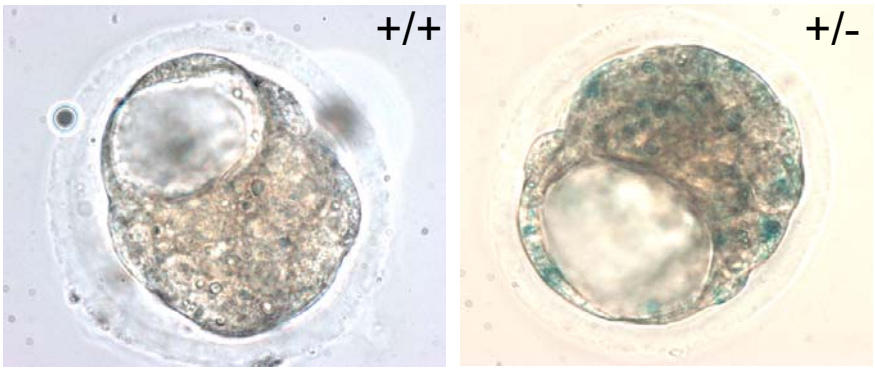


Figure 3.30: Staining of blastocyst stage embryos derived from heterozygous intercrosses with X-gal. Heterozygous embryos were derived from crosses between wild-type males and heterozygous females. βgal expression pattern was detected in both the inner cell mass and trophectoderm of heterozygous blastocysts. Wild type blastocysts showed no staining.

3.2.3 Phenotypical Analysis of the *Arfgef2* Gene

3.2.3.1 *Arfgef2* Mutation Results in Early Embryonic Lethality

Heterozygous males and females displayed no overt phenotype and were fertile. Heterozygous animals were intercrossed, and the offspring were genotyped by PCR analysis. No viable *Arfgef2*$^{-/-}$ mice were identified among over 200 mice. These results indicate that *Arfgef2* deficiency results in embryonic lethality.

For genotyping of preimplantation embryos, a PCR method was established to detect wild type and mutant alleles. Analysis of the genomic structure of the *Arfgef2* gene gave opportunity to localize the primers locating just on the flanking regions of the GT vector integration site in intron 7 of *Arfgef2* gene (Fig. 3.31 A).

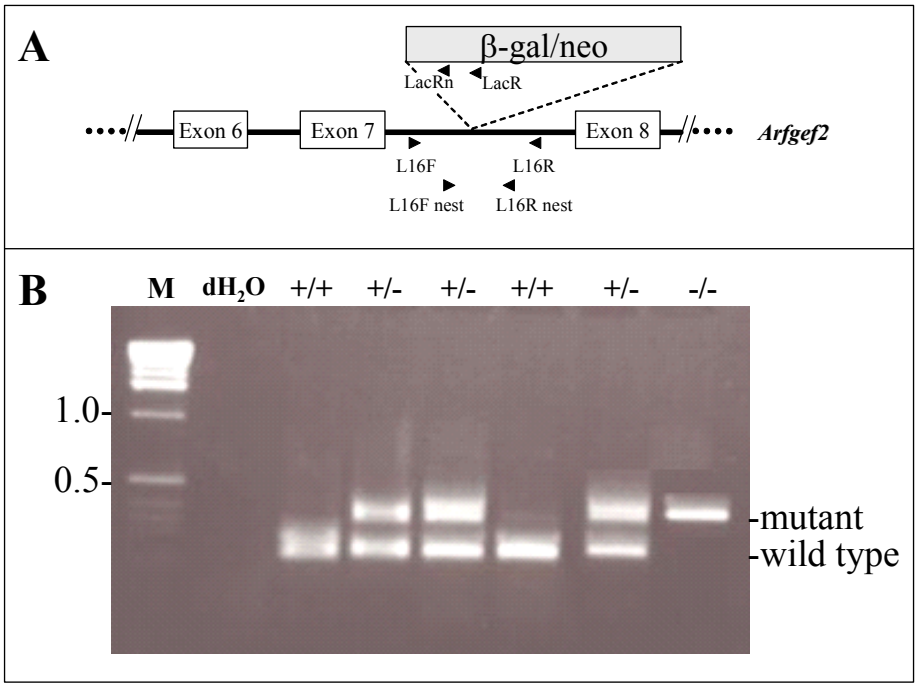


Figure 3.31: Genotyping strategy. (**A**) The schematic diagram of the gene trap event in *Arfgef2* gene. GT integration occurred in intron 7 of *Arfgef2* gene. Primers were designed in flanking regions of the GT vector integration as well as in the GT vector. (**B**) A 325 bp fragment was amplified from mutant allele using L16F nest and LacRn primers and a 167 bp fragment was amplified from wild type allele using L16F nest and L16R nest primers.

To determine the time of embryonic lethality, embryos E0.5-E3.5 were isolated from heterozygous intercrosses and genotypes were determined by PCR. No homozygous embryos were detected in 2-cell, 4-cell, 8-cell, morula and blastocyst stage embryos. At E0.5 in syngamy stage homozygous embryos were detected but not in Mendelian ratio. Syngamy stage embryos were prepared by following steps: 1. Fertilized oocytes were isolated and tested for the pronuclear stages using light microscopy. 2. Pronuclear stage embryos were incubated for 4-6 hours in M16 medium to allow the pronuclei to fuse (syngamy). Than syngamy stage oocytes were collected individually for genotyping (Table 3.2).

	Genotype of the embryos				Total
Age (days)	**Wild type**	**Heterozygous**	**Homozygous**	**unknown**	
syngamy	8	6	2	16	**34**
2 cell	36	23	0	5	**64**
4 cell	27	15	0	3	**45**
8 cell	3	4	0	0	**7**
morula	11	16	0	1	**27**
blastocyte	10	18	0	0	**28**

Table 3.2 Genotyping of preimplantation embryos

3.2.3.2 Golgi Apparatus Distribution in Early Embryonic Stage

Human BIG2 is known to be involved in trafficking from the trans-Golgi network by regulating membrane association of AP-1 (adaptor protein 1) protein through activating ADP-rybosilation factor. The γ-adaptin is a subunit of the AP-1. It was demonstrated that the dominant-negative mutation in human *BIG2* results in redistribution of the γ-adaptin protein in human cell line (Shinotsuka et al., 2002). To address a question whether gene trap insertional mutagenesis in *Arfgef2* could also leads a redistribution of the AP-1 complex in preimplantation embryos. Immunocytochemistry was performed using anti γ-adaptin antibody in cultured cells and preimplantation stage embryos. However no specific signals were observed when using commercially available anti γ-adaptin tested in control embryos. It could indicate that whether the γ-adaptin is not expressed in preimplantation embryos or due to the formation of a mature trans-Golgi network. GM-130 (a Golgi marker) antibody was used in this experiment as a positive control. NIH 3T3 cells were immunostained to demonstrate the antibodies specificity. The γ-adaptin and GM-130 were localized in Golgi apparatus (Fig 3.32). In this experiment, we have observed the Golgi apparatus distribution in very early embryonic stage. Mini Golgi stacks were observed in early stage embryos including fertilized oocyte and 2 cell stage embryos. However the mini Golgi stacks were aggregated slightly in 2 cell stage embryo compare to fertilized oocytes (Fig 3.32 A and B). Any specific staining was observed on the embryos which were stained with g-adaptin antibody (E-G).

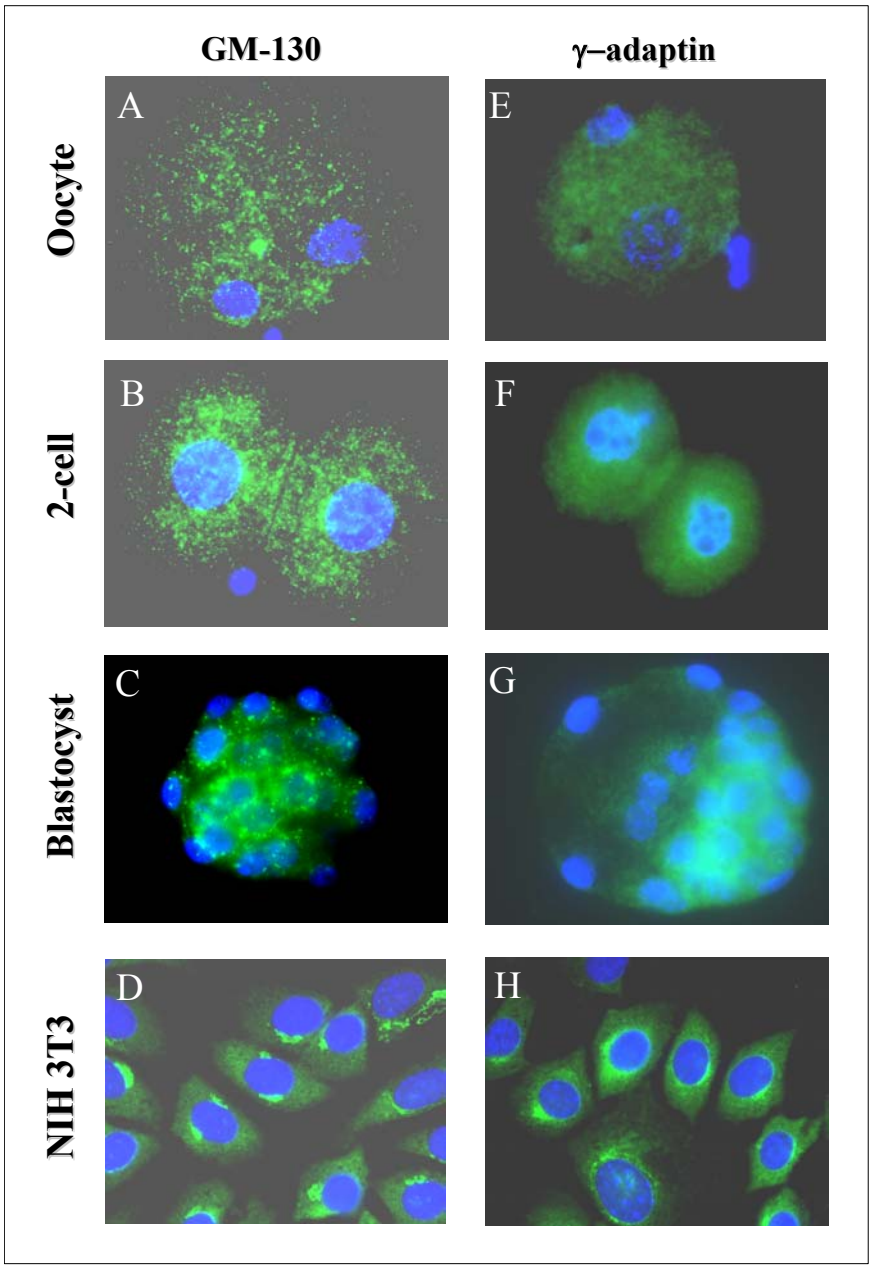


Figure 3.32: Immunofluorescence experiment demonstrating the localization of the γ-adaptin in preimplantation embryos. (A and E) fertilized oocytes containing two pronuclear. Mini Golgi stacks visible in GM-130 stained oocyte. (A and B). 2 cell stage embryos, nuclear as well as polar bodies are stained with DAPI. (C and G) blastocyst stage embryos, more complex Golgi could be visible in this stage. (D and H) NIH 3T3 GM–130 and γ-adaptin were localized in Golgi apparatus.

4. DISCUSSION

Several gene trap mouse lines were obtained and characterized for testicular gene expression using X-gal staining in male germ cells. Two of these gene trap lines (line 16 and line 8) showed strong expression in testis. For phenotypical studies, heterozygous animals were bred to obtain homozygous animals. However, no homozygous offsprings were obtained indicating that deficiency of trapped genes by in result in embryonic lethality. 5'RACE experiment allowed us to identify fragments of the trapped genes. This information able to us clones the whole cDNA of both genes. Analysis of the genomic DNA in these lines revealed that the gene trap vector has been integrated in mouse *Gtl8* and *Arfgef2* genes. The two genes were characterized further at the molecular level.

The *Gtl8* gene encodes a protein showing high similarity to the Protein 4.1 superfamily. Common sign of this family is the FERM domain and it was demonstrated that these protein could play an important role in regulation of cytoskeleton-plasma membrane interactions. Gtl8 could have a similar function because the protein was localized near to membrane, especially in cell to cell contact region. The *Gtl8* gene consists of 25 exons and was localizes on mouse chromosome 1. The expression was observed in preimplantation embryos including fertilized and unfertilized oocyte, 2- to 4-cell stage embryos, morula and blastocyst stage embryos as well as gene expression was detected ubiquitously in adult mouse tissues.

The *Arfgef2* is a mouse homology of the human *ARFGEF2* that encodes the BIG2 protein. The BIG2 protein plays a role in intracellular vesicular transport. The *Arfgef2* consists of 40 exons and the transcript was detected ubiquitously in adult mouse tissues as well as in preimplantation embryos including fertilized oocyte.

Further analysis focused on developmental stages when homozygous embryos. It was demonstrated that *Gtl8*$^{-/-}$ homozygous embryos die between morula and blastocyst stage. The *Arfgef2*$^{-/-}$ embryos die just after fertilization.

4.1 Methods for Functional Analysis of Genes

4.1.1 “Knockout” Mouse Models

"Knockout" mouse models have been developed to disrupt the expression of a targeted gene as a result of the insertion by another gene or nucleotide sequences; hence they are often called "loss of function" models.

They are created by the process of targeted mutagenesis (Majzoub et al., 1996). This relies on a "targeting" construct that has flanking sequences matching those in the gene of interest and contains markers for positive and negative selection, such as the phosphoglycerate kinase/bacterial neomycin resistance gene (*PGK-neo*) and phosphoglycerate kinase/viral thymidine kinase gene (*PGK-tk*), respectively. The construct is transfected into embryonic stem (ES) cells obtained from the inner cell mass of a blastocyst, which have the potential to differentiate into any tissue of the body. "Homologous recombination" then occurs resulting in exchange of the artificial DNA sequence for the corresponding area of genomic DNA as the DNA breaks and rejoins. The probability of this happening is increased by the presence of the matching sequences that flank the target gene. The ES cells (from agoti mice like 129) that have correctly recombined with the construct are selected out using neomycin and ganciclovir (due to the presence of the *PGK-neo* gene and absence of the *PGK-tk* gene). These cells are then injected into the host mouse blastocyst (from black mice like C57Bl) and impregnated into a foster mother. The result is a chimeric mouse (agoti and white). Chimeras that have incorporated recombinant ES cells in their germ line (for instance, the sperm), when bred with white mice will produce offspring heterozygous for the gene deletion, and interbreeding of these in turn will lead to homozygous offspring for the required mutation. The targeted mutation can thus be followed by coat colour. This strategy results in mice that develop in utero with a single gene deletion. The major problem using this approach is that the gene defect can affect prenatal and postnatal development leading to changes that may be relevant to disease expression in later life. Alternatively, there may be a degree of overlapping function in metabolic processes so that normal development is protected against the loss of the gene (redundancy). However,

absence of one gene during development may lead to upregulation of components in compensatory pathways, thereby complicating the final phenotype of the "knockout" mouse.

Therefore, "knockout" animals should be allowed to develop normally before the gene of interest is deleted. This can be achieved with the Cre-lox P technology (Gu et al., 1994) which involves the insertion of specific nucleotide sequences called lox P either side of the targeted gene. The Cre enzyme (Cre recombinase, derived from a bacteriophage P1 that infects *Escherichia coli*) recognises the sequences flanked by the lox P sequences and causes DNA cleavage at these points, thereby deleting the gene. The advantage of this system is that relatively large sections of DNA can be excised after transfection of ES cells with Cre. Also, the Cre gene itself can be inserted into the mouse genome and modified so that it can be expressed only in specific tissue by being linked to a cell specific promoter. When this mouse is crossed with a mouse that has a targeted gene flanked by lox P sequences, this gene will only be "knocked out" in those specific tissues cells. The Cre gene can also be made inducible by using a cell specific promoter that can be activated by the presence of another molecule. For instance, promoter Mx1 in T cells can be induced by interferon (IFN)α leading to the expression of the linked DNA-polymerase-β gene (Kuhn et al., 1995). Gene deletion can therefore be controlled in time and place by administration of the inducing agent (such as IFNα) to the mice leading to the controlled expression of Cre and hence cleavage of the targeted gene again only in the cells of interest. These sophisticated techniques remove the complicating issues of possible interference with developmental processes by "knocking out" the gene after development has occurred and lack of tissue specificity by only deleting the product in cells of interest. However, generation of conditional knock out mice is time consuming and the inactivation of the gene is not complete.

4.1.2 Transgenic Mouse Models

Transgenic mice have genes of interest inserted into their genome with the resultant expression or overexpression of the relevant product, and hence are often called "gain of

function" models. One of the simplest ways to study gene function in a mouse is exogenous expression of a protein in some or all tissues. For this type of genetic modification, a new piece of DNA is introduced into the mouse genome. This piece of DNA includes the structural gene of interest, a strong mouse gene promoter and enhancer to allow the gene to be expressed and vector DNA to enable the transgene to be inserted into the mouse genome. Successful integration of this DNA results in the expression of the transgene in addition to the wild type gene. Depending on the goal of the experiment, the transgenic mouse will exhibit over-expression of a non-mutated protein, expression of a dominant-negative form of a protein, or expression of a fluorescent-tagged protein. By definition, transgenesis is the introduction of DNA from one species into the genome of another species. Many of the transgenic mice in the literature fit this description well as they were generated to study the overexpression of a human protein, often an oncogene (Robertson et al., 1986). Currently, the phrase "transgenic mouse" generally refers to any mouse whose genome contains an inserted piece of DNA which is transmitted to the germ line. The piece of DNA can originate from the mouse genome or from the genome of another species. The term includes the standard transgenic mouse as well as a knockin or knockout mouse.

To generate a standard transgenic mouse, a bacterial or viral vector containing the transgene and any desired markers are injected into the pronucleus of a fertilized mouse egg. The DNA usually integrates into one or more loci during the first few cell divisions of preimplantation development. The number of copies of the transgenic fragment can vary from one to several hundred, arranged primarily in head-to-tail arrays, and the transgenic founder mice are mosaic for the presence of the transgene. Founders are very likely to have germ cells with the integrated transgene, and therefore will be able to vertically transmit the integrated gene, and all cells of the progeny of a transgenic mouse contain the transgene. This method is relatively quick, but includes the risk that the DNA may insert itself into a critical locus, causing an unexpected, detrimental genetic mutation (insertional mutation; in about 15% of cases). Alternatively, the transgene may insert into a locus that is subject to gene silencing. If the protein being expressed from the transgene causes toxicity, excessive overexpression from multiple insertions can be lethal to some tissues or even to the entire mouse. For these reasons, several independent mice lines containing the same transgene

must be created and studied to ensure that any resulting phenotype is not due to toxic gene-dosing or to mutations created at the site of transgene insertion.

4.1.3 Gene Trap Mouse Models

Gene trap (GT) strategies in mouse embryonic (ES) cells are increasingly being used for detecting patterns of gene expression (Gossler at al., 1991; Wurst et al., 1995), isolating and mutating endogenous genes (Skarnes et al., 1995), and identifying targets of signalling molecules and transcription factors (Stoykova et al., 1998). The general term gene trap refers to the random integration of a reporter gene construct (called gene trap vector) into the genome such that "productive" integration events bring the reporter gene under the transcriptional regulation of an endogenous gene. In some cases this also simultaneously generates an insertional mutagenesis. Gene trap vectors were originally developed in bacteria (Casadaban et al., 1980) and applied in *Drosophila* to identify novel developmental genes and/or regulatory sequences (Bier et al., 1989). Subsequently, a modified strategy was developed for mouse in which the reporter gene mRNA becomes fused to an endogenous transcript.

In last few years there has been a significant shift of gene trap approaches in mouse to much broader, large scale applications in the context of analysis of mammalian genomes and "functional genomics". Although at least 35,000 human genes have been sequenced and mapped, adequate expression or functional information is available for only ~15% of them. This vast sequence information contrasts with a rather limited understanding of the *in vivo* functions of these genes. Whereas DNA sequence can provide some indication of the potential functions of these genes and products, their physiological roles in the organism have to be determined by mutational analysis. Thus, the sequencing effort of the human genome project has to be complemented by efficient functional analyses of the identified genes. One potentially powerful complementation to the efforts of the human genome project would be a strategy whereby large scale random mutagenesis in mouse is combined with the rapid identification of the mutated genes. Random mutagenesis in ES cells using gene trap approach appears to be well suited for this purpose, because tens of thousands of

ES cell clones carrying random gene trap vector integrations can readily be generated *in vitro*, and cloned cell lines can be stored in liquid nitrogen. The integration sites (trapped genes) can be efficiently cloned using PCR-based approaches and then sequenced. Finally, selected integrations can be introduced into mice by the generation of germ line chimeras with the ES cell clones.

Primary advantages of gene trap strategy are easy to clone mutated gene and to study endogenous gene expression pattern. Primary disadvantage of gene trap strategy is it creates unpredictable phenotypes.

Three types of trap vector — the enhancer-, promoter- and gene-trap vector which can be introduced into the genome by electroporation or by retroviral infection— were simultaneously being developed for ES-cell mutagenesis while targeted mutagenesis mediated by homologous recombination was still in its infancy.

Enhancer-trap vector contains a minimal promoter that requires the vector to insert near to a *cis*-acting enhancer element to produce expression of a *lacZ* reporter gene. Enhancer traps were first tested by introduction into fertilized oocytes by pronuclear injection. More than 20% of the resulting transgenic lines that were generated showed restricted patterns of reporter gene expression during embryogenesis, with at least 5% of insertions being mutagenic, which indicates that this strategy can be used to trap various loci.When enhancer trapping was carried out in ES cells, this approach, on the basis of *lacZ* expression, was found to be more efficient at trapping genes than the use of pronuclear injection to introduce the enhancer-trap vectors. Cloning insertion sites from ES-cell lines that showed reporter expression *in vivo* indicated that insertions occurred adjacent to coding sequences. The mutagenicity rate from enhancer trapping has not been reported, but the nature of the insertions indicates that loss-of-function mutations from enhancer traps might be rare. Therefore, enhancer-trap vectors have not been widely exploited in the mouse.

Promoter-trap vectors consist of a promoterless reporter gene (*lacZ*) and selectable marker (*neo*). Reporter expression and mutagenesis occur when the vector inserts into an exon to generate a fusion transcript that comprises upstream endogenous exonic sequences and the reporter gene. Because transcription of the reporter requires that the vector inserts

into an exon, the mutagenicity rate of promoter-trap vectors should be very high, as indicated by there having been very few reports of hypomorphic mutations from promoter traps. In addition, because the insertion site is in transcribed DNA, cloning the insertion site will identify the disrupted gene. However, the frequency with which promoter-trap vectors insert into exons is exceedingly low, at least 200-fold lower than that of the enhancer trap. Therefore, promoter-trap vectors generally contain a selectable marker, such as the neomycin resistance gene (*neo*) or the β-galactosidase- Neo^R (β-geo) fusion marker, as a reporter, so that only ES-cell clones that contain vector insertions can be selected. This approach, however, means that only insertions into those genes that are transcriptionally active in ES cells will be selected.

Gene-trap vector contains a splice acceptor site immediately upstream of a promoterless reporter. On transcriptional activation of the endogenous *cis*-acting promoter and enhancer elements of the trapped gene, a fusion transcript is generated from the upstream coding sequence and the reporter gene, simultaneously mutating the trapped gene and reporting its expression pattern. The fusion transcript also serves as the template for PCR-based cloning by a technique called 5′RACE. The relative size of introns to exons increases the efficiency of gene traps per vector insertion by at least 50-fold in comparison to promoter traps. The efficiency of gene trapping is sufficiently high for some groups, including ours, to have developed strategies to trap genes that are not expressed in undifferentiated ES cells. The main disadvantage to using gene-trap vectors is that, because the insertion occurs in an intron, alternative splicing can sometimes take place, leading to lower levels of wild-type transcripts and often resulting in hypomorphic alleles.

4.1.4 RNAi

Gene interference is an effect similar to loss-of function mutations in organisms, as if the genes being studied were inactivated. Both sense and antisense RNA are already known to produce interference with the expression of the genes they correspond to by blocking protein synthesis. Antisense RNA is single-stranded RNA that is complementary to a

particular mRNA sequence. Sense RNA, also single-stranded, is a shorter version of a particular mRNA strand. Another mechanism for gene interference using RNA has been developed in the past few years. This process, called RNA interference (RNAi) involves double-stranded RNA (dsRNA), and was first developed for use in invertebrates, later vertebrates, and now after much doubt, has been proved to work for mammals, specifically mice.

RNAi proves to be a very good alternative to the gene "knockout" method, in which homologous recombination is used to produce an organism with a disrupted version of a gene. The knockout method has several flaws when used for developmental studies. First, it can cause complications from the contributions of maternal transcripts and their translation products. These complications can mask the effects of the gene inactivation. Knockout also eliminates the targeted gene permanently and universally throughout the embryo. This prevents scientists from studying the effects of the gene on different parts of the developing organism or how it effects different developmental stages. RNAi does not cause these problems because the targeted gene is only blocked for a number of replications, not eliminated.

The RNAi method was developed in 1998, using *Caenorhabditis elegans*, a type of nematode. This research was described in the article, "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans " (Fire et. al., 1998). This experiment was originally intended to examine the mechanisms of gene interference by sense and antisense RNA. The gene studied in this experiment was *unc-22*, which encodes a myofilament protein that is abundant in muscle cells of *C. elegans* but is not necessary for survival of the organism. A decrease in function of this gene has been shown to produce twitching, and the complete loss of function produces structural defects in muscles (Moerman, and Baillie, 1979). Adult animals were injected with sense RNA, antisense RNA, and a mixture of the two, and the resulting gene interference was studied in both the adults and their progeny. Both the sense and antisense RNA separately produced little interference activity, and large amounts of RNA had to be injected before an observable effect appeared. However, the sense-antisense mixture produced highly effective

interference with the gene activity. The interference disappeared in the progeny when the embryos reached approximately 500 cells in size, at which point cell division had caused the injected RNA to be too dilute to have any effect. The researchers suspected that the increased interference activity was due to the formation of double-stranded RNA (dsRNA). Electrophoresis revealed that the injected sense-antisense mixture was indeed mostly dsRNA, and when the dsRNA was purified, it retained the powerful interference activity.

Naturally, the researchers were interested in the specificity of the interference produced by the dsRNA, or whether or not the phenotype produced corresponded to the loss of function of the gene that the dsRNA was specific to. The specificity of the dsRNA was tested by co-injecting dsRNA segments that were not related to *unc-22*. This was found not to enhance the interference with the *unc-22* specific dsRNA. dsRNA specific for *unc-54, fem-1, gfp*, and *hlh-1* were tested, and each was found to produce the known null-mutant phenotype, or the phenotype seen in individuals lacking the gene, while single-stranded RNA specific for the same genes had no significant effect. In conclusion, this study showed that in *C. elegans*, dsRNA formed from the combination of sense and antisense RNA, is much more effective than either sense or antisense RNA for producing gene interference activity. Also, the interference produced is specific for the gene to which the dsRNA corresponds.

The RNA interference technique has also been shown to work for *Drosophila melanogaster* (Kennerdell and Carthew, 1998), *Trypanosoma brucei* (Ngo et. al., 1998), *Planarians* (Sanchez Alvarado and Newark, 1999), and plants (Waterhouse et. al., 1998). The only vertebrate in which RNAi has been studied, previous to the mouse study discussed later here, has been zebrafish, and only limited success has been demonstrated with this species. This research was conducted by Wargelius et. al. in 1999, and is described in the paper, "Double-stranded RNA induces specific developmental defects in zebrafish embryos." Wargelius, et. al. (1999) were interested in finding an effective mechanism for studying loss-of-function mutations in zebrafish and other vertebrates. They studied the effects of microinjecting dsRNA fragments corresponding to sequences of several genes for which the null-mutant effects were known. Three genes were studied. The first, *flh* (floating head), prevents the formation of the notochord and causes somites to fuse medially under the

neural tube (Talbot et. al., 1995). The second gene, *ntl* (no tail), causes the mutants to lack a tail and notochord and produces abnormal somite patterning (Halpern et. al., 1993). The third gene, *pax2.1* causes a deletion of brain tissues in the midbrain-hindbrain boundary region.

Two scientists from Cambridge, Wianny and Zernicka-Goetz, used this technique in mouse oocyte, or egg cells, and in preimplanted embryos. The article "Specific interference with gene function by double -stranded RNA in early mouse development" from the February 2000 edition of Nature Cell Biology describes their research. In this experiment, three genes were tested: *MmGFP*, *C-mos*, and *E-cadherin*. For the first gene, a line or transgenic mice was created in which the *MmGFP* gene was paternally inherited to prevent complications from maternal transcripts and translation products. GFP stands for green fluorescent protein, and that is exactly what this gene codes for. The expression of this gene causes green fluorescence in the cells that can easily be observed, making this a particularly good gene to study. Tests showed that when embryos were injected with dsRNA specific for *MmGFP*, the fluorescence was significantly diminished, indicating that expression of the gene had been blocked. Also, when the embryos were injected with dsRNA specific for *C-mos* or *E-cadherin*, no effect on the fluorescence occurred, although changes resulting from the blockage of these two genes were observed, which indicates that in mice, as in invertebrates, the interference effect is specific.

A similar test was done with dsRNA specific for *E-cadherin*. The disruption of this gene leads to compaction, a severe preimplantation defect, which prevents the embryo from developing correctly (Larue, et. al., 1994 and Riethmacher, et. al., 1995). Similar effects to the *MmGFP* study were found. dsRNA specific for *E-cadherin* resulted in compaction of the embryos, and dsRNA specific for *C-mos* or *MmGFP* did not. The final test involved *C-mos*, a maternally inherited gene which arrests maturing oocytes at metaphase during the second meiotic division. The injection of dsRNA specific for C-mos caused 63% of the injected cells to fail to maintain arrest in metaphase, whereas 1-2% of the control group failed to maintain arrest. This demonstrated that, unlike the knockout method, dsRNA can block expression of maternally provided gene products. RNA interference is important

because it allows researchers to study the effects of gene loss of function in developing embryos without the complications of the gene knockout method. The application of this mechanism to vertebrates and then to mammals is likely to provides better models for studying the effects of genes and inactivation of genes in humans. It also brings researchers closer to applying the process to humans.

4.2 Characterization of the Murine *Gtl8* Gene

4.2.1 The Murine *Gtl8* Gene is a Member of the Protein 4.1 superfamily

In this study we have identified and cloned the full-length cDNA of the *Gtl8* gene. It is expressed ubiquitously in adult mouse tissue and during preimplantation embryonic development including unfertilized oocyte.

Gtl8 is a novel mouse gene and shows similarity to the Protein 4.1 superfamily. Members of Protein 4.1 superfamily are characterized by the presence of a conserved FERM (Four.1 protein, Ezrin, Radixin, Moesin) domain at the N-terminus of the molecule (Chishti et al., 1998) and in many cases with a spectrin/actin binding domain (SABD). The name Protein 4.1 derives from the gel position of the founding member of the protein 4.1 superfamily, Protein 4.1R, following 2-dimensional SDS gel electrophoresis. The identity of the amino acid sequences in the FERM domain of deduced *Gtl8* and those of other Protein 4.1 superfamily proteins is up to 74% (3.1.5.1).

The FERM domain was originally identified in protein 4.1 (also known as band 4.1) that was isolated from human erythrocyte ghosts. *In vitro*, the purified FERM domain of human erythrocyte protein 4.1 has been reported to bind to phosphatidylinositol 4,5-biphosphate (Anderson et al., 1985), phosphatidylserine (Cohen et al., 1988), calmodulin (Kelly et al., 1991) and p55 (Marfatia et al., 1995). Furthermore, the FERM domain mediates the attachment of protein 4.1 to the plasma membrane, by binding to the cytoplasmic domains of the transmembrane proteins glycophorin A (Anderson et al., 1984), glycophorin C (Marfatia et al., 1995; Jons et al., 1992), band 3 and CD44 (Numonura et al., 1997). Evidence supporting the membrane localization of the FERM domain comes from

its occurrence in a family of membrane associated proteins termed ezrin, radxin, and moesin (ERMs). The ERMs localize just beneath the plasma membrane and its function as molecular linkers that connect cell-surface transmembrane proteins such as CD44, CD43, ICAM-2 and ICAM-3 to the actin cytoskeleton, in a variety of non-erythroid cells (Arpin et al., 1994; Tsukita et al., 1994 Serrador et al., 1997). The FERM domain is located at the N-terminus in the majority of FERM domain containing proteins (Fig 4.1). Its presence in PTP-BAS, a protein tyrosine phosphatase that associates with cytoplasmic domain of the Fas cell-surface receptor, marked the first identification of a FERM domain is located near the centre of a proteins primary structure (Fig 4.1). The presence of a FERM domain at the C-terminus was documented in myosin VIIA, mutation in this gene shows Usher-1B syndrome that causes deafness and blindness in humans.

The FERM domain is also found in the brain tumor suppressor neurofibromatosis 2 (NF2) gene product termed merlin (schwanomin) (Rouleau et al., 1993), a putative tumor suppressor gene product termed DAL-1 (Tran et al., 1999), and a cell-cell contact protein, talin (Rees et al., 1990).

In many cases, Protein 4.1 superfamily proteins also have spectrin/actin binding domain (SABD) and C-terminal domain (CTD). Protein 4.1 binds spectrin and actin and potentiates interactions of spectrin tetramers with F-actin through its SABD. However *Gtl8* does not contain such domains. But some Protein 4.1 family members including NBL4 subfamily also do not contain any spectrin actin binding domains.

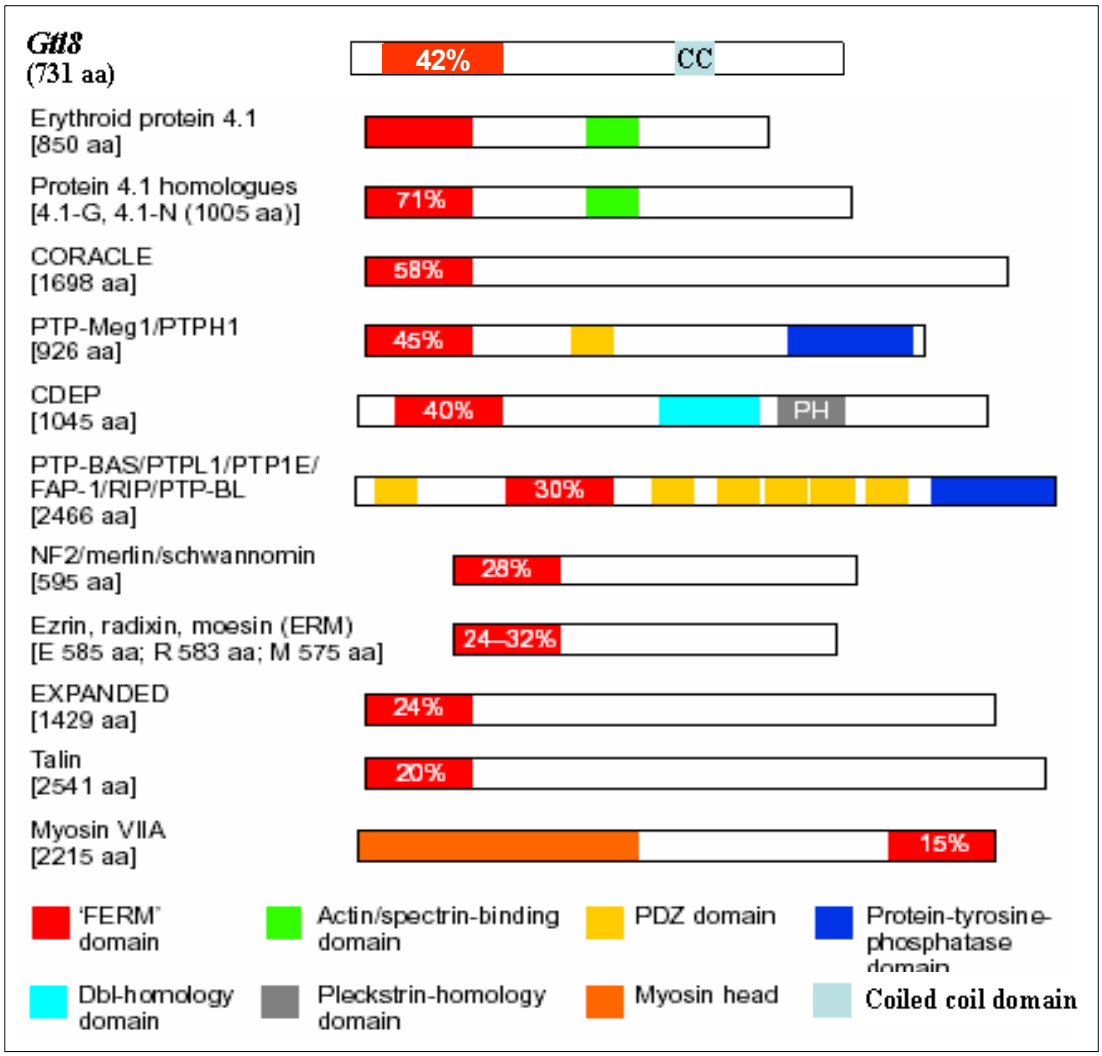


Figure 4.1: Location of the FERM domain in members of the Protein 4.1 family. The degree of amino acid sequence similarity between each FERM domain of several proteins and the FERM domain of erythrocyte protein 4.1 (founder protein) are indicated. *Gtl8* FERM domain has 42 % identity to the FERM domain of the founder protein.

The Protein 4.1 superfamily has grown significantly since the identification of original founding member, with more than 40 members identified to date. On the basis of protein sequence similarity, this Protein 4.1 superfamily can be subdivided into five subgroups: Protein 4.1 molecules, ezrin/radaxin/meosin (ERM) proteins, talin-related molecules, PTPH (protein tyrosine phosphotases) proteins and NBL4 (novel bind 4.1 like 4) (Takeuchi et al., 1994). One of the main functions of Protein 4.1 family members is the

structural stabilization of the cell membrane, which has been extensively characterised in the erythrocyte. A decrease in 4.1R expression as a consequence of a chromosomal mutation results in hereditary elliptocytosis, a disorder characterized by pronounced hemolysis, splenomegaly and abnormally shaped red blood cells. This abnormal erythrocyte phenotype has also been documented in Protein 4.1R null mice (Shi et al., 1999).

The results of our GFP experiments indicate that *Gtl8* is localized in the plasma membrane, specifically in cell to cell contact regions. (Fig 3.25 B and B´). The FERM domain is likely to be responsible for this localization. GFP-Gtl8 fusion construct contains C-terminal part of *Gtl8* and showed green fluorescence signal in the cytoplasm (Fig 3.25 C and C´). From those observations we can speculate that *Gtl8* may interact with a transmembrane protein possibly involved in cell-cell interaction.

Our preliminary result from Yeast Two Hybrid (YTH) experiment supports this speculation. For the YTH assay we used the FERM domain containing part of *Gtl8* cDNA as bait. One of the putative interaction partners found by YTH assay was *endoglin* (or CD105). Endoglin is a transmembrane glycoprotein that binds the transforming growth factor-β (TGF-β). Loss-of-function mutations in the human *endoglin* gene *ENU* cause hereditary hemorrhagic telangiectasia (HHT1), a disease characterized by vascular malformations. The HHT is an autosomal dominant vascular dysplasia that affects 1 in 10,000 individuals. Bourdeau et al. (1999) generated mice lacking *endoglin* gene. Endoglin null embryos die at gestational day 10.0-10.5 due to defects in vessel and heart development.

Putative coiled coil domain was detected in the C-terminal part of *Gtl8* by using computer analysis, and no homology with other members of the Protein 4.1family was found. Coiled-coil structures are known to participate in protein-protein interactions.

4.2.2 The *Gtl8* Gene Expression and Genomic Organization

The *Gtl8* gene expression was found to be ubiquitously in adult mouse tissue including brain, kidney, liver, heart, spleen, lung, ovary and testis (Fig 3.9). Expression of *Gtl8* in embryonic level was traced by the β-galactosidase expression using X-gal assay. Blue staining was detected in all stages of preimplantation embryos including unferilized and fertilized oocyte, 2- to blastocyst stages (Fig 3.10). Expressed high metastasis 2 (*Ehm2*) gene, other member of the Protein 4.1 superfamily which encodes the highest level of similarity (74%) with the *Gtl8* protein (3.1.5.1), was identified and cloned by screening of a cDNA library from high-metastatic M-2 cells. *Ehm2* is expressed in high-metastasic but not in low-metastatic K-1735 murine melanoma cells. The *Ehm2* gene encodes 527 amino acid residues, showing up to 41% amino acid identity with FERM domain of NF2/ERM/4.1 superfamily proteins. *Ehm2* gene was mapped to chromosome 4 and was found to be expressed in liver lung, kidney, and testis and in 7- to 17-days old embryos (Shimizu et al., 2000). However, the biochemical mechanism as well as the functional role of interaction between the *Ehm2* protein and cytoskeletal proteins is currently unknown.

The murine *Gtl8* gene consists of 25 exons in the genomic sequence ranging from 28 to 1065 bp in size (Fig 3.12). The *Gtl8* gene is a single copy gene in murine genome and it mapped to mouse chromosome 1 region E2-F (Fig 3.14) and the region is synthenic to human chromosome 2q14.1. We deduced the cDNA structure of the human *GTL8* gene from human genomic DNA and EST sequences deposited in the databases. A high level of sequence homology (83%) between the human and the mouse *Gtl8* proteins (3.1.3.3) leads to the prediction that this gene is highly conserved among different species. The presence of rat, *Canis familiaris* and *Xenopus laevis* cDNA and EST clones homologous to the mouse *Gtl8* cDNA sequence further supports this prediction.

4.2.3 Possible Underlying Causes of Embryonic Lethality of *Gtl8* in Homozygous Gene Trap Mice

It is demonstrated that *Gtl8*$^{-/-}$ embryos die between morula and blastocyst stage embryos (Table 3.1). Therefore *Gtl8* function is important for development of mouse preimplantation embryos. Sequence homology to Protein 4.1 indicates that *Gtl8* function in regulation of transmembrane and cytoskeletal protein interaction. Subcellular localization was revealed in cell to cell contact region. This observation could indicate that *Gtl8* could regulate cell adhesion molecule. It is well known that "compaction" is the first cell adhesion event at the late 8 cell stage or morula stage in mammals. *Endoglin* is the putative interaction partner of *Gtl8* could support its function in early embryonic stage lethality. Accessory protein *endoglin* interacts with transforming growth factor (TGF) receptors. *TGFβ_1*, one of the growth factors detected in the mouse embryo and reproductive tract, is a multifunctional polypeptide hormone that influences numerous physiological processes (Akhursh et al., 1991). Its expression increased gradually, peaked at 8 cell stage. The mRNA levels of *TGFβ_1* at 8 cell and morula stages were significantly higher than at other cell stages. *TGFβ_1* type I and type II receptors which interacts with *endoglin* were detected in embryos from 1 cell to blastocyst stage by immunohistochemistry.

Therefore recessive loss of function mutation of *Gtl8* may modulate the signalling cascade of adhesion molecule or may cause disruption of the interaction between transmembrane and cytoskeletal molecules which have important role for preimplantation development.

4.3 Characterization of the Murine *Arfgef2* Gene

4.3.1 The Murine *Arfgef2* Gene is Involved in Intracellular Vesicular Transport

Cells contain 10 or more chemically distinct membrane-bound compartment and the vesicular transport mediates a continuous exchange of components among them. Most transport vesicles formed from specialized coated regions of membranes are called coated vesicles.

There are two well-characterized types of coated vesicles: clathrin-coated and coatomer-coated. Clathrin coated vesicles mediate endocytosis and transport from the trans-Golgi network to prelysosomal/endosomal compartments. The major compartment of clathrin-coated vesicles is clathrin itself, a protein complex which consists of three large and three small polypeptide chains that together form a three-legged structure called the triskelion. The second major coat protein in this vesicle is a multisubunit complex called “adaptor protein” (AP) and has at least two functions: it provides mechanical force to pull the membrane into a bud, and it helps to capture membrane receptors and their bound cargo molecules. In this way a selected set of cargo molecules bound to their specific cargo receptors is incorporated into the lumen of each newly formed clathrin-coated transport vesicle (Figure 4.2). Hille-Rehfeld et al. demonstrated that AP-1 containing clathrin-coated vesicles are formed at the trans-Golgi network and transport mannose-6-phosphate receptors (MPRs), the sorting receptors for lysosomal hydrolases that contain mannose-6-phosphate residues (Hille-Rehfeld et al., 1995) to endosomes.

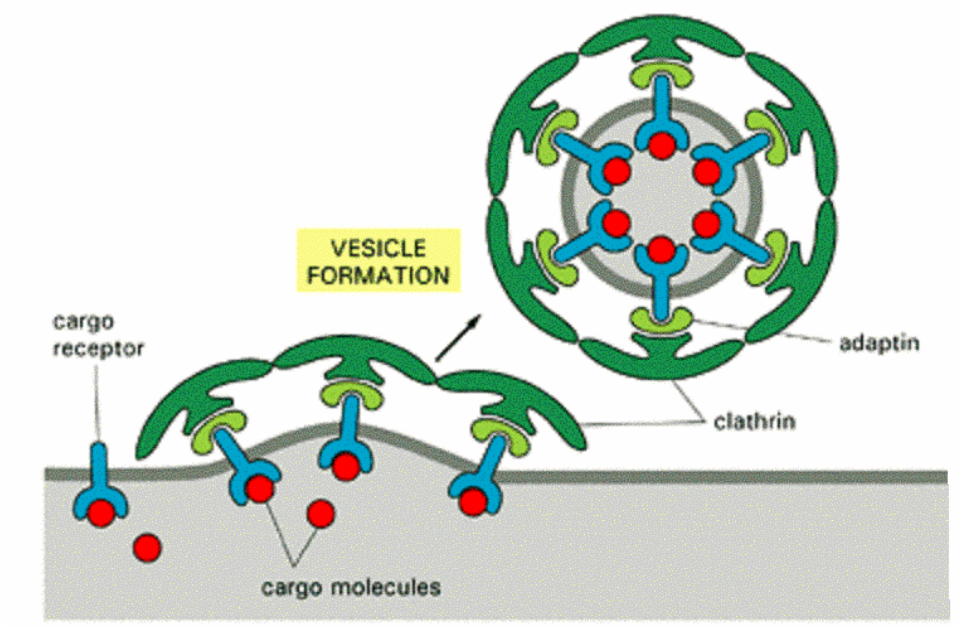


Figure 4.2: Selective transport mediated by clathrin-coated vesicles. The adaptins (AP1) bind both clathrin triskelions and cargo receptors (picture comes from Alberts et al., "Molecular biology of the Cell" 3rd edition)

The AP-1 complex consists of two ~100 kDa units, γ-adaptin and β1-adaptin, plus a 47 kDa adaptin, μ1, and 19 kDa adaptin, σ1 subunits (Stamnes et al., 1993). Of the four AP1 subunits, γ-adaptin as well as β1-adaptin interact with clathrin (Schmid et al., 1997).

ADP-ribosylation factors (ARFs) are monomeric small (~20 kDa) GTPases with fatty acid tail, that trigger budding of clathrin/AP1 coated vesicles from the trans-Golgi network (TGN), by inducing the assembly of coat proteins onto membranes and possibly by activating lipid mediators. ARF1 mediates AP-1 binding to the trans-Golgi membrane. The inactive GDP-bound form of ARF1 is located in the cytosol, whereas GTP-bound ARF1 is active and associated with Golgi membranes. The alteration of ARF1 between active and inactive forms is controlled by two types of regulatory proteins, guanine nucleotide-exchange factors (GEFs) that catalyze the replacement of GDP with GTP (Fig 4.3 A), and GTPase-activating proteins that accelerate the hydrolysis of bound GTP to GDP (Moss et al., 1998). The binding of GTP is thought to cause the ARF to expose its fatty acid tail,

which insert into lipid bilayer of the donor membrane. The tightly bound ARF recruits adaptin subunits to the trans-Golgi membrane (Fig 4.3 B)

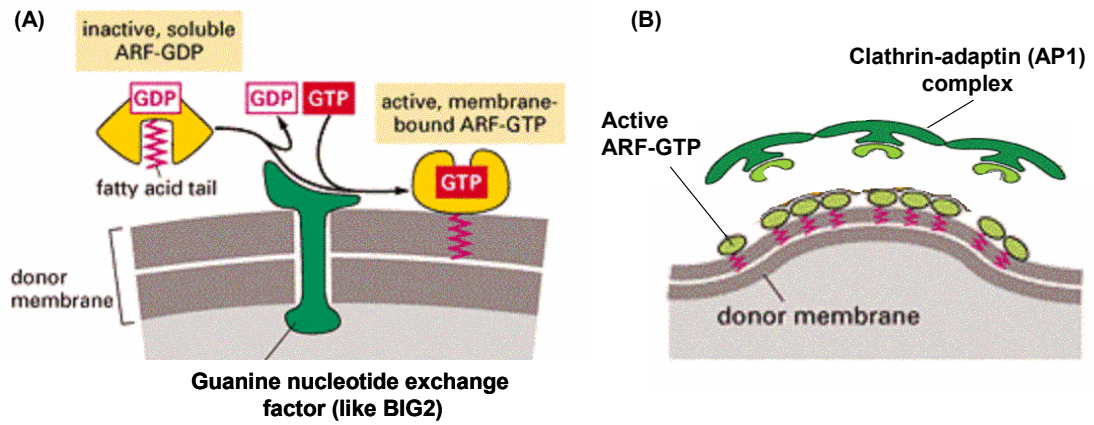


Figure 4.3: A model of clathrin coated vesicle budding. (A) Inactive soluble ARF-GDP binds to Guanine nucleotide exchange factor (like BIG2) in the donor membrane, causing the ARF to release its GDP and bind GTP. A GTP-triggered conformational change in the ARF exposes its fatty acid chain, which inserts into the donor membrane. (B) Membrane bound, active ARF-GTP mediates AP-1 binding to the membrane. This causes the membrane to form a bud. A subsequent membrane fusion event pinches off releases the coated vesicle (Fig 4.2). The drug Brefeldin A (BFA) blocks coatomer and clathrin-coated assembly by inhibiting the exchange of GDP to GTP. This blocks further vesicular trafficking from the donor membrane to the receptor membrane. (Picture comes from Alberts et al., "Molecular biology of the Cell" 3rd edition)

All ARF-GEFs identified so far posses a Sec7 domain of ~200 amino acids that is a minimum unit for the catalysis (Jackson et al., 2000). The ARF-GEF gene family can be divided into high and low molecular weight groups, on the basis of sequence similarity and functional differences. Three mammalian high molecular weight ARF-GEFs: GBF1, BIG1 and BIG2, have been identified to date and all except GBF1 have been reported to be sensitive to Brefeldin A (BFA) (Morinaga et al., 1996).

Brefeldin A (BFA) is a fungal metabolite that is known to inhibit ARF-GEFs thereby blocking assembly of ARF1 and the COPI and AP-1 complexes onto membranes and inducing formation of membrane tubules from the Golgi complex and the TGN that subsequently fuses with the endoplasmic reticulum and with endosomes, respectively (Jackson et al., 2000).

4.3.2 Analysis of the Murine *Arfgerf2*

By analyses of gene trap mouse line 16 using 5'RACE PCR method we demonstrated that GT vector integration has occurred in a gene which has 93% similarity (Fig 3.28) to the human ADP-ribosylation factor guanine nucleotide-exchange factor-2 (*ARFGEF2*). *ARFGEF2* gene encodes a Brefeldin A-inhibited guanine nucleotide-exchange protein 2 (*BIG2*). Based on this high homology between mouse and human genes, the trapped gene in the mouse line 16 was named *Arfgef2* which encodes mouse BIG2. The *BIG2* was initially purified from bovine brain cytosol associated with more that 600-kDa macromolecular complex (Yamaji et al., 1999). Identification and functional analysis of bovine and human BIG2 have been reported previously (Morinaga et al., 1996; Togawa et al., 1999; Yamaji et al., 2000; Shinotsuka et al., 2002), but there is no report concerning mouse BIG2.

Mouse BIG2 consists of 1798 amino acids (Fig 3.27) with calculated molecular weight of ~202 kDa. The cDNA for human *ARFGEF2* encodes *BIG2* protein which consists of 1785 amino acids with a molecular weight of also 202 kDa (Togawa et al., 1999). Morinaga et al. reported that the bovine BIG2 was 190 kDa in size (Morinaga et al., 1996). The deduced amino acid sequence of the mouse BIG2 revealed a region of ~185 (637-822) amino acids, which was referred to as Sec7 domain (Fig 3.27). This domain has been reported to be involved in main function of ARF-GEF as catalysing of GDP to GTP exchange. The Sec 7 domain has been found in a human protein B2-1, which has been found in T natural killer cells but not in T helper cells. In addition, the product, EMB30, a gene that is essential for proper development of the *Arabadopsis* embryo also contains a Sec 7 domain (Shevell et al., 1994).

Ubiquitous expression of the BIG2 in adult mouse organs was found by RT-PCR technique. The RT-PCR products were obtained with RNA of all mouse tissues, with exception of muscle tissue (Fig 3.29). Probably the quantity of the RNA from muscle tissue was too less to obtain visible amplification. Tissue distribution of human BIG2 was demonstrated by Northern blot. A 9.4 kb transcript was detected in human heart, brain, placenta, lung, kidney and pancreas with the *ARFGEF2* cDNA probe (Togawa et al., 1999).

In addition, the expression of the mouse *Arfgef2* was demonstrated in RNA from ES cell and oocytes by using the RT-PCR method (Fig 3.29) and in blastocyst stage embryos by using X-gal staining (Fig 3.30).

Shinotsuka et al (2002) studied that transiently transfected cells expressing epitope-tagged (HA- or Myc-tagged) BIG2 and examined subcellular localization of BIG2. When expressed at moderate levels in HeLa cells and rat hepatocytes, HA-BIG2 was localized in the perinuclear region. Similar perinuclear localization was observed for Myc-tagged BIG2. The staining for BIG2 overlapped with that for β-COP (a COPI subunit), Man II (a medial Golgi membrane protein), γ-adaptin (subunit of the AP-1 complex) or MPR (a protein recycling between the trans-Golgi and late endosomes). In contrast BIG2 was not significantly colocalized with Lamp-1 (a marker for late endosomes and lysosomes) or EEA1 (an early endosomal protein).

Chromosomal localization and genomic characterization of mouse *Arfgef2* could be predicted from mouse genomic database (NCBI). Mouse *Arfgef2* is localized on mouse chromosome 2, region of H2-H3 and *Arfgef2* gene consists of 40 exons. Human *ARFGEF2* is localized on human chromosome 20 (Togawa et al., 1999).

Several mutational and functional analyses of human BIG2 have been reported. A fungal metabolite, Brefeldin A, inhibits ARF-GEFs and leads to redistribution of coat proteins from membranes to the cytoplasm and membrane tabulation of Golgi complex and trans-Golgi network. Overexpression of BIG2 blocks BFA-induced redistribution of the AP-1 complex (Shinotsuka et al., 2002). Furthermore, the authors constructed a dominant-

negative BIG2 mutant and found that when expressed in cells it induces the redistribution of AP-1. To construct a potential dominant-negative mutant of BIG2, Glu residue in Sec7 domain was replaced by a Lys residue (E738k). Such type of mutation was first found in the deficient *emb30* allele (E658K) in *Arabidopsis* (Shevell et al., 1994). EMB30 gene was identified to be essential for proper development of the *Arabidopsis* embryo. In our gene trap mouse, integration of the GT vector has occurred in front of the Sec7 domain.

4.3.3 Possible Underlying Causes of Embryonic Lethality of *Arfgef2* in Homozygous Gene Trap Mice

In phenotypical analysis using gene trap mouse line 16, we could demonstrate that mice with a homozygous gene trap vector insertion in *Arfgef2* gene die during preimplantation embryonic development.

Arfgef2$^{-/-}$ embryos seems to die after fertilization but before development of two cell stage embryos (Table 3.2). No *Arfgef2*$^{-/-}$ embryo was found among 64 normal looking 2-cell stage embryos which derived from heterozygous intercrosses. It was found, that inhibition of BIG2 by Brefeldin A, or by dominant negative *ARFGEF2* cDNA, decreases cell proliferation *in vitro* (Shinotsuka et al., 2002). It is known that BIG2 is involved in trafficking from the trans-Golgi network by regulating membrane association of AP-1 through activating ADP-ribosylation factor. γ-Adaptin (subunit of AP-1) deficient embryos develop until day 3.5 post coitus and die during the preimplantation period, revealing that AP-1 is essential for viability (Zizioli et al., 1999). Golgi apparatus dynamics have been studied during mouse oocyte in vitro maturation (Moreno et al., 2002). In germinal vesicle oocytes the Golgi consist of a series of structures, possibly cisternal stacks, dispersed in ooplasm, but slightly more concentrated in the interior than at the cortex.

Therefore fertilized oocyte may require also the active vesicular transport ARF and AP-1 dependent manner. It is suggestive to ascribe lethality of the very early embryos on our gene trap line 16 is could be due to the missorting or any sorting of some functionally important proteins caused by the absence of BIG2.

4.4 Examples of Preimplantation Embryonic Death in Other Genes

The very early preimplantation embryo depends entirely on maternal mRNAs and proteins deposited and stored in the oocyte prior to its ovulation. If oocyte is not sufficiently equipped with maternally stored products, or if zygotic gene expression does not commence at the correct time, the embryo will die. Mutations in some genes have been found which lead to lethality in preimplantation development.

Mice with a homozygous gene trap vector insertion in *mgcRacGap* die during preimplantation development. MgcRacGap protein is a GTPase-activating protein and was identified first in human. In this mouse, maternally derived protein was found to be present in the oocyte, and *mgcRacGap* transcription start at the four-cell stage. E3.5 old *mgcRacGap*$^{-/-}$ embryos display a dramatic reduction in cell number, but undergo compaction and form blastocyst. At 3.0-3.5, binucleated blastomers with nuclei partially interconnected are frequently observed, suggesting that mgcRacGap is required for normal mitosis and cytokinesis in the preimplantaion embryo. Severe phenotype of *null* embryos indicates that mgcRacGap is functionally non-redundant and cannot be substituted by other GTPase-activating proteins (Van et al., 2000).

Targeted disruption of the *Rad51* gene leads to lethality in preimplantation embryos (Tsuzuki et al., 1996) The mouse Rad 51 gene is a mammalian homolog of the *E.coli recA* and yeast *RAD51* genes. These genes are known to be involved in homologous recombination and DNA repair. Domain II of RAD51, which is conserved in RecA-related proteins, was targeted by homologous recombination. The high level of mouse *Rad51* gene expression in testis, ovary, spleen, and thymus suggests that RAD51 protein might play an important role in meiosis (Shinohara et al., 1993)

The neurogenic gene *brainiac* was first isolated in *Drosophila melanogaster*. It interacts with members of the Notch signalling cascade. The Notch signalling pathway has been implicated in cell fate decisions in a variety of developmental aspects in *Drosophila*

melanogaster, Caenorhabtidis elegans, and in vertebrates. Targeted deletion of mouse *Brainaic 1* expression leads to preimplantation lethality. The *Brainaic 1* gene is expressed in ovarian granulosa cells in a stage specific manner. This indicated that at least in the mouse, the gene has a function during murine follicular development and modulate the adhesion properties of granulosa cells. Expression of *Brainaic 1* was also detected in brain, kidney and lung (Vollrath et al., 2001).

Study of genes which cause embryonic lethality in the mouse could be of value to understand human embryonic failure. In human, most miscarriages occur due to chromosomal aberrations. The rate of chromosomal aberrations decreases during embryonic development, from 37% in the fertilized oocyte to 0.6% in the newborn (Plachot et al., 1987). Due to chromosomal aberrations, 16.4% of preimplantation embryos fail to develop to further stages. Spontaneous abortion in human pregnancy occurs during the peri-implantation period with a frequency of ~30%. Embryonic lethality in human showed not only depends on the chromosomal aberrations but could also depend on monogenic defects.

In the following list more examples for genes are given which result in preimplantation death in case of mutation. The two genes (*Gtl8* and *Arfgef2*) which are studied in our work can be added to this list. It might be a very important although laborious work to check these genes concerning mutation in human couples suffering from unexplained infertility.

Name of the gene	Expression	Function	Day of the embryonic lethality	Reference
γ-adaptin	widely expressed	post-Golgi vesicular transport and sorting processes	E3.5	Zizioli et al., 2000
Zar1	testis, ovary, oocyte	oocyte-to-embryo transition	E0.5	Wu et al., 2003
mgcRacGAP	preimplantation embryos, intra- and extra embryonic tissue,	GTP-ase activating	E3.5	Van et al., 2001
Rad51	testis, ovary, spleen, thymus	cell proliferation	E2.5-3.5	Tsuzuki et al., 1996

Brainic1	developing central nervous system, in retina, adult hippocampus, brain, kidney, ovary	Notch signalling cascade,	E3.5-4.5	Vollrath et al., 2001
Fen1	Epidermis, small intestinal epithelium	DNA repair as well as DNA replication	E.3.5	Larsen et al., 2003
muCdc7	ES cell, testis, spleen lung, thymus, stomach, brain	DNA replication	E3.5-6.5	Kim et al., 2002
Ltb2	testis, lung, dermis, heart valves	structural role within elastic fibers	E3.5-6.5	Shipley et al., 2000
ICLn	ubiquitously expressed	regulation of the cell cycle and RNA processing	E3.5-7.5	Pu et al., 2000

4.5 Future Studies

In order to determine precise localization of Gtl8 and BIG2 proteins in cell line or preimplantation embryos, specific antibodies have to be generated. Antibodies can also be applied for genotyping of the preimplantation embryos. Any alteration of the protein localization in cell line (siRNA treated) or in blastomers of homozygous preimplantation embryos could be determined by using specific antibodies. It is possible that any cargo molecules or its receptors are arrested or mislocalized in BIG2 deficient embryo. This can be investigated by the immunofluorescence study using specific antibodies against BIG2 and for the cargo molecule. Putative cargo molecule receptors involved in ARFGEF dependent pathway is the mannose 6-phosphatase receptor (MPR46 and MPR300) which shuttles between trans-Golgi network and early endosome (Zizioli et al.,2000).

Electron microscopical analysis could be useful to detect any abnormalities in homozygous preimplantation embryos.

RNAi approach can be performed in cell culture or in preimplantation embryos to knock down both genes. *Gtl8* is believed to have role in transmembrane protein and cytoskeleton interaction. Function of this gene may involve in stable structure of the cytoskeletal

network and cell shape. It is difficult to determine the cytoskeletal network in preimplnatation embryos. But it might be easy to look the structure or morphology of cytoskeleton in cell line. Therefore siRNAi approach can be used to inhibit the Gtl8 expression in cell line and to observe possible structural or any other defects in the treated cells. There is a hind that dsRNAi approach can be used for functional analysis in preimplantation embryos including oocyte (it is discussed in 4.2.4). It is demonstrated that *Arfgef2*$^{-/-}$ mice die in oocyte stage. dsRNA treatment can be performed on fertilized oocyte using *Arfgef2* specific construct and could be compared between *Arfgef2*$^{-/-}$ from gene trap mice and *Arfgef2* dsRNA treated oocytes. For this reason, antibody specific for BIG2 is required.

Yeast two hybrid or any other alternative approach can be used to determine interaction partners. Gtl8 protein contains FERM domain which could interact with transmembrane protein like *endoglin*. C-terminal part of the Gtl8 could have a domain which might interact with cytoskeletal molecules like actin, spectrin and microtubules. Therefore it is necessary to find out more interaction partners for Gtl8 protein. The *Endoglin* and the Gtl8 interaction should be confirmed in vivo by any other alternative methods like coimmunoprecipitation, coimmunostaining and GST pull down assay.

5. Synopsis

For better understanding of the mechanisms regulating mammalian embryonic development, the isolation and characterization of new genes involved in this mechanism is important. A useful tool for studying gene function during development is to generate mutant mouse lines lacking a fully functional gene. One of the useful methods for functional analysis of genes is gene trap approach. Gene trapping in mouse embryonic stem cells offers a rapid, but essentially random, method to identify and simultaneously mutate genes expressed during mouse embryonic development.

Two gene trap mouse lines (Line 8 and Line 16) were obtained from Max Planck Institute. Initially, heterozygous mice were breeded to obtain homozygous animals. No homozygous mice were obtained from these breedings. This indicated that trapped genes in these gene trap mouse lines have crucial roles for mouse embryonic development and homozygous mice die before birth.

In the first part, *Gtl8* which was trapped by gene trap integration in gene trap mouse line 8 was identified using 5'RACE method. The gene was characterized by using molecular biological and biochemical analyses. The *Gtl8* is located on chromosome 1 E2-F of the mouse, composed of 25 exons and encodes for a protein which has high similarity to Protein 4.1 superfamily. The *Gtl8* is expressed ubiquitously in adult mouse tissue as well as in mouse preimplantation embryos including unfertilized and fertilized oocyte, 2-cell to blastocyst stage embryos. Expression in preimplantation embryos was traced by β-galactosidase using X-gal assay. Genomic organization of the *Gtl8* gene was characterized and gives opportunity to find out precise integration point of the gene trap vector integration site in this gene. Definite time of the embryonic lethality in gene trap mouse line 8 was demonstrated. *Gtl8*$^{-/-}$ embryos die between morula and blastocyst stages. Putative human homolog gene of the *Gtl8* was revealed by computer analysis using human genomic database. Mouse and human genes are 83% identical on the amino acid level. The homology between Gtl8 and Protein 4.1 was restricted to highly conserved FERM domain which localizes at the N-terminal part of the Gtl8 protein. The

FERM domains are known to interact with transmembrane proteins. Using GFP experiment in cultured cells it was shown that Gtl8 protein is localized in the membrane, specifically in cell to cell contact regions. It was demonstrated that FERM domain of the Gtl8 protein is responsible for the membrane localization. We have tried to generate antibody against Gtl8-GST fusion protein. However, it did not give any specific signals when we applied for Western blot and immunofluorescence analyses.

In the second part, analysis of the gene trap mouse line 16 and its trapped gene was detailed. The trapped gene in this line is ADP-ribosylation factor guanine nucleotide-exchange factor-2 (*Arfgef2*). The amino acid sequence of mouse *Arfgef2*, as deduced from the cDNA, shows high homology (93%) with human Brefeldin A inhibited guanine nucleotide exchange factor 2 (BIG2) which contains the conserved Sec7 domain. The BIG2 has function in intracellular vesicular transport. The *Arfgef2* expression was detected in most tissues of adult animal. RT-PCR experiments using RNA from oocyte and ES cells revealed transcripts of the *Arfgef2* gene in these embryonic stages. The *Arfgef2* expression in preimplantation embryos including oocyte, 2 cell to blastocyst stage embryos was traced by β-galactosidase activity. β-galactosidase activity was detected in blastocyst stage embryo. Function of the *Afrgef2* is essential for embryonic development because it was demonstrated that the *Afrgef2*$^{-/-}$ embryos die just after fertilization. No *Afrgef2*$^{-/-}$ embryo was found in normal looking 2 cell stage embryos indicating that *Afrgef2* is crucial for very early development of the mouse embryo.

6. REFERENCES

Adham IM, Nayernia K, Engel W. 1997. Spermatozoa lacking acrosin protein show delayed fertilization. *Mol Reprod Dev* **46**:370-6.

Adham IM, Steding G, Thamm T, Bullesbach EE, Schwabe C, Paprotta I, Engel W. 2002. The overexpression of the insl3 in female mice causes descent of the ovaries. *Mol Endocrinol* **16**:244-52.

Akhurst RJ, FitzPatrick DR, Gatherer D, Lehnert SA, Millan FA. 1990. Transforming growth factor betas in mammalian embryogenesis. *Prog Growth Factor Res* **2**:153-68.

Anderson RA, Marchesi VT. 1985. Regulation of the association of membrane skeletal protein 4.1 with glycophorin by a polyphosphoinositide. *Nature* **318**:295-8.

Arpin M, Algrain M, Louvard D. 1994. Membrane-actin microfilament connections: an increasing diversity of players related to band 4.1. *Curr Opin Cell Biol* **6**:136-41.

Bier E, Vaessin H, Shepherd S, Lee K, McCall K, Barbel S, Ackerman L, Carretto R, Uemura T, Grell E, et al. 1989. Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. *Genes Dev* **3**:1273-87.

Bonaldo P, Chowdhury K, Stoykova A, Torres M, Gruss P. 1998. Efficient gene trap screening for novel developmental genes using IRES beta-geo vector and in vitro preselection. *Exp Cell Res* **244**:125-36.

Bourdeau, A. Dumont, D. J. Letarte, M. 1999. A murine model of hereditary hemorrhagic telangiectasia. *J. Clin. Invest.* **104**: 1343-1351.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248-54.

Casadaban MJ, Chou J, Cohen SN. 1980. In vitro gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: Escherichia coli plasmid vectors for the detection and cloning of translational initiation signals. *J Bacteriol* **143**:971-80.

Chien A., Edgar D.B., Trela J.M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *J Bacteriol.* **127**: 1550-57.

Chishti AH, Kim AC, Marfatia SM, Lutchman M, Hanspal M, Jindal H, Liu SC, Low PS, Rouleau GA,Mohandas N, Chasis JA, Conboy JG, Gascard P, Takakuwa Y, Huang SC, Benz EJ Jr, Bretscher A, Fehon RG, Gusella JF, Ramesh V, Solomon F, Marchesi VT, Tsukita S, Tsukita S, Hoover KB, et al. 1998. The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane. *Trends Biochem Sci* **23**:281-282.

Chowdhury K, Bonaldo P, Torres M, Stoykova A, Gruss P. 1997. Evidence for the stochastic integration of gene trap vectors into the mouse germline. *Nucleic Acids Res* **25**:1531-6.

Cohen AM, Liu SC, Lawler J, Derick L, Palek J. 1988. Identification of the protein 4.1 binding site to phosphatidylserine vesicles. *Biochemistry* **27**:614-9.

Fire A, Siqun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver & Craig C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**:806 – 811

Forrester LM, Nagy A, Sam M, Watt A, Stevenson L, Bernstein A, Joyner AL, Wurst W. 1996. An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid in vitro. *Proc Natl Acad Sci U S A* **93**:1677-82.

Frohman MA, Dush MK and Martin GR. 1988. Rapid production of full-length cDNAsfrom rare transcripts by amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences of the USA* **85**: 8998–9002.

Gossler A, Joyner AL, Rossant J, Skarnes WC. 1989. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* **244**:463-5.

Gu, H, J. D. Marth, P. C. Orban, H. Mossmann, and K. Rajewsky. 1994. Deletion of a DNA polymerase ß gene segment in T cells using cell type-specific gene targeting. *Science* **265:**103-106.

Halpern MD, Craven SY, Cohen PL, Eisenberg RA. 1993. Regulation of anti-Sm autoantibodies by the immunoglobulin heavy chain locus. *J Immunol* **151**:7268-72.

Hille-Rehfeld A. 1995. Mannose 6-phosphate receptors in sorting and transport of lysosomal enzymes. *Biochim Biophys Acta* **1241**:177-94.

Hinners I, Tooze SA. 2003. Changing directions: clathrin-mediated transport between the Golgi and endosomes. *J Cell Sci* **116**:763-71.

Hirsch DS, Stanley KT, Chen LX, Jacques KM, Puertollano R, Randazzo PA. 2003. Arf regulates interaction of GGA with mannose-6-phosphate receptor. *Traffic* **4**:26-35.

Holzschu D, Lapierre L, Neubaum D, Mark WH. 1997. A molecular strategy designed for the rapid screening of gene traps based on sequence identity and gene expression pattern in adult mice. *Transgenic Res* **6**:97-106.

Hoover KB, Bryant PJ. 2000. The genetics of the protein 4.1 family: organizers of the membrane and cytoskeleton. *Curr Opin Cell Biol* **12**:229-34.

Hugenholtz AP, Eng VWS and Bruce WR. 1984. *Mouse News Letter* **71**: 34-35

Jackson, C.L., Casanova, J.E. 2000. Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors. *Trends Cell Biol* **10**:60-67

Jons T, Drenckhahn D. 1992. Identification of the binding interface involved in linkage of cytoskeletal protein 4.1 to the erythrocyte anion exchanger. *EMBO J* **11**:2863-7.

Kelly GM, Zelus BD, Moon RT. 1991. Identification of a calcium-dependent calmodulin-binding domain in Xenopus membrane skeleton protein 4.1. *J Biol Chem* **266**:12469-73.

Kennerdell JR, Carthew RW. 1998. Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* **95**:1017-26.

Kim JM, Nakao K, Nakamura K, Saito I, Katsuki M, Arai K, Masai H. 2002. Inactivation of Cdc7 kinase in mouse ES cells results in S-phase arrest and p53-dependent cell death. *EMBO J* **21**:2168-79.

Klemm M, Engel W. 1991. On the capacity of mouse spermatozoa for spontaneous acrosome reaction in the male and female genital tract. *Andrologia* **23**:427-33.

Kuhn R, Schwenk F, Aguet M, Rajewsky K. 1995. Inducible gene targeting in mice. *Science* **269**:1427-9.

Laird PW, Zijderveld A, Linders K, Rudnicki MA, Jaenisch R, Berns. 1991. Simplified mammalian DNA isolation procedure. *A.Nucleic Acids Res* **19**:4293.

Lankes WT, Furthmayr H. 1991. Moesin: a member of the protein 4.1-talin-ezrin family of proteins. *Proc Natl Acad Sci U S A* **88**:8297-301.

Larsen E, Gran C, Saether BE, Seeberg E, Klungland A. 2003. Proliferation failure and gamma radiation sensitivity of Fen1 null mutant mice at the blastocyst stage. *Mol Cell Biol* **23**:5346-53.

Larue L, Ohsugi M, Hirchenhain J, Kemler R. 1994. E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proc Natl Acad Sci U S A* **91**:8263-7.

Majzoub JA, Muglia LJ. 1996. Knockout mice. N *Engl J Med* **334**:904-7.

Marfatia SM, Leu RA, Branton D, Chishti AH. 1995. Identification of the protein 4.1 binding interface on glycophorin C and p55, a homologue of the Drosophila discs-large tumor suppressor protein. *J Biol Chem* **270**:715-9.

McCartney BM, Kulikauskas RM, LaJeunesse DR, Fehon RG. 2000. The neurofibromatosis-2 homologue, Merlin, and the tumor suppressor expanded function together in Drosophila to regulate cell proliferation and differentiation. *Development* **127**:1315-24.

Mendoza-Lujambio I, Burfeind P, Dixkens C, Meinhardt A, Hoyer-Fender S, Engel W, Neesen J. 2002. The Hook1 gene is non-functional in the abnormal spermatozoon head

shape (azh) mutant mouse. *Hum Mol Genet* **11**:1647-58.

Moerman, D. Baillie, D. 1979. Genetic organization in Caenorhabditis elegans: fine structure analysis of the unc-22 gene. *Genetics* **91**: 95-104

Moreno RD, Schatten G, Ramalho-Santos J. 2002. Golgi apparatus dynamics during mouse oocyte in vitro maturation: effect of the membrane trafficking inhibitor brefeldin A. *Biol Reprod* **66**:1259-66.

Morinaga N, Tsai SC, Moss J, Vaughan M. 1996. Isolation of a brefeldin A-inhibited guanine nucleotide-exchange protein for ADP ribosylation factor (ARF) 1 and ARF3 that contains a Sec7-like domain. *Proc Natl Acad Sci U S A* **93**:12856-60.

Moriniere M, Ribeiro L, Dalla Venezia N, Deguillien M, Maillet P, Cynober T, Delhommeau F, Almeida H, Tamagnini G, Delaunay J, Baklouti F. 2000. Elliptocytosis in patients with C-terminal domain mutations of protein 4.1 correlates with encoded messenger RNA levels rather than with alterations in primary protein structure. *Blood* **95**:1834-41.

Neesen J, Kirschner R, Ochs M, Schmiedl A, Habermann B, Mueller C, Holstein AF, Nuesslein T, Adham I, Engel W. 2001. Disruption of an inner arm dynein heavy chain gene results in asthenozoospermia and reduced ciliary beat frequency. *Hum Mol Genet* **10**:1117-28.

Ngo H, Tschudi C, Gull K, Ullu E. 1998. Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. *Proc Natl Acad Sci U S A* **95**:14687-92.

Nunomura W, Takakuwa Y, Tokimitsu R, Krauss SW, Kawashima M, Mohandas N.J. 1997. Regulation of CD44-protein 4.1 interaction by Ca2+ and calmodulin. Implications for modulation of CD44-ankyrin interaction. *Biol Chem* **272**:30322-8.

Ohara R, Yamakawa H, Nakayama M, Ohara O. 2000. Type II brain 4.1 (4.1B/KIAA0987), a member of the protein 4.1 family, is localized to neuronal paranodes. *Brain Res Mol Brain Res* **85**:41-52.

Perez-Ferreiro CM, Luque CM, Correas I. 2001. 4.1R proteins associate with interphase microtubules in human T cells: a 4.1R constitutive region is involved in tubulin binding. *J Biol Chem* **276**:44785-91.

Plachot M, de Grouchy J, Junca AM, Mandelbaum J, Turleau C, Couillin P, Cohen J & Salat-Baroux J. 1987. From oocyte to embryo: a model, deduced from *in vitro* fertilization, for natural selection against chromosome abnormalities. *Annales de Genetique* 30:22–32.

Pratt HP, Chakraborty J, Surani MA. 1981. Molecular and morphological differentiation of the mouse blastocyst after manipulations of compaction with cytochalasin D. *Cell* **26**:279-92.

Pu WT, Wickman K, Clapham DE. 2000. ICln is essential for cellular and early embryonic viability. *J Biol Chem* **275**:12363-6.

Rees, D.J.G., S.E. Ades, S.J. Singer, R.O. Hynes. 1990. Sequence and domain structure of talin. *Nature* **347**: 685-689

Riethmacher D, Brinkmann V, Birchmeier C. 1995. A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development. *Proc Natl Acad Sci U S A* **92**:855-9.

Robertson JA, Chen HC, Nebert DW. 1986. NAD(P) H:menadione oxidoreductase. Novel purification of enzyme cDNA and complete amino acid sequence, and gene regulation.
J Biol Chem **261**:15794-9.

Rouleau GA, Merel P, Lutchman M, Sanson M, Zucman J, Marineau C, Hoang-Xuan K, Demczuk S, Desmaze C, Plougastel B, Pulst S, Lenoir G, Bijlsma E, Fashold R, Dumanksi J, de Jong P, Parry D, Eldridge R, Aurias A, Delattre O, Thomas G. 1993. Alteration in a New Gene Encoding a Putative Membrane Organzing Protein Causes Neurofibromatosis Type 2. *Nature* **363**:515-521.

Salminen M, Meyer BI, Gruss P. 1998. Efficient poly A trap approach allows the capture of genes specifically active in differentiated embryonic stem cells and in mouse embryos. *Dev Dyn* **212**:326-33.

Sambrook J, Fritsch E.F. Maniatis T. (1989). Molecular cloning: a laboratory manual. 2nd Ed., (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NewYork).

Sanchez Alvarado, A. and Newark, P. A. 1999. Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc. Natl Acad. Sci. USA* **96**: 5049-5054

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**:5463-7.

Sata M, Moss J, Vaughan M. 1999. Structural basis for the inhibitory effect of brefeldin A on guanine nucleotide-exchange proteins for ADP-ribosylation factors. Proc Natl Acad Sci U S A **96**:2752-7.

Serrador JM, Alonso-Lebrero JL, del Pozo MA, Furthmayr H, Schwartz-Albiez R, Calvo J, Lozano F, Sanchez-Madrid F. 1997. Moesin interacts with the cytoplasmic region of intercellular adhesion molecule-3 and is redistributed to the uropod of T lymphocytes during cell polarization. *J Cell Biol* **138**:1409-23.

Sheen VL, Ganesh VS, Topcu M, Sebire G, Bodell A, Hill RS, Grant PE, Shugart YY, Imitola J, Khoury SJ, Guerrini R, Walsh CA. 2004. Mutations in ARFGEF2 implicate vesicle trafficking in neural progenitor proliferation and migration in the human cerebral cortex. *Nat Genet* **36**:69-76.

Shinotsuka C, Waguri S, Wakasugi M, Uchiyama Y, Nakayama K. 2002 Dominant-negative mutant of BIG2, an ARF-guanine nucleotide exchange factor, specifically affects membrane trafficking from the trans-Golgi network through inhibiting

membrane association of AP-1 and GGA coat proteins. *Biochem Biophys Res Commun* **294**:254-60.

Shi Zheng-Tao, Veena Afzal, Barry Coller, Dipti Patel. Joel A. Chasis. 1999. Protein 4.1R–deficient mice are viable but have erythroid membrane skeleton abnormalities. *J Clin Invest* **103**:331-340

Shimizu, K., Nagamachi, Y., Tani, M., Kimura, K., Shiroishi, T. Wakana, S., and Yokota, J. 2000. Molecular Cloning of a Novel NF2/ERM/4.1 Superfamily Gene, Ehm2, That Is Expressed in High-Metastatic K1735 Murine Melanoma Cells. *Genomics* **65** : 113-20.

Shinotsuka C, Waguri S, Wakasugi M, Uchiyama Y, Nakayama K. 2002. Dominant-negative mutant of BIG2, an ARF-guanine nucleotide exchange factor, specifically affects membrane trafficking from the trans-Golgi network through inhibiting membrane association of AP-1 and GGA coat proteins. *Biochem Biophys Res Commun* **294**:254-60.

Shinotsuka C, Yoshida Y, Kawamoto K, Takatsu H, Nakayama K. 2002. Overexpression of an ADP-ribosylation factor-guanine nucleotide exchange factor, BIG2, uncouples brefeldin A-induced adaptor protein-1 coat dissociation and membrane tubulation. *J Biol Chem* **277**:9468-73.

Shipley JM, Mecham RP, Maus E, Bonadio J, Rosenbloom J, McCarthy RT, Baumann ML, Frankfater C, Segade F, Shapiro SD. 2000. Developmental expression of latent transforming growth factor beta binding protein 2 and its requirement early in mouse development. *Mol Cell Biol.* **20**:4879-87.

Skarnes WC, Auerbach BA, Joyner AL. 1992. A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev* **6**:903-18.

Southern E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J Mol Biol.* **98**: 503-17.

Stamnes MA, Rothman JE. 1993. The binding of AP-1 clathrin adaptor particles to Golgi membranes requires ADP-ribosylation factor, a small GTP-binding protein. *Cell* **73**:999-1005.

Stoykova A, Chowdhury K, Bonaldo P, Torres M, Gruss P. 1998. Gene trap expression and mutational analysis for genes involved in the development of the mammalian nervous system. *Dev Dyn* **212**:198-213.

Shevell DE, Kunkel T, Chua NH. 2000. Cell wall alterations in the arabidopsis emb30 mutant. *Plant Cell* **12**:2047-60.

Takeuchi, K., Kawashima, A., Nagafuchi, A. and Tsukita, S. 1994. Structural diversity of band 4.1 superfamily members. *J. Cell Sci* **107**:1921 -1928.

Talbot WS, Trevarrow B, Halpern ME, Melby AE, Farr G, Postlethwait JH, Jowett T, Kimmel CB, Kimelman D. 1995. A homeobox gene essential for zebrafish notochord development. *Nature* **378**:150-7.

Togawa A, Morinaga N, Ogasawara M, Moss J, Vaughan M. 1999. Purification and cloning of a brefeldin A-inhibited guanine nucleotide-exchange protein for ADP-ribosylation factors. *J Biol Chem* **274**:12308-15.

Torres M, Stoykova A, Huber O, Chowdhury K, Bonaldo P, Mansouri A, Butz S, Kemler R, Gruss P. 1997. An alpha-E-catenin gene trap mutation defines its function in preimplantation Development. *Proc Natl Acad Sci U S A* **94**:901-6.

Tran YK, Bogler O, Gorse KM, Wieland I, Green MR, Newsham IF. 1999. A novel member of the NF2/ERM/4.1 superfamily with growth suppressing properties in lung cancer. *Cancer Res* **59**:35-43.

Tsukita S, Hieda Y, Tsukita S. 1989. A new 82-kD barbed end-capping protein (radixin) localized in the cell-to-cell adherens junction: purification and characterization. *J Cell Biol* **108**:2369-82.

Tsukita S, Oishi K, Sato N, Sagara J, Kawai A, Tsukita S. 1994. ERM family members as molecular linkers between the cell surface glycoprotein CD44 and actin-based cytoskeletons. *J Cell Biol* **126**:391-401.

Tsukita S, Yonemura S. 1997. ERM (ezrin/radixin/moesin) family: from cytoskeleton to signal transduction. *Curr Opin Cell Biol* **9**:70-5.

Tsuzuki T, Fujii Y, Sakumi K, Tominaga Y, Nakao K, Sekiguchi M, Matsushiro A,Yoshimura Y, MoritaT. 1996. Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. *Proc Natl Acad Sci USA* **93**:6236-40.

Van de Putte T, Zwijsen A, Lonnoy O, Rybin V, Cozijnsen M, Francis A, Baekelandt V, Kozak CA, Zerial M, Huylebroeck D. 2001. Mice with a homozygous gene trap vector insertion in mgcRacGAP die during pre-implantation development. *Mech Dev* **102**:33-44.

Voss AK, Thomas T, Gruss P. 1998. Efficiency assessment of the gene trap approach. *Dev Dyn* **212**:171-80.

Waterhouse PM, Graham MW, Wang MB. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci U S A* **95**:13959-64.

Wu X, Viveiros MM, Eppig JJ, Bai Y, Fitzpatrick SL, Matzuk MM. 2003. Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nat Genet* **33**:187-91.

Wurst W, Rossant J, Prideaux V, Kownacka M, Joyner A, Hill DP, 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.

Yageta M, Kuramochi M, Masuda M, Fukami T, Fukuhara H, Maruyama T, Shibuya M, Murakami Y. 2002. Direct association of TSLC1 and DAL-1, two distinct tumor suppressor proteins in lung cancer. *Cancer Res* **62**:5129-33.

Yamaji R, Adamik R, Takeda K, Togawa A, Pacheco-Rodriguez G, Ferrans VJ, Moss J, Vaughan M. 2000. Identification and localization of two brefeldin A-inhibited guanine nucleotide-exchange proteins for ADP-ribosylation factors in a macromolecular complex. *Proc Natl Acad Sci U S A* **97**:2567-72.

Zimmermann S, Steding G, Emmen JM, Brinkmann AO, Nayernia K, Holstein AF, Engel W, Adham IM. 1999. Targeted disruption of the Insl3 gene causes bilateral cryptorchidism. *Mol Endocrinol* **13**:681-91.

Zizioli D, Meyer C, Guhde G, Saftig P, von Figura K, Schu P. 1999. Early embryonic death of mice deficient in gamma-adaptin. *J Biol Chem* **274**:5385-90.

ACKNOWLEDGEMENT

I wish to express my gratitude in the first place to Prof. Dr. W. Engel for offering me great opportunities in relation to my Ph.D study in the Institute for Humangenetics under his supervision and I am very grateful to him personally for all he has done on my behalf.

Thank you most sincerely and cordially in the second place to PD. Dr. J. Neesen, at whose helpful support the work was carried out, and who throughout has gently supervised its progress.

I should greatly appreciate to PD. Dr. S. Hoyer-Fender for having accepted to be my co-referee in this study. I also thanks to Prof. Dr. M. Mühlenberg and Prof. Dr. D. Gradmann for having accepted to be my examiners.

Very many thanks to my lab colleagues and all of the co-workers in my Institute for their numerous advices, discussions and good working environment as well as to Mr. C. Muller for his efficient technical assistance, all which were a really impressive contributions in an entirely my study.

Finally, I am glad indeed to my parents, all my brothers and sister as well as relatives. Specially, I deeply thank to my dear wife O. Azbileg and my little daughter B. Anu for their invaluable help.

CURRICULUM VITAE

PERSONAL DATA

Full Name:	Batjargal Gundsambuu
E-mail:	gbachka@yahoo.com
Date of birth:	June 28, 1972
Sex:	Male
Nationality:	Mongolian
Marital status:	Married, 1 child

EDUCATIONAL BACKGROUND

1. National University of Mongolia, Fuculty of Biology, Ulaanbaatar, Mongolia 1993-2000.
2. Georg-August University of Göttingen, Germany 1998-1999 and 2000-

DEGREES

B. Sci and *M. Sci* in Biology in 1999 and 2000 from National University of Mongolia.
M.Sci. Thesis title **"**Isolation and characterization of embryonic development related gene of mice"

WORK EXPERIENCE

1. ***August 2000-***

   Ph.D student of Georg-August University of Göttingen

2. *June 1999-July 2000*
   Researcher and teaching assistant, Department of Human Genetics, Mongolian Medical University,
   Ulaanbaatar, Mongolia

3. *Junuary 1998-March 1999*
   Graduate student (guest student), Institut of Human genetics, Georg-August University of Göttingen

4. *October1991-September 1993*
   Research assistant, Department of Human Genetics, Mongolian Medical University, Ulaanbaatar, Mongolia

5. *June1989-September 1991*
   Research assistant, Department of Human Genetics, Institute of Biotechnology, Mongolian Academy of Science, Ulaanbaatar, Mongolia