

**Elucidation of *Theg* Gene Role in Spermatogenesis and Characterisation of
a Novel Spontaneous Mutation Named “*nax*” in Mouse**

Dissertation

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



**vorgelegt von
Ashraf-ul Mannan
aus Murshidabad, Indien**

Göttingen 2002

D7

Referent: Prof. Dr. W. Engel

Korreferentin: PD Dr. S. Hoyer-Fender

Tag der mündlichen Prüfungen: 29.01.03

Contents

	Page
ABBREVIATIONS.....	V
1. Introduction.....	1
1.1 Structural relationships of Sertoli cells and germ cells.....	1
1.2 Sertoli cell functions in spermatogenesis.....	3
1.3 In vitro study: Sertoli cell-germ cell interaction.....	4
1.4 Description of Testicular haploid expressed gene (<i>Theg</i>).....	6
1.5 Objectives of this study.....	7
2. Materials and Methods.....	8
2.1 Materials.....	8
2.1.1 Chemicals.....	8
2.1.2 Solutions, buffers and media.....	11
2.1.2.1 Agarose gel electrophoresis.....	11
2.1.2.2 SDS-PAGE.....	11
2.1.2.3 Frequently used buffers and solutions.....	12
2.1.3 Laboratory Materials.....	15
2.1.4 Sterilisation of solutions and equipments.....	15
2.1.5 Media, antibiotics and agar-plates.....	16
2.1.6 Antibiotics.....	17
2.1.7 IPTG / X-Gal plate.....	17
2.1.8 Bacterial strains.....	18
2.1.9 Plasmids.....	18
2.1.10 Synthetic oligonucleotide primers.....	18
2.1.11 cDNA probes.....	20
2.1.12 Eucaryotic cell lines.....	20
2.1.13 Mouse strains.....	20
2.1.14 Antibodies.....	21
2.1.15 Enzymes.....	21
2.1.16 Kits.....	21
2.1.17 Instruments.....	22
2.2 Methods.....	23
2.2.1 Isolation of nucleic acids.....	23
2.2.1.1 Isolation of plasmid DNA.....	23
2.2.1.1.1 Small-scale isolation of plasmid DNA.....	23
2.2.1.1.2 Large-scale preparation of plasmid DNA.....	23
2.2.1.1.3 Endotoxin free preparation of plasmid DNA.....	24
2.2.1.2 Isolation of genomic DNA from tissue samples.....	25
2.2.1.3 Isolation of genomic DNA from ES cells.....	25
2.2.1.4 Isolation of total RNA from tissue samples and cultured cells.....	26
2.2.2 Determination of the nucleic acids concentration.....	26
2.2.3 Gel electrophoresis.....	27
2.2.3.1 Agarose gel electrophoresis of DNA.....	27
2.2.3.2 Agarose gel electrophoresis of RNA.....	27
2.2.3.3 Polyacrylamide gel electrophoresis (PAGE) of DNA.....	28
2.2.3.4 SDS-PAGE for the separation of proteins.....	28
2.2.4 Isolation of DNA fragments after agarose gel electrophoresis.....	29
2.2.4.1 Glass Silica Method.....	29

2.2.4.2 QIAquick Gel Extraction method	29
2.2.5 Enzymatic modifications of DNA	30
2.2.5.1 Restriction enzyme digestion of DNA	30
2.2.5.2 Ligation of DNA fragments	30
2.2.5.3 TA-Cloning	30
2.2.5.4 Filling-up reaction	31
2.2.6 Preparation of competent <i>E.coli</i> bacteria	31
2.2.7 Transformation of competent bacteria	31
2.2.8 Polymerase Chain Reaction (PCR)	32
2.2.8.1 PCR amplification of DNA fragments	32
2.2.8.2 Genotyping of the knock-out mice by using PCR	33
2.2.8.3 PCR amplification of microsatellite markers	34
2.2.8.4 Reverse transcription PCR (RT-PCR)	35
2.2.9 Protein and biochemical methods	35
2.2.9.1 Isolation of total proteins	35
2.2.9.2 Isolation of nuclear proteins	36
2.2.9.3 Determination of protein concentration	37
2.2.9.4 Coupling of the synthetic peptide to BSA	37
2.2.10 Blotting techniques	37
2.2.10.1 Southern blotting of DNA onto nitrocellulose filters	37
2.2.10.2 Northern blotting of RNA onto nitrocellulose filters	38
2.2.10.3 Western blotting of protein onto PVDF membrane	38
2.2.11 “Random Prime” method for generation of ³² P labelled DNA	39
2.2.12 Non-radioactive dye terminator cycle sequencing	39
2.2.13 Hybridisation of nucleic acids	40
2.2.14 Generation of polyclonal antibody against peptide	40
2.2.14.1 Peptide analysis	40
2.2.14.2 Immunisation of rabbit	41
2.2.14.3 Determination of specificity of the polyclonal antibody	41
2.2.14.4 Affinity purification of the polyclonal antibody	42
2.2.14.4.1 Ligand coupling	42
2.2.14.4.2 Washing and deactivation	43
2.2.14.4.3 Purification	43
2.2.15 Histological techniques	43
2.2.15.1 Tissue preparation for paraffin-embedding	43
2.2.15.2 Sections of the paraffin block	44
2.2.15.3 Staining of the histological sections (Nissl staining)	44
2.2.16 Immunofluorescence staining of mouse testis	45
2.2.17 Fluorescence in situ hybridisation (FISH)	45
2.2.18 Transfection of NIH3T3 cells with the <i>Gfp-fusion</i> construct	46
2.2.19 Techniques for production of targeted mutant mice	46
2.2.19.1 Production of targeted embryonic stem cell clones	47
2.2.19.1.1 Preparation of EMFI feeder layers	47
2.2.19.1.2 Growth of ES cells on feeder layer	47
2.2.19.1.3 Electroporation of ES cells	47
2.2.19.1.4 Growing ES cells for Southern blot analysis	48
2.2.19.2 Production of chimeras by injection of ES cells into blastocyst	48
2.2.19.3 Detection of chimerism and mice breeding	49
2.2.20 Sperm isolation and counting	49
2.2.21 Sperm motility	49
2.2.22 Acrosome reaction	49

2.2.23 Techniques involved in linkage analysis.....	50
2.2.23.1 Breeding Scheme.....	50
2.2.23.2 Framework linkage map: evaluation of swept radius.....	51
2.2.23.3 A stratified approach to high-resolution mapping.....	52
2.2.23.4 Genotyping of microsatellite marker.....	53
2.2.23.4.1 Normalisation of genomic DNA.....	53
2.2.23.4.2 Pooling of PCR reaction for analysis.....	53
2.2.23.4.3 Analysis of microsatellite markers.....	54
2.2.24 Computer Analysis.....	54
3 Results.....	55
3.1 Introduction to result section.....	55
3.2 Alternative splicing of <i>Theg</i> , cDNA and deduced amino acid sequence analysis.....	56
3.3 <i>Theg</i> protein analysis.....	58
3.3.1 Expression pattern of <i>Theg</i> protein in adult mouse.....	58
3.3.2 <i>Theg</i> protein expression during different developmental stages of mouse.....	59
3.3.3 <i>Theg</i> protein expression in different mutant mice.....	60
3.3.4 <i>Theg</i> protein expression in subcellular organelles.....	60
3.3.5 <i>Theg</i> expression during male germ cell differentiation.....	62
3.3.6 Intracellular distribution pattern of <i>Theg</i>	62
3.4 Targeted inactivation of mouse <i>Theg</i> gene.....	64
3.4.1 Genomic structure and restriction digestion analysis of the <i>Theg</i> Gene.....	64
3.4.2 Construction of <i>Theg</i> gene targeting vector.....	64
3.4.3 Identification of a 5' upstream external probe.....	66
3.4.4 Electroporation of <i>Theg</i> targeting construct into ES cells and screening of ES clones for homologous recombination events.....	67
3.4.5 Generation of chimeric mice.....	68
3.4.6 Generation of <i>Theg</i> deficient mice.....	69
3.5 Analysis of <i>Theg</i> expression in knock-out mice.....	69
3.6 Phenotypic analysis of <i>Theg</i> knock-out mice.....	71
3.6.1 Statistical Analysis.....	71
3.6.2 Growth curve of testes during mouse development.....	72
3.6.3 Sperm motility analysis.....	73
3.6.4 Subcellular localisation of mutant <i>Theg</i> protein.....	75
3.7 Generation of <i>Theg</i> knock-out mice (Th14) in C3H/J background.....	75
3.8 Generation of new <i>Theg</i> knock-out mice deleting 5' end of the gene.....	77
3.8.1 Restriction digestion analysis.....	77
3.8.2 Cloning of 5' wing of the new knock-out construct.....	77
3.8.3 Cloning of 3' wing of the new knock-out construct.....	77
3.8.4 Identification of 5' external probe.....	78
3.8.5 Electroporation of new <i>Theg</i> targeting construct into ES cells and screening of ES clones for homologous recombination events.....	79
3.8.6 Generation of chimeric mice.....	81
3.8.7 Generation of <i>Theg</i> deficient mice.....	81
3.9 Characterisation of human <i>THEG</i> gene.....	81
3.9.1 Cloning of full-length <i>THEG</i> cDNA.....	81
3.9.2 Chromosome localisation and genomic organization of <i>THEG</i>	83
3.9.3 Expression and alternative splicing of the <i>THEG</i> gene.....	85
3.10 Identification and characterisation of <i>nax</i> mutant mice.....	86
3.10.1 About <i>nax</i> mutation.....	86
3.10.2 Identification of <i>nax</i> locus in mouse genome.....	86
3.10.3 Phenotypic description of <i>nax</i> mutant mice.....	88

3.10.4 Growth curve analysis	89
3.10.6 Analysis of cerebellum sagittal section of developing <i>nax</i> mice	90
3.10.7 Linkage analysis	92
3.10.7.1 Breeding strategy	92
3.10.7.2 Genome scan analysis	93
3.10.7.3 Haplotype analysis	94
3.10.7.4 Characterisation of <i>nax</i> locus	97
3.10.7.5 Identification of putative polymorphic markers	98
3.10.7.6 New haplotype analysis	98
3.10.7.7 Identification of gene responsible for <i>nax</i> phenotype	99
4. Discussion	101
4.1 Functional characterisation of <i>Theg</i> and its role in spermatogenesis	101
4.1.1 Expression analysis, cellular distribution and subcellular localisation of the <i>Theg</i> protein	101
4.1.2 Generation of <i>Theg</i> deficient mice to characterise its role in spermatogenesis	103
4.1.3 An insertional mutation in <i>Theg</i> locus caused a defect in spermatogenesis	106
4.1.4 Generation of a new <i>Theg</i> knock-out mice	109
4.1.5 Conclusion	110
4.2 Characterisation of human <i>THEG</i> gene	111
4.3 Identification and characterisation of a novel mutation named <i>nax</i>	114
4.3.1 Overview of cerebellum development	114
4.3.1.1 Genesis of cerebellum	114
4.3.1.2 Maturation of the cerebellum	116
4.3.1.3 A genetic dissection of the origin of the cerebellum	118
4.3.1.4 Migration of cerebellar neurons and neuron-glia interaction	121
4.3.1.5 Corollary remarks	124
4.3.2 Characterisation of <i>nax</i> phenotype	125
4.3.2.1 <i>nax</i> cerebellum and neurodevelopment	126
4.3.3 Identification of gene/s responsible for <i>nax</i> phenotype	127
4.3.3.1 Mapping of <i>nax</i> locus by linkage analysis	128
4.3.3.2 Determining gene order: generating a physical map	129
4.3.3.3 A new haplotype analysis	130
4.3.3.4 In pursuit for candidate gene?	131
4.3.3.5 Future perspective	132
5. Synopsis	134
References	136
Publication	156
Acknowledgements	157
Resume	159

ABBREVIATIONS

ABI	Applied Biosystem Instrument
APS	Ammonium peroxodisulfate
ATP	Adenosintriphosphate
BAC	Bacterial Artificial Chromosome
BCP	1-bromo-3-chloropropane
bp	base pair
BSA	Bovine serum albumin
°C	Degree Celsius
CASA	Computer Assisted Sperm Analysis
cDNA	complementary DNA
Cy3	indocarbocyanine
dATP	Desoxyriboadenosintriphosphate
dH ₂ O	distil Water
DAPI	Diamidino-2-phenylindole dihydrochloride
dCTP	Desoxyribocytosintriphosphate
DMSO	Dimethyl sulfoxide
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
Dnase	deoxyribonuclease
dNTP	deoxynucleotidetriphosphate
dpc	day post coitus
dT	deoxythymidinate
DTT	Dithiothreitol
E	Embryo
EDTA	Ethylene diamine tetraacetic acid
EGL	External granular layer
ES	Embryonic stem
F	Filial generation
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
g	gravity
Gfp	Green fluorescence protein
gm	gram

Contents

HEPES	N-(2-hydroxymethyl)piperazine, N'-3-propanesulfonate
HPLC	High performance liquid chromatography
hr(s)	hour(s)
IGL	Internal granular layer
IPTG	Isopropyl- β -thiogalactopyranoside
IVF	In vitro fertilisation
JL	Jackson Laboratory
kb	kilobase
LB	Luria-Bertrani
LIF	Recombinant leukaemia inhibitory factor
LPS	lipopolysaccharides
M	molarity
Mb	Mega base pair
ML	Molecular layer
MOPS	3-[N-Morpholino]-Propanesulfate
mRNA	messenger Ribonucleic acid
mg	milligram
ml	milliliter
μ l	microliter
μ m	micrometer
min	minute
NaAc	Sodium acetate
NBT	Nitro-blue tetrazolium
NCBI	National Center for Biotechnology Information
<i>Neo</i>	<i>Neomycin</i>
ng	nanogram
NLS	Nuclear localisation sequence
nm	nanometer
NTP	Nucleotidetriphosphate
OD	Optimal density
ORF	Open Reading Frame
Pa	Pascal
PAC	Bacteriophage P1 Artificial Chromosome
PAGE	Polyacrylamide Gel Electrophoresis

Contents

PCR	Polymerase chain reaction
pH	Prepondirance of hydrogen ions
pmol	pmol
PBS	Phosphatebuffersaline
PBT	Phosphatebuffersaline + Tween
PMSF	Phenylmethysulfonyl fluoride
RNA	Ribonucleic acid
Rnase	Ribonuclease
Rnasin	Ribonuclease inhibitor
rpm	revolution per minute
RT	Room temperature
RT-PCR	Reverse transcriptase-PCR
SDS	Sodium Dodecylsulfate
SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis
sec	second
SV 40	Simian Virus 40
<i>Taq</i>	<i>Thermus aquaticus</i>
TBE	Tris-Borate-EDTA-Electrophoresis buffer
TE	Tris-EDTA buffer
TEMED	Tetramethylethylene diamine
<i>Theg</i>	<i>Thesticular Hapliod Expressed gene</i>
Tris	Trihydroxymethylaminomethane
U	Unit
UV	Ultra violet
V	Voltage
w/v	weight/volume
WIBR	Whitehead Institute of Biomedical Research
X-Gal	5-bromo-4-chloro-3-indolyl- β -galactosidase

Symbol of amino acids

A	Ala	Alanine
B	Asx	Asparagine or Asparatic acid
C	Cys	Cysteine
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	Gaunosine
T	Thymidine
U	Uridine

1. Introduction

The formation of functional spermatozoa is dependent upon a remarkable cascade of intracellular and intercellular regulatory events. Diploid spermatogonia differentiate and pass through meiosis as tetraploid spermatocytes, which then undergo two cell divisions without DNA synthesis, as they become haploid germ cells. During the haploid interval of spermatogenesis, spermiogenesis, round spermatid differentiates into species-specific shaped spermatozoa. In addition to marked morphological changes, the differentiating meiotic spermatocytes move from the basal membrane compartment of the seminiferous tubule to the adluminal compartment or in structural terms from the periphery towards the lumen of the seminiferous tubule (figure 1.1). This places the subsequent stages of male germ cells inside a blood-testis barrier created by the somatic Sertoli cells. Numerous investigations support the hypothesis that the complex functional interdependence of germ cells and Sertoli cells in seminiferous tubules may play a pivotal role in the regulation of spermatogenesis (reviewed in Russell, 1993).

1.1 Structural relationships of Sertoli cells and germ cells

Anatomical relationships of Sertoli cells and germ cells were first described by Enrico Sertoli (Sertoli, 1865). Morphological studies have given rise to two very important concepts. Firstly, the seminiferous tubules are separated into two compartments by tight junctional complexes between adjacent Sertoli cells (Aoki and Fawcett, 1975; Dym and Fawcett, 1970; Hagenas et al., 1977; Neaves, 1973; Ross, 1970; Russell and Peytersen, 1985). The array of junctional complexes is completed at puberty, resulting in a seminiferous tubule barrier that separates the adluminal compartment from the basal compartment of the seminiferous tubule. Throughout spermatogenesis, developing germ cells migrate from basal lamina of the seminiferous epithelium to the adluminal compartment (de Kretser and Kerr, 1988; Setchell and Waites, 1986). All spermatogonial divisions take place in the basal compartment, while zygotene spermatocytes complete meiosis in the adluminal compartment. Finally mature spermatozoa are released into tubular lumen during spermiation (figure 1.1b). Serum macromolecules are effectively excluded from the inner or adluminal compartment, and the extracellular milieu in this compartment is composed of secretions from Sertoli cells and germ cells. The second key morphologically derived concept is that the highly organised, precisely timed sequences of germ cell development involve interaction of groups of germ cells at

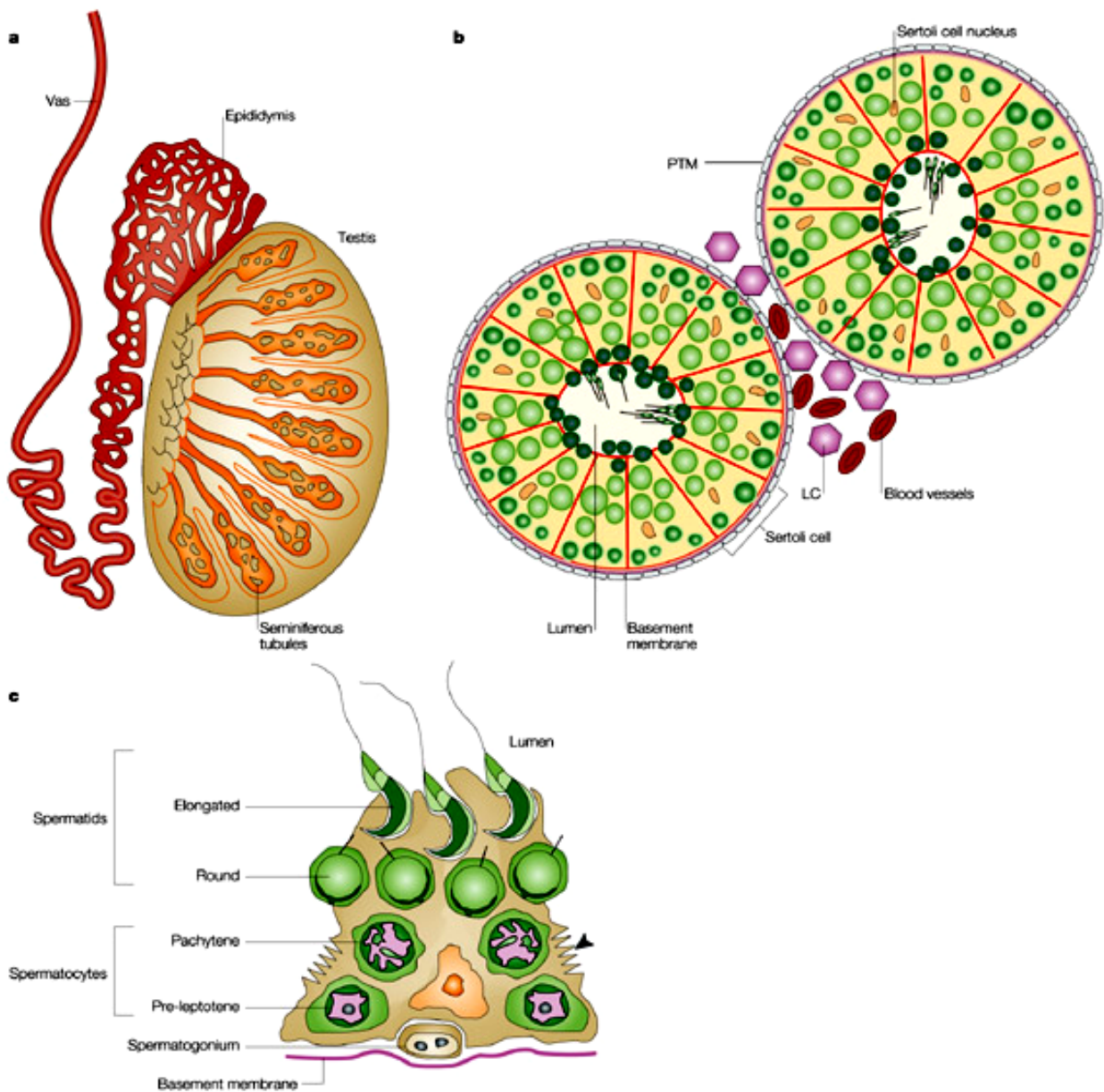


Figure 1.1: Organization of the testis. **a)** A cross-section through a testis showing the location of the seminiferous tubules, the vas deferens and the epididymis. **b)** A diagrammatic cross-section through a testicular tubule showing the germ cells (green) at different stages of maturation developing embedded in somatic Sertoli cells (each Sertoli cell is outlined in red). Leydig cells-where testosterone is synthesised-are present in the interstitium. Maturing sperm are shown in the lumen of the tubules. **c)** A single Sertoli cell with its associated germ cells. Note that tight junctions between Sertoli cells (arrowhead) define two compartments: the stem cells and the pre-meiotic cells (spermatogonia) are found on one side of the junction, whereas the meiotic (spermatocytes) and the post-meiotic (round and elongating spermatids) cells are found organized in strict order of maturation towards the lumen (cytoplasm is shown in dark green, DNA is shown in pink, Sertoli cell nucleus is shown in orange) (adapted from Cooke and Saunders, 2002).

different stages of maturation with Sertoli cells (Leblond and Clermont, 1952). Specialised junctions between Sertoli and germ cells such as ectoplasmic specialisation (Grove et al., 1990; Russell, 1977), tubulobular complex (Russell, 1979a; Russell, 1979b), desmosomes (Enders and Millette, 1988), and gap junctions (McGinley et al., 1979; Pelletier, 1988; Russell et al., 1983) are continuously disassembled and reassembled to facilitate the migration of developing germ cells. Evidence for specific recognition and adhesion between Sertoli cells and germ cells has come from several *in vitro* studies (DePhilip and Danahey, 1987; Tres and Kierzenbaum, 1983). Germ cells added to cultures attach preferentially to Sertoli cells but not to peritubular cells or cell-free culture plates. These attached germ cells are capable of extended viability and some degree of differentiation. Neither somatic cells from the testis other than Sertoli cells nor culture medium conditioned by Sertoli cells can substitute for actual attachment of germ cells to Sertoli cells (Tres and Kierzenbaum, 1983). Multiple germ cell and Sertoli cell surface molecules evidently are involved in regulation of cellular adhesion within the seminiferous epithelium. Newton and Millite (1992) identified a Sertoli cell surface polypeptide with an apparent molecular weight of 55 kDa that is involved in Sertoli cell-spermatogenic cell adhesion. D'Agostino et al. (1984) demonstrated components with apparent molecular weights of 78 and 51 kDa as potential participants in adhesion between Sertoli cells and pachytene spermatocyte cells. However, which molecules identified or as yet unidentified, are most important for Sertoli cell-germ cell adhesion is unknown.

1.2 Sertoli cell functions in spermatogenesis

Sertoli cells synthesize specific products that are necessary for germ cell survival. Together, these products form a unique environment in the adluminal compartment that is essential for meiosis and spermiogenesis (Grootegoed et al., 1984; Setchell, 1986). Some of these products are testis-specific (Cheng et al., 1986; Kissinger et al., 1982; Wilson and Griswold, 1979); other factors are homologous to serum proteins (Wright et al., 1983). Glycoprotein secreted by the Sertoli cells can be placed in several categories according to their known biochemical properties:

- 1) Transport or bioactive proteins. These are secreted in relatively high abundance and include androgen-binding protein and the metal ion transport proteins such as transferrin and ceruloplasmin. Other transport proteins include sulphated glycoproteins: SGP-1 and SGP-2.

- 2) Proteases and protease inhibitors. These inhibitors, such as plasminogen activator, are important in tissue remodelling as plasminogen activator, are important in tissue remodelling processes that occur during spermiation and movement of preleptotene spermatocytes into the adluminal compartment (Fritz et al., 1993). Cyclic protein 2 is also involved in sperm release (Wright et al., 1983).
- 3) Extracellular matrix components. Extracellular matrix components (collagen type IV, laminin, and unique proteoglycans) are among the glycoproteins that form the basement membrane between Sertoli cell and peritubular cells (Fritz and Tung, 1986; Fritz et al., 1993; Skinner et al., 1989).
- 4) Growth factors, cytokines, and hormones. Glycoproteins function as growth factors. These include Müllerian inhibiting substance, meiosis-inducing substance (Lacroix et al, 1981; Pravinen, 1982; Shabanowitz et al., 1986), activins, inhibin (Skinner, 1993), insulin-like growth factor (Sylvester, 1993; Tries and Kierzenbaum, 1983), transforming growth factors α (TGF- α) and β (TGF- β) (Bellve and Zheng, 1989; Griswold, 1993; Pravinen, 1982; Sharpe, 1993; Skinner, 1991), basic fibroblast growth factor (bFGF), interleukin (IL-1), and seminiferous growth factor (SGF). Estrogen produced by Sertoli cells may be important in adult endocrine regulation or developing seminiferous tubules (de Kretser and Kerr, 1988).
- 5) Energy metabolites. Sertoli cells can secrete lactate and pyruvate, metabolites required by germ cells (Griswold, 1993; Grootegoed et al., 1984; Jutte et al., 1981; Jutte et al., 1983; Mita and Hall, 1982) because germ cells cannot use glucose as an energy source.

Finally, intimate Sertoli cell-germ cell contact is necessary for prolonged germ cell survival and function (D'Agostino et al., 1984; Palombi et al., 1979; Tries and Kierzenbaum, 1983).

1.3 In vitro study: Sertoli cell-germ cell interaction

Our knowledge of the mediators of Sertoli cell-germ cell interactions remains, however, imperfect. Owing to its complex architecture, the seminiferous epithelium has been a difficult area for biological studies, and many of the modes of signaling between the germ cells and their somatic partners remain to be identified at the molecular level. In vitro analysis conducted on simplified culture systems may offer useful alternatives. One such system

described previously by Rassoulzadegan et al. (1993) is based on the properties of a Sertoli differentiated cell line 15P-1. These cells express a series of Sertoli cell-specific genes, and form with male germ cells in co-cultures multicellular complexes which support the progression of pachytene spermatocytes to the haploid state (Rassoulzadegan et al., 1993; Vincent et al., 1998). In previous studies, our group took advantage of the ability of 15P-1 cells to interact with male germ cells to devise a general strategy based on the mRNA differential display technique (Nayernia et al., 1999) for the identification of genes whose expression in germ cells is regulated by Sertoli cells. Differential display represents an effective method for the identification and separation of cDNAs, which are differentially expressed between various cell types or in a defined cell type under altered conditions (Liang and Pardee, 1992). We analysed changes in gene expression while co-culturing 15P-1 cells and spermatids by using the mRNA differential display technique, this involved the comparison of RNA populations from the Sertoli cell line 15P-1, from isolated spermatids, and from co-cultures of these two cell types (figure 1.2). This enabled us to identify differentially expressed genes in co-culture of Sertoli cells and germ cells (spermatids/spermatocytes). Using combinations of primer sets, five differentially expressed cDNAs were identified. One cDNA fragment identified by above method was named Testicular haploid expressed gene (*Theg*).

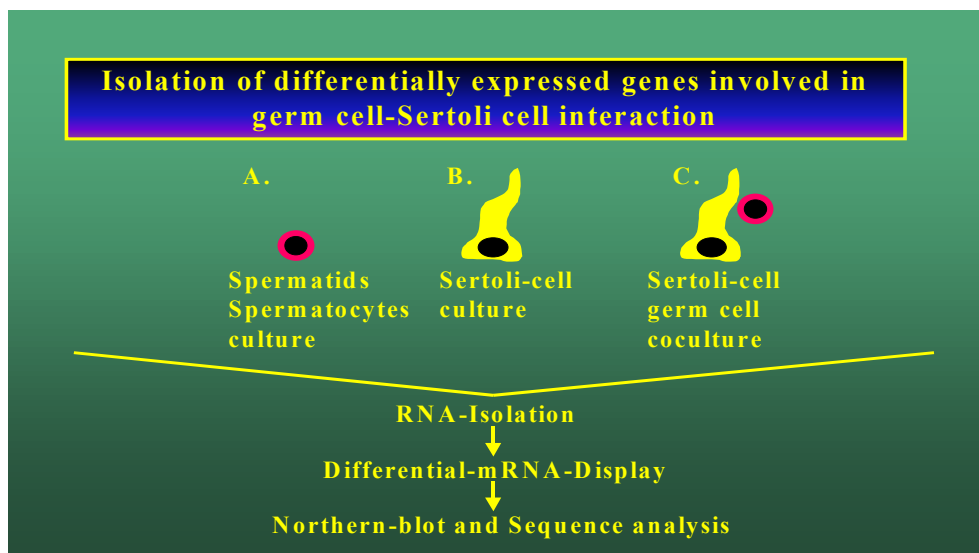


Figure 1.2: Schematic representation of the strategy used for the isolation of differentially expressed genes involved in Sertoli cell-germ cell interaction.

1.4 Description of Testicular haploid expressed gene (*Theg*)

Theg is specifically expressed in spermatids and regulated by Sertoli cells (Nayernia et al., 1999). In Sertoli cells-germ cells (spermatocytes and spermatids) coculture study, it was shown that the expression of *Theg* is maintained at basal level in germ cells only in the presence of Sertoli cells. However, if isolated germ cells (spermatocytes and spermatids) were cultured for 16 hrs alone, the expression of *Theg* was down regulated, suggesting that some factor/s from Sertoli cells is required for induction/maintenance of *Theg* expression in spermatid cells (Nayernia et al., 1999). Expression pattern analysis of *Theg* revealed a testis specific expression in mouse and expression in testis is only detectable at stage P20 of mouse, the time when haploid germ cells appear in mouse testis. When the cDNA and deduced amino acid sequences of *Theg* were compared with the Genebank/EMBL and EST database, *Theg* was shown to encode for a novel protein. In the deduced amino acid sequence of *Theg* protein, two putative nuclear targeting signals (NLS) of SV40 large T antigen type were detected (Hicks and Raikhel, 1995), suggesting that *Theg* encodes for a nuclear protein. Several phosphorylation and glycosylation sites are also present in the deduced protein sequence, which might be an indication that *Theg* undergoes post-translation modification. Further studies using gene-targeting technologies may elucidate the function of *Theg* during spermatogenesis. In recent years, much knowledge about the functions of defined genes in spermatogenesis has been gained by making use of mouse transgenic and gene knock-out models

During the generation of *Theg* knock-out mice, an intriguing but nevertheless very interesting mutation was discovered in F2 generation of knock-out mice. The inheritance of this mutation was independent of that of *Theg*-deleted locus and the mode of inheritance was autosomal recessive. Our preliminary genetic analysis revealed that this mutation arose in ES cells. The mutation was named as *nax*. *nax* mice exhibit complete lack of hair after birth and displays growth retardation during development. However there is a delayed appearance of hair but the posterior part of adult mice remains naked. Strikingly *nax* mice starts to show an ataxic gait around stage P10, they walk around incessantly in the cage. Macroscopic inspection showed a pronounced reduction in the brain size.

We decided to identify and characterise *nax* mutation in parallel with our functional analysis of *Theg* gene.

1.5 Objectives of this study

The aims and experimental approaches devised in this study can be placed in several categories, which are described below:

- 1) The expression pattern of Theg protein during spermatogenesis in mouse and also the cellular type and subcellular localisation of Theg protein in mouse testis were studied.
- 2) To elucidate the function of Theg protein and its role in spermatogenesis, we decided to disrupt the *Theg* gene by homologous recombination in mouse. Theg was identified as a novel protein and therefore for functional analysis of Theg protein's domains, two different knock-out approaches were undertaken. In the first knock-out mice the C-terminal of Theg protein, which includes exons 3-8, was deleted. In the second knock-out mice, we intended to delete the N-terminal domain of Theg protein by homologous recombination, which involved deletion of exons 1-4 of *Theg*.
- 3) Identification and cloning of the human homologue of murine *Theg*. The genomic organization, chromosomal localisation and expression pattern of *THEG* were studied.
- 4) Characterisation of *nax* mutant mice. To understand the aetiology of the abnormalities in *nax* mutant mice, we sought to determine the development stage of mouse at which the abnormalities became apparent. In an attempt to identify the gene/s responsible for *nax* phenotype, linkage analysis and physical cloning approach were undertaken.

2. Materials and Methods

2.1 Materials

2.1.1 Chemicals

Acrylamide	Serva, Heidelberg
Acetic acid	Merck, Darmstadt
Agar	Difco, Detroit, USA
Agarose	GibcoBRL, 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 powder	Boehringer, Mannheim
Bromophenol blue	Sigma, Deisenhofen
BSA	Biomol, Hamburg
Coomasie G-250	Sigma, Deisenhofen
Choloroform	Baker, Deventer
DAPI	Vector, Burlingame
Dextran sulfate	Amersham, Freiburg
Diethyl pyrocarbonate (DEPC)	Sigma, Deisenhofen
Dimethyl sulfoxide (DMSO)	Merck, Darmstadt
Dithiothreitol	Sigma, Deisenhofen
DNA Markers	GibcoBRL, Karlsruhe
dNTPs (100 mM)	GibcoBRL, Karlsruhe
Dye Terminator Mix	Applied Biosystems
Ethanol	Baker, Deventer
Ethidium bromide	Sigma, Deisenhofen
Ficoll 400	Amersham, Freiburg
FCS	Gibco/BRL, Karlsruhe

Formaldehyde	Gibco/BRL, Karlsruhe
Formamide	Fluka, Neu Ulm
Glutaraldehyde	Sigma, Deisenhofen
Glycerol	Gibco/BRL, Karlsruhe
Glycine	Biomol, Hamburg
Goat serum	Sigma, Deisenhofen
HCl	Merck, Darmstadt
H ₂ O ₂	Merck, Darmstadt
HEPES	Merck, Darmstadt
Ionophore A23187	Calbiochem, Karlsruhe
IPTG	Biomol, Hamburg
Isopropanol	Merck, Darmstadt
IVF Media	Medicult, Hamburg
KCl	Merck, Darmstadt
Lambda DNA	Boehringer, Mannheim
Methanol	Merck, Darmstadt
MgCl ₂	Merck, Darmstadt
MOPS	Merck, Darmstadt
Methyl benzoate	Fulka, Neu Ulm
β-Mercaptoethanol	Serva, Heidelberg
Mineral oil	Sigma, Deisenhofen
Na acetate	Merck, Darmstadt
Na citrate	Merck, Darmstadt
NaCl	Merck, Darmstadt
Na ₂ HPO ₄	Merck, Darmstadt
NaH ₂ PO ₄	Merck, Darmstadt
NaHCO ₃	Merck, Darmstadt
NaN ₃	Merck, Darmstadt
NaOH	Merck, Darmstadt
NBT	Boehringer, Mannheim
Orange G	Sigma, Deisenhofen

Materials and Methods

PBS	Gibco/BRL, Karlsruhe
Phosphoric acid	Merck, Darmstadt
Picric acid	Fulka, Neu Ulm
Phenol	Gibco/BRL, Eggenstein
Proteinase K	Boehringer, Mannheim
Protein marker	Biorad, Sigma
Radioactive substances: [γ - ^{32}P]-ATP [α - ^{32}P]-dCTP	Amersham, Braunschweig
Rediprime TM II	Amersham, Freiburg
RNase Inhibitor	Boehringer, Mannheim
RNA length standard	Gibco/BRL, Eggenstein
RNase away	Biomol, Hamburg
Salmon sperms DNA	Sigma, Deisenhofen
SDS	Serva, Heidelberg
Select Peptone	Gibco/BRL, Eggenstein
TEMED	Serva, Heidelberg
Triton X-100	Serva, Heidelberg
Tris	Sigma, Deisenhofen
Tween-20	Sigma, Deisenhofen
X-Gal	Biomol, Hamburg
Xylencyanol	Bio-Rad, München
Cell culture media	Gibco/BRL, Eggenstein

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

2.1.2 Solutions, buffers and media

2.1.2.1 Agarose gel electrophoresis

5X TBE buffer	450 mM Trisbase 450 mM Boric acid 020 mM EDTA (pH 8)
Glycerol loading buffer -I	10 mM Tris/HCl (pH 7.5) 10 mM EDTA (pH 8) 0.025% Bromophenol blue 0.025% Xylenecyanol 30% Glycerol
Glycerol loading buffer -II	10 mM Tris/HCl (pH 7.5) 10 mM EDTA (pH 8) 0.025% Orange G 30% Glycerol

2.1.2.2 SDS-PAGE

40% Acrylamide stock solution	Acrylamide 29.2% (w/v) Bis-acrylamide 0.8% (w/v)
10% Ammonium persulfate solution in H ₂ O	
Sample buffer (2x)	0.5 M Tris/HCl (pH 6.8) 20% Glycerol 04% SDS 10% β-Mercaptoethanol
Running buffer (5x)	025 mM Tris/HCl (pH 8.3) 192 mM Glycine 0.1% SDS

Stacking gel buffer (4x) 0.5 M Tris/HCl (pH 6.8)
0.4% SDS

Separating gel buffer (4x) 1.5 M Tris/HCl (pH 8.3)
0.4% SDS

2.1.2.3 Frequently used buffers and solutions

Denaturation solution 1.5 M NaCl
0.5 M NaOH

Denhardt's solution (50x) 1% BSA
1% Polyvinylpyrrolidon
1% Ficoll 400

Depurisation solution 0.25 N HCl

E-buffer (10x) 300 mM NaH₂ PO₄
050 mM EDTA

Elution buffer 1.5 M NaCl
20 mM Tris/HCl (pH 7.5)
01 mM EDTA

NBT- Solution 75 mg/ml NBT
70% Dimethyl formamide

BCIP-Solution 50 mg/ml BCIP
70% Dimethyl formamide

Bouin's solution 15 volumes picric acid (in H₂O)
05 volumes Formaldehyde
1 volume Acetic acid

Materials and Methods

Hybridisation solution I	5x SSPE solution 5x Denhardt's solution 0.1% SDS
Hybridisation solution II	5x SSC 5x Denhardt's solution 10% Dextran sulfate 0.1% SDS
Kinase buffer (10x)	100 mM Tris/HCl (pH 7.4) 100 mM MgCl ₂ 100 mM DTT 0.01 mM ATP
Ligation buffer (10x)	600 mM Tris/HCl (pH 7.5) 80 mM MgCl ₂ 100 mM DTT
Lysis buffer I	100 mM Tris/HCl (pH 8.0) 100 mM NaCl 100 mM EDTA 0.5% SDS
Lysis-buffer II	100 mM Tris/HCl (pH 8.0) 0.05 mM EDTA 200 mM NaCl 0.2% SDS 100 µg/ml proteinase K
10 X MOPS Buffer	41.8 gms MOPS 16.6 ml 3.0 M Sodiumacetate 20.0 ml 0.5 M EDTA in 1 liter of DEPC Water adjust pH to 6.75

Materials and Methods

Neutralisation solution	1.5 M NaCl 1.0 M Tris/HCl (pH 7.0)
PBS buffer	130 mM NaCl 007 mM Na ₂ HPO ₄ 004 mM NaH ₂ HPO ₄
PBT buffer	0.1% Tween-20 in PBS (1x)
SSC (20x)	3.0 M NaCl 0.3 M Na citrate (pH 7.0)
SSPE (20x)	0.02 M EDTA 0.20 M NaH ₂ PO ₄ 3.60 M NaCl (pH 7.0)
Stop-Mix I	95% Formamide 20 mM EDTA 0.05% Bromphenol blue 0.05% Xylene cyanol
Stop-Mix II	15% Ficoll 400 200 mM EDTA 0.1% Orange G
TE-buffer	10 mM Tris/HCl (pH 8.0) 01 mM EDTA
Washing solution	2x SSC 0.1% SDS

2.1.3 Laboratory Materials

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

Whatman blotting papers (GB 002, GB 003 and GB 004)	Schleicher and Schüll, Dassel
Cell culture flasks	Greiner, Nürtingen
Dialysis hoses	Serva, Heidelberg
Disposable filter Minisart NMI	Sartorius, Göttingen
Filter papers 0858	Schleicher and Schüll, Dassel
Aluminum folio	Merck, Darmstadt
HiTrap NHS activated columns	Amersham, Braunschweig
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

2.1.4 Sterilisation of solutions and equipments

All solutions that are not heat sensitive were sterilised at 121°C, 10⁵ Pa for 60 min in an autoclave (Webeco, Bad Schwartau). Heat sensitive solutions were filtered through a disposable sterile filter (0.2 to 0.45 µm pore size). Plastic wares were autoclaved as above. Glass wares were sterilised overnight in an oven at 220°C.

2.1.5 Media, antibiotics and agar-plates

2.1.5.1 Media for bacteria

LB Medium (pH 7.5):	1.0% Bacto-trypton
	0.5% Yeast extract
	1.0% NaCl
LB-Agar:	1.0% Bacto-trypton
	0.5% Yeast extract
	1.0% NaCl
	1.5% Agar

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

2.1.5.2 Media for cell culture

ES-cell medium:

DULBECCO's MEM (DMEM)	
0.1 mM	Non essential amino acids
1.0 mM	Sodium pyruvate
10.0 µM	β-Mercaptoethanol
2.0 mM	L-Glutamine
20%	Fetal calf serum (FCS)
1000 U/ml	Recombinant leukaemia inhibitory factor (LIF)

Fibroblast cell medium (EmFi):

DULBECCO's MEM (DMEM)	
2 mM	L-Glutamine
10%	FCS

For long time storage of the cells in liquid nitrogen, the following freezing media were used:

ES cell – freezing medium: 30% ES cell medium
 50% FCS
 20% DMSO

EmFi cells – freezing medium: 30% EmFi cell medium
 50% FCS
 20% DMSO

2.1.6 Antibiotics

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

	Master solution	Solvent	Final concentration
Ampicillin	50 mg/ml	H ₂ O	050 $\mu\text{g/ml}$
Kanamycin	25 mg/ml	H ₂ O	050 $\mu\text{g/ml}$
G 418	40 mg/ml	PBS	400 $\mu\text{g/ml}$
Gancyclovir	100 mM	PBS	002 μM
Mitomycin C	01 mg/ml	PBS	010 $\mu\text{g/ml}$
Penicillin	01 mg/ml	PBS	010 $\mu\text{g/ml}$
Streptomycin	01 mg/ml	PBS	010 $\mu\text{g/ml}$

2.1.7 IPTG / X-Gal plate

LB-agar with 50 $\mu\text{g/ml}$ ampicillin, 100 μM IPTG and 0.4% X-Gal was poured into petri dishes. The dishes were stored at 4°C .

2.1.8 Bacterial strains

<i>E. coli</i> JM 109	(Promega)
<i>E. coli</i> DH5 α	(GibcoBRL)
<i>E. coli</i> TOP10	(Invitrogen)

2.1.9 Plasmids

pBluescript SK (+/-)	(Stratagene)
pBluescript KS (+/-)	(Stratagene)
pGEM-T	(Promega)
pGEM-T Easy	(Promega)
pTKneo	Dr. N. Brose, MPI für Experimentelle Medizin, Göttingen
pZERO-2	(Invitrogen)
pEGFP-C1	(Clontech)
Lawrist 7	RZPD, Berlin

2.1.10 Synthetic oligonucleotide primers

The synthetic oligonucleotide primers used in this study were obtained either from Operon (Köln, Germany) or Roth (Karlsruhe, Germany) and dissolved in water to a final concentration of 100 pmol/ μ l.

J0:	5' GGT GTG CCC CTC CAG TGG GTG GG 3'
J1:	5' CGC CCT GTA GTG GCC TTA GAG CC 3'
J2:	5' GTC CAT CAC TGA GCG GAC TCC C 3'
J3:	5' GGG CTA TGC CTG GAT TTC CCC ACG 3'
J4:	5' GGC ATA GGG TCC CAC TCT TCC 3'
J5:	5' CCC AAG CCT TAC GTG TCA GAC 3'
J6:	5' CTG GAC CGA GCG ATT TAT CGA GG 3'
J7:	5' GCG AAC CTT CGG GGT AGC CAG 3'
J8:	5' GTG CGT TCC TGA TGC CAG TCC 3'
J9:	5' GGG GGT GGC AAG CTC ATA ATT CG 3'

J10: 5' CCC AGC GCA TGG AGG AGC AAG 3'
 J11: 5' TTC CTG GTA GAA CCG CTT GGG CC 3'
 J12: 5' CGA GGG CTA CAA CCC ATA CTA 3'
 J13: 5' GGA GAT GGC TGG GGA AGA GCT CC 3'
 J14: 5' CGC CAG CTC CAG CAG CCT CCG 3'
 J15: 5' CAA TAC AGG CAG AAC TGC AAG 3'
 J16: 5' GGG ACC GTG ATG GTC AAC GTG G 3'
 J17: 5' CCA TCC CTC GGT CCA CTC TGG AG 3'
 J18: 5' CTT GGT GAC ATC CAG GAC CTC CC 3'
 JH3: 5' CTG GCA GAG CCC AAG ATA AAC TGG CAA G 3'
 JH4: 5' GGG CAT GCT CCA GAA GTT ATC ACG AAT C 3'
 expF: 5' CCC TCT GAA ACT GTA AGC AAG 3'
 expR: 5' CCT CGG CAG CAT TCC TTA AC 3'
 GenoF: 5' AAT CTG TGT TTC CCC TGG TG 3'
 GenoR: 5' GAT CCC ATT TGG GAA GGA AG 3'
 hTHEG1: 5' CTC CTC ATC CTT CTT TTG GCC 3'
 hTHEG1a: 5' GAG GAG TAG GAA GAA AAC CGG 3'
 hTHEG2: 5' GCC ACA TGT GTC AGA CCA TAA C 3'
 hTHEG2a: 5' CGG TGT ACA CAG TCT GGT ATT G 3'
 pTKNf: 5' ATT GTC TGA GTA GGT G 3'
 pTKNr: 5' GCG CGA ATT CGA TGA TCC TGA ACG GC 3'
 pTKR: 5' AAC AGC TAT GAC CAT GAT TAC G 3'
 NeoRI: 5' AGG AGC AAG GTG AGA TGA CAG 3'
 5COS: 5' GAA TCG GCC CGT CAC CAC CGT GTG TGG 3'
 3COS: 5' GGG TGA CTC CGG TCT GGT CCG TCA GG 3'
 5NC: 5' TCA AGC TCC GGA TGT TTC ACT C 3'
 3NC: 5' GTT CAG AGA ACT CCC AGT CCG CG 3'
 5 ASTHEG1: 5' CTG GCA GAG CCC AAG ATA AAC TGG CAA G 3'
 3 ASTHEG1: 5' GGG CAT GCT CCA GAA GTT ATC ACG AAT C 3'
 5 ASTHEG2: 5' GGG ACG CTG TGG TAA GGG GTA TG 3'
 3 ASTHEG2: 5' GCC AGA CAG GAG TCG TCC TGT TG 3'
 mTHEG1: 5' GGT TGA AAT GAG CCA TCT GTC CAT CAC 3'
 mTHEG2: 5' CAG ATG GGC GTT GTC CTG TTA TTG TTG 3'
 expol: 5' GTG CCA TTG TAT GTT GGA AAT ATG 3'

mTHEGf: 5' GCT GAG GAG GGA GTG AGG TTA AAG 3'
mTHEGr: 5' GGG GAG ATG TAG TAT GGG TTG TAG 3'
T7: 5' TAA TAC GAC TCA CTA TAG GG 3'
T3: 5' ATT AAC CCTT CAC TAA AG 3'
SP6: 5' AGG TGA CAC TAT AGA ATA C 3'
Poly T: TTTTTTTTTTTTTTTTTTT

Primer sequence for microsatellite markers were taken from Mouse Genome Informatics, Jackson Laboratory (<http://www.informatics.jax.org>). The primers for putative microsatellite markers were designed by Primer3 program provided by Whitehead Institute of Biomedical Research (http://www-genome.wi.mit.edu/cgi_bin/primer/primer3_www.cgi). Due to space constraint primer sequence of putative markers and candidate genes are not provided here.

2.1.11 cDNA probes

<i>EF-2</i> cDNA	Hanes et al.,1992
Human 3' <i>THEG</i> cDNA	Image clone
O1 mouse <i>Theg</i> cDNA	In this study
U1 mouse <i>Theg</i> cDNA	In this study
<i>β-actin</i> cDNA	Clontech

2.1.12 Eucaryotic cell lines

NIH3T3 mouse fibroblast cell line, S.A. Aaronson, Bethesda, U.S.A.

RI mouse embryonic stem cell line, Dr. A. Nagi, Toronto, Canada

2.1.13 Mouse strains

Mice strains C57BL/6J, 129X1/SvJ, C3H/J and DBA/2J were initially ordered from Charles River Laboratories, Wilmington, USA and further inbred in animal facility of Institute of Human Genetics, Göttingen.

2.1.14 Antibodies

Mouse monoclonal against γ -tubulin, goat anti-rabbit-IgG-conjugated to alkaline phosphatase and Cy3-conjugated goat anti-rabbit antibodies were purchased from Sigma Deisenhofen. Rabbit anti-mouse Theg polyclonal antibody was generated in present study.

2.1.15 Enzymes

Restriction enzymes (with supplied buffers)	(GibcoBRL, NEB)
Klenow Fragment	(GibcoBRL)
Mung bean exonuclease	(GibcoBRL)
Proteinase K	(Sigma)
Platinum <i>Taq</i> polymerase	(GibcoBRL)
<i>Pfu</i> polymerase	(Stratagene)
Rnase A	(Qiagen)
Rnase H	(GibcoBRL)
Rnase inhibitor	(GibcoBRL)
Superscript-II	(GibcoBRL)
<i>Taq</i> polymerase	(GibcoBRL)
T4 DNA ligase	(Promega)
T4 RNA ligase	(Invitogen)
Tyrpsin	(GibcoBRL)

2.1.16 Kits

BigDye Terminator Cycle Sequencing Ready Reaction Kit	(Applied Biosystems)
DYEnamic ET-Terminator mix	(Amersham Pharmacia)
Endo Free Plasmid Maxi Kit	(Qiagen)
JETsorb Gel Extraction Kit	(Genomed)
JETstar Plasmid MIDI Kit	(Genomed)
Large Construct Plasmid Kit	(Qiagen)
Megaprime DNA Labeling Kit	(Amersham Pharmacia)
Maxi Plasmid Kit	(Qiagen)

Mega Plasmid Kit	(Qiagen)
Mini Plasmid Kit	(Qiagen)
PCR Purification Kit	(Qiagen)
QIAquick Gel Extraction Kit	(Qiagen)
5`RACE Kit	(GibcoBRL)
5`and 3`RACE Kit	(Invitrogen)
RNA Easy Kit	(Qiagen)
Rediprime™ II Random Prime Labeling System	(Amersham Pharmacia)

2.1.17 Instruments

ABI PRISM 377 DNA Sequencer	(Applied Biosystem)
ABI 3100 Genetic Analyser	(Applied Biosystem)
Microscope BX60	(Olympus)
GeneAmp PCR System 9600	(Perkin Elmer)
Microtiterplate-Photometer	(BioRad)
Molecular Imager FX	(BioRad)
Phosphoimager Screen	(Kodak)
Semi-Dry-Blot Fast Blot	(Biometra)
Spectrophotometer Ultraspec 3000	(Amersham Pharmacia)
SpeedVac concentrator SVC 100H	(Schütt)
Thermomixer 5436	(Eppendorf)
Turboblotter™	(Schleicher & Schüll)
UV Stratalinker™ 1800	(Leica)
Video-Documentationsystem	(Herolab, Heidelberg)
X-Ray Automatic Processor Curix 60	(Agfa)
CASA system	(Hamilton Thorne Research)

2.2 Methods

2.2.1 Isolation of nucleic acids

2.2.1.1 Isolation of plasmid DNA

(Sambrook et al., 1989)

2.2.1.1.1 Small-scale isolation of plasmid DNA

A single *E.coli* colony was inoculated in 5 ml of LB medium with the appropriate antibiotic and incubated in a shaker for 16 hrs at 37°C with a speed of 160 rpm. 1 ml of this saturated culture was used for making glycerol stock and rest of the culture was centrifuged at 5000xg for 15 min. The pellet was resuspended in 150 µl of solution P1. The bacterial cells were lysed with 300 µl of P2 solution and then neutralised with 200 µl. The precipitated solution was incubated on ice for 15 min, and centrifuged at 13000xg at 4°C. The supernatant was transferred into a new tube, and 1 ml of 100% ethanol was added to precipitate the DNA. It was then stored in ice for 15 min, centrifuged at full speed for 20 min, and finally the pellet was washed with 70% ethanol and after air-drying dissolved in 30 µl of TE buffer (adapted from Birnboim and Doly, 1979).

<u>P1</u> :	50 mM	Tris/HCl (pH 8.0)
	10 mM	EDTA
	100 µg/ml	RNase A
<u>P2</u> :	200 mM	NaOH
	1%	SDS
<u>P3</u> :	3.0 M	Potassium acetate, pH 5.5

2.2.1.1.2 Large-scale preparation of plasmid DNA

A single clone was inoculated in 2 ml LB medium with appropriate antibiotic as a pre-culture for 8 hrs in 37°C shaker. In 100 ml LB medium with appropriate antibiotic this pre-culture was added in a dilution of 1/100 fold and incubated overnight at 37°C with shaking. The

saturated culture was centrifuged at 6000xg for 15 min. The pellet was resuspended in 5 ml of solution P1 and cells were lysed with P2 and P3 as described above. The precipitated solution was centrifuged at 20000xg for 30 min at 4°C. Meanwhile, the column (Qiagen-tip) that was provided with the midi preparation kit was equilibrated with 10 ml of QBT solution. After centrifugation the lysate was poured into this equilibrated column thus allowing the DNA to bind with the resin present in the bed of the column. The column was then washed twice with 10 ml of solution QC. Finally, the DNA was eluted with 5 ml of QF solution. To precipitate the DNA, 3.5 ml of isopropanol was added and mixed thoroughly and centrifuged at 14000xg for 30 min at 4°C. The DNA pellet was washed with 70% ethanol and dissolved in 100 µl of TE.

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

2.2.1.1.3 Endotoxin free preparation of plasmid DNA

Endotoxins, also known as lipopolysaccharides or LPS, are cell membrane components of Gram-negative bacteria (e.g., *E.coli*). During lysis of bacterial cells for plasmid preparations, endotoxin molecules are released from the outer membrane of bacterial cells into the lysate. Endotoxins strongly influence transfection of DNA into primary cells and sensitive cultured cells like embryonic stem (ES) cell culture, increased endotoxin levels lead to sharply reduced transfection efficiencies. Endofree plasmid preparation kit integrates endotoxin removal into standard plasmid preparation procedure. The neutralised bacterial lysate was filtered through a QIAfilter cartridge (provided in kit) and incubated on ice with a specific Endotoxin Removal buffer (patented by Qiagen). The endotoxin removal buffer prevents LPS molecules

from binding to the resin in the columns (Qiagen-tips) thus allowing purification of DNA containing less than 0.1 endotoxin units per μg plasmid DNA.

2.2.1.2 Isolation of genomic DNA from tissue samples
(Laird et al., 1991)

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

The method employed was the same as that of Laird et al. (1991). 1 to 2 cm of the tail from a mouse was incubated in 700 μl of lysis buffer containing 35 μl Proteinase K (10 $\mu\text{g}/\mu\text{l}$) at 55°C for overnight in Thermomixer 5436. To the tissue lysate, 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 and incubated at 60° C for 15 min.

2.2.1.3 Isolation of genomic DNA from ES cells

Lysis buffer II:	100 mM	Tris/HCl (pH 8.5)
	5 mM	EDTA
	200 mM	NaCl
	100 $\mu\text{g}/\text{ml}$	Proteinase K
	0.2%	SDS

To isolate the DNA from the ES cells, cells in a 24 well plate were washed with PBS and incubated overnight in 500 μl lysis buffer II at 55° C. Equal volume of isopropanol was added and mixed for 15 min to precipitate the DNA. After washing with 70% ethanol, the DNA was transferred into a microcentrifuge cup containing 80 μl TE buffer and incubated at 60° C for 15 min.

2.2.1.4 Isolation of total RNA from tissue samples and cultured cells

Total RNA isolation reagent is an improved version of the single-step method for total RNA isolation. The composition of reagent includes phenol and guanidine thiocyanate in a monophasic solution. 100-200 mg tissue sample was homogenised in 1-2 ml of total RNA Reagent by using a glass-teflon homogeniser. The sample volume should not exceed 10% of the volume of the reagent used for the homogenisation. To isolate total RNA from cultured cells, 350 µl of reagent was added to the petri dish (6 cm diameter). Cells were homogenised with a rubber stick and the lysate was transferred into a microcentrifuge tube. The homogenate was incubated at 4°C for 5 min to permit the complete dissociation of nucleoprotein complexes. Then, 0.2 ml of chloroform was added, mixed vigorously, and stored at 4°C for 10 min. After centrifuging at 12000xg for 15 min at 4°C, the colourless upper aqueous phase was transferred into a new tube. The RNA was precipitated by adding 0.5 ml of isopropanol to aqueous solution. Finally, the pellet was washed twice with 75% ethanol, and dissolved in 80-100 µl DEPC-H₂O.

2.2.2 Determination of the nucleic acids concentration

The concentration of nucleic acids was determined spectrophotometrically by measuring absorption of the samples at 260 nm and 280 nm. The quality of nucleic acids i.e. contamination with salt and protein was checked by the measurements at 230, 280, and 320 nm. The concentration was calculated according to the formula:

$$C = (E_{260} - E_{280}) \times fc$$

C = concentration of sample (µg/µl)

E₂₆₀ = ratio of extinction at 260 nm

E₂₈₀ = ratio of extinction at 280 nm

f = dilution factor

c = concentration (standard) / absorption (standard)

for double stranded DNA : c = 0.05 µg/µl

for RNA : c = 0.04 µg/µl

for single stranded DNA : c = 0.03 µg/µl

2.2.3 Gel electrophoresis

Gel electrophoresis is the technique by which mixture of charged macromolecules, especially nucleic acids and proteins, are separated in an electrical field according to their mobility which is directly proportional to macromolecule's charge to mass ratio.

2.2.3.1 Agarose gel electrophoresis of DNA

Agarose gels are used to electrophorese nucleic acid molecules from as small as 50 bases to more than 50 kilobases, depending on the concentration of the agarose and the precise nature of the applied electrical field (constant or pulse). Usually, 1 gm of agarose was added in 100 ml 0.5x TBE buffer, and boiled in the microwave to dissolve the agarose, then cooled down to about 60°C before adding 3 µl ethidium bromide (10 mg/ml). This 1% agarose gel was poured into a horizontal gel chamber.

2.2.3.2 Agarose gel electrophoresis of RNA (Hodge, 1994)

Single-stranded RNA molecules often have complementary regions that can form secondary structures. Therefore, RNA was run on a denaturing agarose gel that contained formaldehyde, and before loading, the RNA was pre-treated with formaldehyde and formamide to denature. 1.25 gm of agarose was added in 100 ml of 1x MOPS Buffer and dissolved by heating in microwave. After cooling it to about 50°C, 25 ml of formaldehyde (37%) was added, stirred and poured into a horizontal gel chamber.

RNA samples were treated as follows:

10-20 µg	RNA
2 µl	10 x MOPS Buffer
3 µl	Formaldehyde
8 µl	Formamide (40%)
1.5 µl	Ethidium bromide

Samples were denatured at 65°C for 10 min and chilled on ice before loading into the gel. The gel was run at 40 V at 4°C for about 12 hrs.

2.2.3.3 Polyacrylamide gel electrophoresis (PAGE) of DNA

Polyacrylamide gel electrophoresis was employed to separate and analyse very small DNA fragments (70-200 bp). The percentage of acrylamide (7-12%) determines the resolving property of the gel. A 10% of gel was prepared as follows:

2.5 ml	40% stock solution
2.5 ml	5x TBE buffer
150 μ l	APS (10% w/v)
15 μ l	TEMED
up to 10 ml with water	

APS in presence of TEMED generates free radicals, which initiate the polymerisation of acrylamide. The gel was poured vertically between two clean glass plates, avoiding any air bubbles. After completion of the electrophoresis, DNA was visualised by the staining the gel in an ethidium bromide solution under UV light.

2.2.3.4 SDS-PAGE for the separation of proteins (Laemmli, 1970)

SDS-PAGE gel electrophoresis is a method for separating proteins within a sample for molecular weight determination. The proteins are denatured and rendered monomeric by boiling in the presence of reducing agents (β -mercaptoethanol or dithiothreitol) and negatively charged detergent (SDS). The proteins, which normally differ according to their charges, are all coated with the SDS molecules, which are negatively charged. Hence, all the proteins in the sample become negatively charged and achieve constant charge to mass ratio. In this way, the separation is according to the size of the proteins. A SDS-PAGE consists of two gels; firstly, a 10-12% separating gel was poured. In order to achieve a smooth boundary between separating and stacking gel, the separating gel was covered with a layer of water. After polymerisation of the separating gel, a 4% stacking gel was poured over it. The samples were boiled in sample buffer for 10 min at 95°C before loading into the gel. The gel was run at 15 mA for 1 hr then at a constant current of 30 mA.

2.2.4 Isolation of DNA fragments after agarose gel electrophoresis

2.2.4.1 Glass Silica Method

(Vogelstein and Gillespie, 1979)

For the isolation of DNA fragments of 300-4000 base pair (bp) in length from agarose gels, the GeneClean kit from Biomol was used. The principle of this method depends on the binding capacity of DNA to silica in high salt concentrations and elution in low salt solutions. After separation of DNA on an agarose gel, the DNA fragment to be isolated was excised with a razor blade and weighed, then 3 volumes of JETSORB solution was added and melted at 55°C. Depending on the DNA amount, required amount of Glassmilk, which is an aqueous suspension of silica matrix, was added and the tube was placed on ice for 30 min. After centrifuging it at full speed for 2 min, the pellet was washed 2 times with “New Wash”, and allowed to dry at room temperature. For elution of DNA, the pellet was resuspended in 30 µl of H₂O and incubated at room temperature for 10 min with continuous shaking. After the final centrifugation at 14000xg for 5 min, the supernatant containing the DNA was transferred into a new tube.

2.2.4.2 QIAquick Gel Extraction method

This method is designed to extract and purify DNA of 70 bp to 10 kilobase pair (kb) in length from agarose gels. Up to 400 mg agarose can be processed per spin column. The principle of this method depends on selective binding of DNA to uniquely designed silica-gel membrane. To the excised DNA fragment from agarose 3 volumes of QG buffer was added and incubated at 50°C for 10 min. After the gel slice was dissolved completely, it was applied over a QIAquick column and centrifuged for 1 min. The flow through was discarded and the column was washed with 750 µl of PE buffer. After drying the column it was placed into a fresh microcentrifuge tube. To elute DNA, 50 µl of EB buffer was applied to the centre of the QIAquick membrane and centrifuged for 1 min.

2.2.5 Enzymatic modifications of DNA

2.2.5.1 Restriction enzyme digestion of DNA

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

2.2.5.2 Ligation of DNA fragments

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

30 ng	vector DNA (digested)
50-100 ng	insert DNA (1:3, vector: insert ratio)
1 μ l	ligation buffer (10x)
1 μ l	T4 DNA ligase (5 U/ μ l)

in a total volume of 10 μ l

Blunt-end ligations were carried out at 16°C for overnight, whereas overhang-end ligations were carried out at room temperature for 2-4 hrs.

2.2.5.3 TA-Cloning

(Clark, 1988; Hu, 1993)

Taq and other polymerases have a terminal transferase activity that results in the non-template addition of a single nucleotide to the 3' ends of PCR products. In the presence of all 4 dNTPs, dATP is preferentially added. This terminal transferase activity is the basis of the TA-cloning strategy. For cloning of PCR products, the pGEM-T or pGEM-T Easy vector systems that has 5' dT overhangs were used.

The followings were mixed:

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

The content was mixed by pipetting and the reaction was incubated overnight at 4°C. For transformation of the ligation reaction, DH5 α competent cells were used.

2.2.5.4 Filling-up reaction (Costa and Weiner, 1994)

0.1-4 μ g of digested DNA was mixed with 0.05 mM dNTPs and 1-5 U of Klenow fragment with reaction buffer in a total volume of 50 μ l. The reaction was incubated at 37°C for 15 min, and then stopped by heating at 75°C for 10 min.

2.2.6 Preparation of competent *E.coli* bacteria (Dagert and Ehrlich, 1979)

The competent bacterial cells are characterised by a physical cell wall modification that facilitates DNA uptake. LB medium (100 ml) was inoculated with a single colony of *E.coli* (strain DH5 α) and the culture was grown at 37°C to OD 600 = 0.6. Bacteria were centrifuged (10 min, 4°C, 3000xg) and the pellet was resuspended in 50 ml of sterile 50 mM CaCl₂ solution (4°C) and incubated on ice for 30 min. The suspension of bacteria was centrifuged (10 min, 4°C, 3000xg) and the pellet was resuspended in 10 ml of sterile 50 mM CaCl₂ (4°C) with 15% glycerol. The mixture was dispensed into aliquots of 100 μ l and stored at -80°C.

2.2.7 Transformation of competent bacteria (Ausubel et al., 1994)

Transformation of the bacteria was done by gently mixing one aliquot of competent bacteria (100 μ l) with 10 μ l of ligation reaction. After incubation for 30 min on ice,

bacteria were heat shocked for 45 sec at 42°C, cooled down for 2 min on ice. After adding 450 µl of LB medium, bacteria were incubated at 37°C in a shaker with a speed of 160 rpm for 1hr to allow recovery of heat shocked bacteria and were plated out on LB-agar plates containing appropriate antibiotic (50 µg/ml) and whenever required, 1 mM IPTG and X-Gal 40 mg/ml. X-Gal for “Blue-White” selection.

2.2.8 Polymerase Chain Reaction (PCR)

Undoubtly, the polymerase chain reaction (PCR) represents the single most important technique in the field of molecular biology. It is a very sensitive and powerful technique (Saiki et al., 1988) that is widely used for the exponential amplification of specific DNA sequences in vitro by using sequence specific synthetic oligonucleotides (primers). The general principle of PCR starts from a pair of oligonucleotide primers that are designed so that a forward or sense primer directs the synthesis of DNA towards a reverse or antisense primer, and vice versa. During the PCR, the *Taq* DNA polymerase (a heat stable polymerase) (Chien et al., 1976) catalyses the synthesis of a new DNA strand that is complementary to a template DNA from the 5' to 3' direction by a primer extension reaction, resulting in the production of the DNA region flanked by the two primers. It allows the rapid and unlimited amplification of specific nucleic acid sequences that may be present at very low concentrations in very complex mixtures.

2.2.8.1 PCR 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 (10 pmol/µl)
1 µl	reverse primer (10 pmol/µl)
1 µl	10 mM dNTPs
5 µl	10x PCR buffer
1.5 µl	50 mM MgCl ₂
1 µl	<i>Taq</i> DNA Polymerase (5U/µl)
Up to 50 µl	H ₂ O

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

Initial denaturation	95 ⁰ C	5 min
Elongation 30-35 cycles	95 ⁰ C	30 sec (denaturation)
	58 ⁰ C	45 sec (annealing)
	72 ⁰ C	1-2 min (extension)
Final extension	72 ⁰ C	10 min

2.2.8.2 Genotyping of the knock-out mice by using PCR

The genotypes of all offspring of the mutant mice were analysed by polymerase chain reaction (PCR). For amplification of the wild-type and the mutant allele, the DNA was extracted from mouse tails as described in section 2.2.1.2 and pipetted to the following reaction mixture:

1 µl	DNA	(300-500 ng)
1 µl	J3	(10 pmol/µl)
1 µl	J16	(10 pmol/µl)
1 µl	NeoRI	(10 pmol/µl)
1 ml	dNTPs	(10 mM)
5 µl	<i>Taq</i> Polymerase buffer (10x)	
0.5 µl	<i>Taq</i> Polymerase (5 U/µl, Gibco)	
Up to 50 µl H ₂ O		

The mixture was subjected to the following program in the thermocycler:

Initial denaturation	94 ⁰ C	4 min
----------------------	-------------------	-------

Elongation

(for 35 cycle)	94°C	30 sec	Denaturation
	59°C	45 sec	Annealing
	72°C	1 min 30 sec	Elongation
Final extension	72°C	10 min	

2.2.8.3 PCR amplification of microsatellite markers

For amplification of microsatellite markers, fluorescent-coupled primers were used. The 5' end of each forward primer was labelled with any one of the four different fluorochromes.

6-FAM
 HEX
 TET
 NED

PCR reaction was carried out in 15 µl of volume. The PCR was performed in a special touch down PCR program in 96 well PCR plates.

Initial denaturation	94 ⁰ C	4 min
Touch down 1 3 cycle	94 ⁰ C	30 sec
	61 ⁰ C	30 sec
	68 ⁰ C	45 sec
Touch down 2 3 cycle	94 ⁰ C	30 sec
	59 ⁰ C	30 sec
	68 ⁰ C	45 sec
Touch down 3 3 cycle	94 ⁰ C	30 sec
	57 ⁰ C	30 sec
	68 ⁰ C	45 sec
Enlogation	94 ⁰ C	30 sec
30 cycles	55 ⁰ C	30 sec

	68 ⁰ C	45 sec
Final extension	68 ⁰ C	20 min

2.2.8.4 Reverse transcription PCR (RT-PCR)

RT-PCR generates cDNA fragments from RNA templates and is very useful to determine the expression of genes in specific tissues or in different development stages. 1-5 µg of total RNA was mixed with 1 µl of oligo (dT)₁₈ primer (10 pmol/µl) in a total volume of 12 µl. To avoid the possible secondary structure of the RNA, which might interfere with the synthesis, the mixture was heated to 70⁰C for 10 min, and then quickly chilled on ice. After a brief centrifugation, the following reagents were added to the mixture.

4 µl	5x First strand buffer
2 µl	0.1 M DTT
1 µl	10 mM dNTPs
1 µl	Rnasin (10 U/µl)

The content of the tube was mixed gently and incubated at 42⁰C for 2 min. Then, 1 µl of reverse transcriptase enzyme (Superscript II) was added, and further incubated at 42⁰C for 50 min for the first strand cDNA synthesis. Next, the reaction was inactivated by heating at 70⁰C for 15 min. One µl of the first strand reaction was used for the PCR reaction (as described above).

2.2.9 Protein and biochemical methods

2.2.9.1 Isolation of total proteins

100 mg of tissue was homogenised in 500 µl of 0.25 M Tris/HCl, pH 7.8, with a Teflon-glass headed pestle. The cell membrane was destroyed by freezing in liquid nitrogen and thawing at 37⁰C, which was repeated three times. The samples were centrifuged at 8000xg for 10 min at 4⁰C and supernatant was distributed in several microcentrifuge tubes. The tubes were frozen in liquid nitrogen, and stored at -80⁰C.

2.2.9.2 Isolation of nuclear proteins

(Deryckere and Gannon, 1994)

Solution A	0.6%	Nonidet P-40
	150.0 mM	NaCl
	10.0 mM	HEPES pH 7.9
	1.0 mM	EDTA
	0.5 mM	PMSF
	0.5 mM	DTT
Solution B	25%	Glycerol
	20.0 mM	HEPES pH 7.9
	420.0 mM	NaCl
	1.2 mM	MgCl ₂
	0.2 mM	EDTA
	0.5 mM	DTT
	0.5 mM	PMSF
	2.0 mM	Benzamidine
	5 µg/µl	Aprotinin
	5 µg/µl	Leupetin
	5 µg/µl	Pepstatin

Between 100 and 500 mg of tissue was homogenised in 15-25 ml of solution A with the use of a 50 ml Dounce tissue homogeniser. It was then centrifuged for 1 min at 800xg at 4⁰C to remove any unhomogenised tissue. The supernatant was centrifuged again for 8 min at 3200xg. The pellet, which contained nuclei, was dissolved in 50-500 µl of solution B, and incubated on ice for 20 min in order to lyse the nuclei. After a final centrifugation at 14000xg for 2 min, the nuclear proteins were aliquoted into tubes and stored at -80⁰C.

2.2.9.3 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 Coomassie Blue G-250 shifts from 494 to 595 nm when the binding to 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 H₂O, 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).

2.2.9.4 Coupling of the synthetic peptide to BSA

20 mg of BSA (~ 0.3 µmol) were dissolved in 0.5 ml of 0.4 M PBS, pH 7.5. 10 µmol of synthetic peptide was dissolved separately in 1.5 ml of water and the pH was adjusted to 7.5. The solutions of BSA and peptide were mixed and added drop by drop during a course of 5 min time into a solution of 10 µm glutaraldehyde under continuous stirring. This composite mixture was stirred further for 30 min. The unincorporated glutaraldehyde was inactivated by adding 0.1 volume of 1 M glycine solution and the sample was stirred for 30 min and finally dialysed against PBS for overnight.

2.2.10 Blotting techniques

2.2.10.1 Southern blotting of DNA onto nitrocellulose filters

(Southern, 1975)

In Southern blotting, the transfer of denatured DNA from agarose gels to nitrocellulose membrane is achieved by capillary flow. 20x SSC buffer, in which nucleic acids are highly soluble, is drawn up through the gel into the nitrocellulose membrane, taking with it the single-stranded DNA that becomes immobilised in the membrane matrix.

After electrophoresis of DNA, the gel was treated with 0.25 M HCl for depurination. It was followed by denaturation solution for 30 min, and 45 min in neutralization solution. The

transfer of the DNA onto the nitrocellulose membrane was done in a Turbo-Blot-apparatus (Schleicher & Schuell, Dassel). About 20 Whatman filter papers (GB 003) were layered on a stack tray followed by 4 Whatman filter papers (GB 002) and 1 Whatman filter paper GB 002 soaked with 2x SSC. The equilibrated nitrocellulose filter that was also soaked with 2x SSC was laid on the top. The agarose gel, which was treated as described above was placed on the filter, and was covered with 3 Whatman filter papers GB 002 soaked with 2x SSC. The buffer tray was placed and filled with 20x SSC. Finally a wick, which was soaked with 20x SSC and the wick cover were put on the top of the blot. The transfer was carried out for overnight. Finally, after disassembling the blot, the filter was washed briefly in 2x SSC and the DNA was fixed onto the filter by either baking it at 80°C for 2 hrs or by UV-crosslinking in UV Stratalinker 1800.

2.2.10.2 Northern blotting of RNA onto nitrocellulose filters

For the transfer of RNA onto a nitrocellulose filter, the same procedure as described above (2.2.10.1) was performed. In this case, however, the gel does not need to be denatured, but was transferred directly onto the filter.

2.2.10.3 Western blotting of protein onto PVDF membrane (Gershoni and Palade, 1982)

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

After the electrophoresis of proteins on a SDS-PAGE, the gel and the PVDF membrane, which was cut at the size of the gel, were first moistened with methanol and then equilibrated in anode II buffer. Six pieces of GB004 Whatman filter paper were also cut at the size of the gel. Two pieces of filter papers were soaked in anode buffer I and one paper in anode II buffer, first, the papers soaked with anode I buffer were placed on semi dry transfer machine's lower plate and than papers soaked with anode II buffer was placed over it. The equilibrated

membrane was placed over them and then the gel was placed avoiding any air bubbles. Another three Whatman paper soaked with cathode buffer was placed over to complete the sandwich model. The upper plate was placed over this sandwich and the transfer was carried out at 3.5 mA/cm^2 for 1 hr.

2.2.11 “Random Prime” method for generation of ^{32}P labelled DNA (Denhardt, 1966; Feinberg and Vogelstein, 1989)

RediprimeTM II Random Prime Labeling System (Amersham Pharmacia) was used for labelling of DNA probes. The method depends on the random priming principle developed by Feinberg and Vogelstein (1989). The reaction mix contained dATP, dGTP, dTTP, Klenow fragment (4-8 U) and random oligodeoxyribonucleotides. Firstly, 25-50 ng of DNA were denatured in a total volume of 46 μl at boiling water for 10 min and quick chilled in ice for 5 min. After pipetting the denatured probe to RediprimeTM II Random Prime Labeling System cup, 4 μl of [α - ^{32}P] dCTP (3000 Ci/mmol) was added to the reaction mixture. The labelling reaction was carried out at 37°C for 1 hr. The labelled probe was purified from unincorporated [α - ^{32}P] dCTP by using microspin columns (Amersham Pharmacia).

2.2.12 Non-radioactive dye terminator cycle sequencing

Non-radioactive sequencing was performed with the Dye Terminator Cycle Sequencing-Kit (ABI PRISM). The reaction products were analysed with automatic sequencing equipment, namely 377 DNA Sequencer (ABI PRISM). For the sequencing reaction, four different dye labelled dideoxy nucleotides were used (Sanger et al., 1977), which, when exposed to an argon laser, emit fluorescent light which can be detected and interpreted. The reaction was carried out in a total volume of 10 μl containing 1 μg plasmid DNA or 100-200 ng purified PCR products, 10 pmol primer and 4 μl reaction mix (contains dNTPs, dideoxy dye terminators and *Taq* DNA polymerase). Elongation and chain termination take place during the following program in a thermocycler: 4 min denaturation followed by 25 cycles at 95°C , 30 sec, denaturation; 55°C , 15 sec, annealing; 60°C , 4 min, elongation. After the sequencing reaction, the DNA was precipitated with 1/10 volume 3 M sodiumacetate and 2.5 volumes of 100% ethanol and washed in 70% ethanol. The pellet was dissolved in 4 μl of loading buffer, denatured at 95°C for 3 min, and finally loaded into the sequence gel.

2.2.13 Hybridisation of nucleic acids

(Denhardt, 1966)

The membrane to be hybridised was equilibrated in 2x SSC and transferred to a hybridisation bottle. After adding 10 ml of hybridisation solution and sheared salmon DNA, the membrane was incubated for 2 hrs in the hybridisation oven at an appropriate temperature, which was usually 65°C. Then, the labelled probe was denatured at 95°C for 10 min, quick chilled, and added to the hybridisation solution. The hybridisation was carried out overnight in the oven. Next day, the filter was washed for 10 min with 2x SSC at room temperature. Finally it was washed with 0.2x SSC containing 0.1% SDS at the hybridisation temperature. After drying the filter, it was sealed in Saran wrap, and exposed to autoradiography overnight at -80°C or to Phosphoimager screen for 1-4 hrs. The film was developed in X-ray Automatic Processor Curix 60 or screen was scanned in Phosphoimager.

2.2.14 Generation of polyclonal antibody against peptide

2.2.14.1 Peptide analysis

Different computational tools were applied to select potential antigenic peptides. Before synthesis of the peptides, a hydrophilicity/hydrophobicity profile analysis was carried out and for further confirmation of antigenic peptides, antigenicity prediction was performed. In next step, predictions of secondary structure such as β -turns and α -helices in combination with the surface probability of the protein region were the parameters, which were analysed; this enabled us to select the best antigenic peptides. In the last step, we compared primary sequence of our protein with Genebank database (<http://www.ncbi.nlm.nih.gov/>) to select unique sequence for antibody generation. Two peptides were selected and synthesised. The sequences of peptides are as follows:

Peptide 1 : N₂H- WLQ GSK ATT GRT SEE C - CONH₂ (16 aa)

Peptide 2: N₂H- CQR MEE LSR PKR IYQ E -CONH₂ (16 aa)

2.2.14.2 Immunisation of rabbit

Eurogentec Company did immunisation under DOUBLE X program. Two peptides were selected and synthesised instead of one. Using modern algorithms for peptide selection, the success rate for peptide immunisation can be as high as 75%. This still means a 25% chance of failure. Under DOUBLE X program the success rate is increased to 93.75%.

5 mg of each peptide were conjugated with carrier protein molecules (BSA) as described in section 2.2.9.4 and mixed together before immunisation. Two rabbits were immunised with 100 µg of antigen mixed with Freund's complete adjuvant in 1:1 ratio. Before injection pre-immune serum were collected from the animals. After 14 days a second booster immunisation was performed with 1:1 ratio of antigen with Freund's incomplete adjuvant. A third booster was given after 28 days and final bleeding was done after 36 days. The antiserum was aliquoted and stored at -80°C.

2.2.14.3 Determination of specificity of the polyclonal antibody

1x PBS II

004 mM	KH ₂ PO ₄
016 mM	Na ₂ HPO ₄
115 mM	NaCl

Blocking solution II

5.0%	Skimmed milk powder in washing stock 1x
0.5%	Tween 20 in 1x PBS II

PBT II

0.1%	Tween 20 in 1x PBS II
------	-----------------------

10x washing stock:

1.4 M	NaCl
5.0 mM	MgCl ₂
100.0 mM	Tris/HCl, pH 7.5
0.5%	Tween 20

Washing Buffer:

2%

Skimmed milk powder
in 1x Washing buffer

After SDS-PAGE and electrotransfer of the total mouse proteins to a PVDF membrane, the membrane was blocked with 20 ml of blocking solution II for 1 hr at RT. Then the filter was cut and each lane was incubated with different dilutions of antiserum (1:25, 1:100, 1:500, 1:1000, and 1: 10000) in washing buffer for overnight at 4⁰C. Afterwards, the unbound antiserum was removed by washing the membrane 3 times for 20 min with washing buffer. The secondary antibody coupled with alkaline phosphatase was diluted 1:10000 in washing buffer and added to the blot and incubated for 1 hr. Again the unbound antibodies were removed by washing 4 times for 15 min with washing buffer. The chromogenic reaction was performed with 10 µl NBT and 60 µl of BCIP solution in buffer D until chromogenic precipitate developed. The reaction was stopped by washing the membrane several times with water. The membrane was air-dried and stored in the dark.

2.2.14.4 Affinity purification of the polyclonal antibody

For antibody purification HiTrap NHS-activated 1 ml columns were used, the columns are made of medical grade polypropylene, which is biocompatible and non-interactive with biomolecules. NHS-Activated SepharoseTM is designed for the covalent coupling of ligands containing primary amino group. Non-specific adsorption of proteins to HiTrap columns is negligible due to the hydrophilic properties of the base matrix. The activated gel is supplied in 100% isopropanol to preserve the stability of the activated gel prior to coupling.

2.2.14.4.1 Ligand coupling

Coupling buffer:

0.2 M NaHCO₃

0.5 M NaCl (pH 8.3)

The peptide (10 mg) was dissolved in the coupling buffer. Isopropanol present in column was washed with 1M HCl, ice cold. The flow rate during the pumping was adjusted to about 1 ml/min. Immediately 1 ml of the ligand solution was injected onto the column, then column was sealed and let it stand for 4 hrs at 4⁰C.

2.2.14.4.2 Washing and deactivation

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

A series of alternate washing (injection 3x2 ml) with buffer A and buffer B was done. Finally, 2 ml of PBS was injected.

2.2.14.4.3 Purification

The column was equilibrated with 10 column volumes of start buffer (PBS). The antiserum was diluted five times with PBS and filtered through a 0.45 µm filter and then applied onto the column. During pumping, a constant flow rate of 0.5 ml/min was maintained. The column was washed with 10 column volumes of PBS. Elution was done with three volumes of 3 M MgCl₂. The purified fraction was desalted by dialysing against PBS. The column was re-equilibrated with 10 volumes of PBS.

2.2.15 Histological techniques

2.2.15.1 Tissue preparation for paraffin-embedding

Bouin's Solution	15 ml Picric acid
	05 ml 37% Formaldehyde
	01 ml Acetic acid

The freshly prepared tissues were fixed in Bouin's solution for 24 hrs to prevent the alterations in the cellular structure. The tissue to be embedded in paraffin should be free of water. The dehydration process was accomplished by passing the tissue through a series of increasing alcohol concentrations. For this purpose, the tissue was let in 30%, 70%, 90%, and 100% ethanol for 1hr at RT. Later, the alcohol was removed from the tissue by incubating it in methylbenzoate overnight. It was then incubated in 5 ml of histoclear (Xylol) for 10-30 min at RT. The second histoclear was not discarded but 5 ml of paraplast was added and the incubation was continued at 60°C for another 30 min. The histoclear and paraffin mixture was discarded, and the tissue was further incubated in 5 ml of paraplast at 60°C overnight. Before embedding, the paraffin was changed at least three times. Finally, the tissue was placed in

embedding mold and melted paraffin was poured into the mold to form a block. The block was allowed to cool and was then ready for sectioning.

2.2.15.2 Sections of the paraffin block

The paraffin blocks were pre-cut to the optimal size (1-2 cm³) and clamped into the microtome. The sections were cut in thickness of about 7-10 µm. The sections were floated on 40°C water to allow actual spread and subsequently put onto pre-treated slides. In order to achieve a better adhesion of the sections, the slides were treated with a drop of serum-formalin. A fine brush was used to transfer the sections to the pre-treated slides. After complete evaporation at 37°C for 2-5 days, slides were stored at 4°C for further analysis.

Serum-Formalin:	15 ml	Fresh serum
	15 ml	Glycerin
	06 ml	Formalin 6%

2.2.15.3 Staining of the histological sections (Nissl staining)

The stored slides with the paraffin sections were stained by the following method:

1. Slides were incubated three times in histoclear (Xylol) for 3 min.
2. Incubation in 100%, 96%, 80%, 70% and 50% ethanol for 2 min in each.
3. 1 min in H₂O and then 7 min in Crystal Violet.
4. Running tap water (control) for 10 min, then 1 min in dH₂O.
5. Eosin 0.1% + 2 drops acetic acid for 5 min, then in dH₂O for 1 min.
6. Incubation in 50%, 70%, 80%, 90%, 96% and 100% ethanol for 2 min in each.
7. Slides were incubated two times in histoclear (Xylol) for 3 min.

2.2.16 Immunofluorescence staining of mouse testis

Fixation and subsequent treatment of mouse testicular tissue was performed as described in section 2.2.15.1. Adult mouse testis cross sections (7-10 μm) were dewaxed with roticlear solution (Roth) and rehydrated by descending ethanol concentrations. For immunofluorescence staining, sections were washed in PBS and were then incubated with a blocking solution containing 5% sheep serum and 0.02% Tween-20 in PBS (PBT) for 1 hr at RT. The testis sections were incubated with affinity-purified primary polyclonal antibody for 16 hrs at 4⁰C. The tissue sections were rinsed four times in PBT and subsequently incubated with indocarbocyanine (Cy3)-conjugated sheep anti rabbit Ig (1:100; Sigma) for 1 hr at RT. After the tissue sections were washed three times with PBS, the nuclei were counterstained with DAPI (Vector). Immunostaining of the sections was examined using a fluorescence-equipped microscope (BX60; Olympus).

For immunocytochemical staining, cellular suspensions were prepared from mouse testes by using the collagenase/trypsin method according to published procedure (Romrell et al., 1976). Testes from 60 days old mice were collected aseptically in serum-free culture medium, rinsed in 0.1 M PBS, pH 7.2. After removal of the tunica albuginea, seminiferous tubules were cut into pieces and interstitial cells were removed by mechanical agitation. Subsequently, seminiferous tubules were washed with medium treated with collagenase (Roche, 0.2 mg/ml) in PBS for 15 min at 37⁰C. The medium was replaced by PBS containing Dnase (Sigma, 10 $\mu\text{g}/\text{ml}$) and the mixture was pipetted several times until a cell suspension was diluted (10^7 cells/ml), spread onto pre-treated glass slides, air-dried and mixed in 100 % methanol for 10 min.

The immunocytochemical staining procedure was performed as described above.

2.2.17 Fluorescence in situ hybridisation (FISH)

Chromosomes were prepared from human peripheral blood lymphocytes. A PAC genomic clone containing the human THEG gene was labeled by standard nick translation with biotin-16-dNTP (Boehringer Mannheim) and used as a probe for hybridisation. Hybridised probes were detected by fluorescein isothiocyanate (FITC)-conjugated avidin (Vector) as described earlier (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.

2.2.18 Transfection of NIH3T3 cells with the *Gfp-fusion* construct

Approximately 4×10^5 fibroblast cells (NIH3T3) were plated in a cell chamber with slide (Lab-Tek) and cultured overnight in 5 ml DMEM medium containing 10% FCS and penicillin/streptomycin at 37° C and 5% CO₂. 1 µg of DNA (*Gfp-fusion*) was diluted with the DNA-condensation buffer (buffer EC of the Effectin 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% CO₂ 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.

2.2.19 Techniques for production of targeted mutant mice (Joyner, 2000)

The discovery that cloned DNA introduced into cultured mouse embryonic stem cells can undergo homologous recombination at specific locus has revolutionised our ability to study gene function in vitro and in vivo. In theory, this technique will allow us to generate any type of mutation in any cloned gene. Over twenty years ago, pluripotent mouse embryonic stem cells (ES) derived from inner cell mass cells of mouse blastocysts were isolated and cultured (Martin, 1981; Evans and Kaufman, 1981). Using stringent culture conditions, these cells can maintain their pluripotent developmental potential even after many passages and following genetic manipulations. Genetic alterations introduced into ES cells in this way can be transmitted into the germline by producing mouse chimeras. Therefore, applying gene targeting technology to ES cells in culture gives the opportunity to alter and modify endogenous genes and study their functions in vivo.

2.2.19.1 Production of targeted embryonic stem cell clones

2.2.19.1.1 Preparation of EMFI feeder layers

A frozen vial of EMFI cell was quickly thawed at 37°C and transferred to 10 ml EMFI medium. After centrifugation at 270xg for 5 min, the cell pellet was gently resuspended in 10 ml of EMFI medium and plated on a 50 mm³ culture flask. Cells were incubated at 37°C, 5% CO₂. When the cells formed a confluent monolayer (three days), they were either trypsinised, transferred to five 150 mm³ dishes and grown until they formed confluent monolayer, or directly treated with mitomycin C. To treat the EMFI with mitomycin C, the medium was removed and 10 ml fresh medium containing 100 µl mitomycin C (1 mg/ml) was added. After 2-3 hrs of incubation, the monolayer of cells was washed twice with 10 ml PBS. The cells were then resuspended with 10 ml medium and gentle pipetting dissolved any cell aggregates. The cells were centrifuged, resuspended in EMFI medium and plated onto dishes, which were treated with 0.1% gelatine for 30 min. The feeder cells were allowed to attach the surface of the dishes by incubating them for overnight at 37°C, 5% CO₂. Before adding ES cells on the feeder layer, the medium of the dishes was changed to ES cell medium.

2.2.19.1.2 Growth of ES cells on feeder layer

One vial of frozen ES cells was quickly thawed and cells were transferred to a 12 ml tube containing 6 ml ES cell medium. After centrifugation, the cell pellet was resuspended in 5 ml ES cell medium and plated on 60 mm dishes containing EMFIs at 37°C, 5% CO₂. Next day the medium was changed. The second day, cells were washed with PBS, treated with 2 ml trypsin/EDTA at 37°C, 5% CO₂ for 5 min. The cells were gently pipetted up and down to dissolve cell clumps, resuspended with 5 ml ES medium and centrifuged. The cell pellet was resuspended in 10 ml ES cell medium and distributed either to 5 or 6 dishes (60 mm³), containing feeder layers or to 2 dishes (100 mm³) containing feeder layers. The cells were passaged every second day.

2.2.19.1.3 Electroporation of ES cells

ES cells, which have grown for two days on 100 mm³ dishes, were trypsinised. The cell pellet was resuspended in 20 ml PBS and centrifuged. The cell pellet was then resuspended in 1 ml

PBS. 0.8 ml of cell suspension was mixed with 40 µg of linearised DNA-construct and transferred into an electroporation cuvette. The electroporation was performed at 240 V, 500 µF with the BIO RAD gene pulser™. After electroporation, the cuvette was placed on ice for 20 min. The cell suspension was transferred from cuvette into 20 ml of ES cell medium and plated onto two 100 mm³ dishes containing feeder layers. The medium was changed every next day. Two days after the electroporation, the drugs for the selection were added (active G418 at 150-250 µg/ml and gancyclovir at 2 µM). The medium was changed every day. After about eight days of selection, drug resistant colonies have appeared and were ready for screening by Southern blot analysis.

2.2.19.1.4 Growing ES cells for Southern blot analysis

The drug resistant colonies that were formed after about eight days of selection were picked with a drawn-out Pasteur pipette under a dissecting microscope. Each colony was transferred into a 24 well plate containing feeders and ES cell medium. After 2 days, the ES cells were trypsinised with 100 µl of trypsin for 5 min and resuspended in 500 µl ES cell medium. Half of the cell suspension in each well was transferred to a well on two different 24 well plates, one gelatinised plate, and the other containing feeder cells (master plate). The gelatinised plate was used for preparing DNA and the master plate was kept frozen.

2.2.19.2 Production of chimeras by injection of ES cells into blastocyst

The ability of mammalian embryos to incorporate foreign cells and develop as chimeras has been exploited for a variety of purposes including the perpetuation of mutations produced in embryonic stem (ES) cells by gene targeting, and the subsequent analysis of these mutations. The standard procedure is to inject 10-20 ES cells, which, are recombinant for targeted locus into the blastocoel cavity of recently cavitated blastocysts that have been recovered by flushing the uteri of day 4 pregnant mice (C57BL/6J). After injection, embryos are cultured for a short period (2-3 hrs) to allow re-expansion of the blastocoel cavity, which collapses upon injection and then transferred to the uterine horns of CD1 pseudopregnant mice. Pseudopregnant females are obtained by mating 6-8 weeks old oestrous females with vasectomised males.

2.2.19.3 Detection of chimerism and mice breeding

The most convenient and readily apparent genetic marker of chimerism is coat colour. Chimeric males (and sometimes females) are generally test bred to ascertain contribution of the ES cells to germline. Once a germline chimera has been identified, the first priority will be to obtain and maintain the targeted allele in living animals (inbred background). The chimeras were bred with C57BL/6J and with 129X1/SvJ background mice to compare the phenotype in two different genetic backgrounds.

2.2.20 Sperm isolation and counting

Epididymes of mice were dissected under aseptic condition and put in 1 ml of IVF medium. Spermatozoa were allowed to swim out of the epididymes for 1 hr at 37⁰C, 5 % CO₂. Sperm suspension was diluted 40 times with PBS before counting. 5 µl of this suspension was put into counting chamber and sperms were counted in 10 independent fields (each having an area of 0.0025 mm²) under the microscope (Olympus BX60) with 20X magnification. Total sperms were calculated by following formula:

$$\text{Total Sperm} = \text{average of sperms} \times 10 \times 40 \times 10^3$$

2.2.21 Sperm motility

Sperms were isolated as described in (2.2.20). 10 µl of sperm suspension was put on a dual sided sperm analysis chamber. Sperm motility was quantified using the computer assisted semen analysis (CASA) system (CEROS version 10, Hamilton Thorne Research). At least 250 motile sperms from each animal were analysed. Mean, standard deviation, median and range were calculated using standard statistical program.

2.2.22 Acrosome reaction

Sperms were isolated as described (2.2.20). They were capacitated by incubating for 1 hr at 37⁰C, 5% CO₂. The sperms were transferred into two microcentrifuge tubes and centrifuged for 2 min at 3000xg. The supernatant was aspirated leaving only 50 µl for

resuspension of sperms. 2.5 µl of Ionophore A23187 (end concentration 10 µM in DMSO) was added to sperm suspension, for negative control 2.5 µl of phosphoric acid (5 mM) was added and incubated at 37°C for 1 hr. The sperms were then fixed in 500 µl of 2% formaldehyde (in PBS) for 30 min at 4°C. After completion of fixation, it was centrifuged at 4000xg for 2 min. Sperms were further washed twice with 0.15 mM Ammonium acetate. Finally they were resuspended in 100 µl of PBS and 30 µl of suspension was spread on superfrost slide and air-dried. The slides were stained with coomassie G-250 in 3.5% H₂O₂ for 2.5 min, unbound dye was removed by washing several times with water. The slides were mounted with 30% glycerol and observed under microscope. Sperms with and without blue head were counted at least for 100 sperms. Here blue head sperms means those sperms which failed to undergo acrosome reaction, so the acrosome is present over the sperm head, which was stained with coomassie G-250 therefore appears blue head. The acrosome reaction was calculated as follows:

$$\text{Acrosome reaction (in percentage)} = \frac{\text{Number of sperms without blue head}}{\text{Total numbers of sperms}} \times 100$$

2.2.23 Techniques involved in linkage analysis (Silver, 1995)

To map a mutationally defined locus, one will have to generate a special panel of criteria by which segregation of the mutant and wild type alleles can be followed phenotypically in animals prior to DNA preparation for marker locus typing.

2.2.23.1 Breeding Scheme

For commencing our linkage study, a two-step four generation cross, the outcross-intercross (F1 x F1) was undertaken. The first cross was an outcross between two parental strain, 129X1/SvJ and C57BL/6J. The original mutation arose in 129X1/SvJ parental strain. In a similar fashion, the animals were bred to generate F2, F3 and F4 affected mice (figure 2.1).

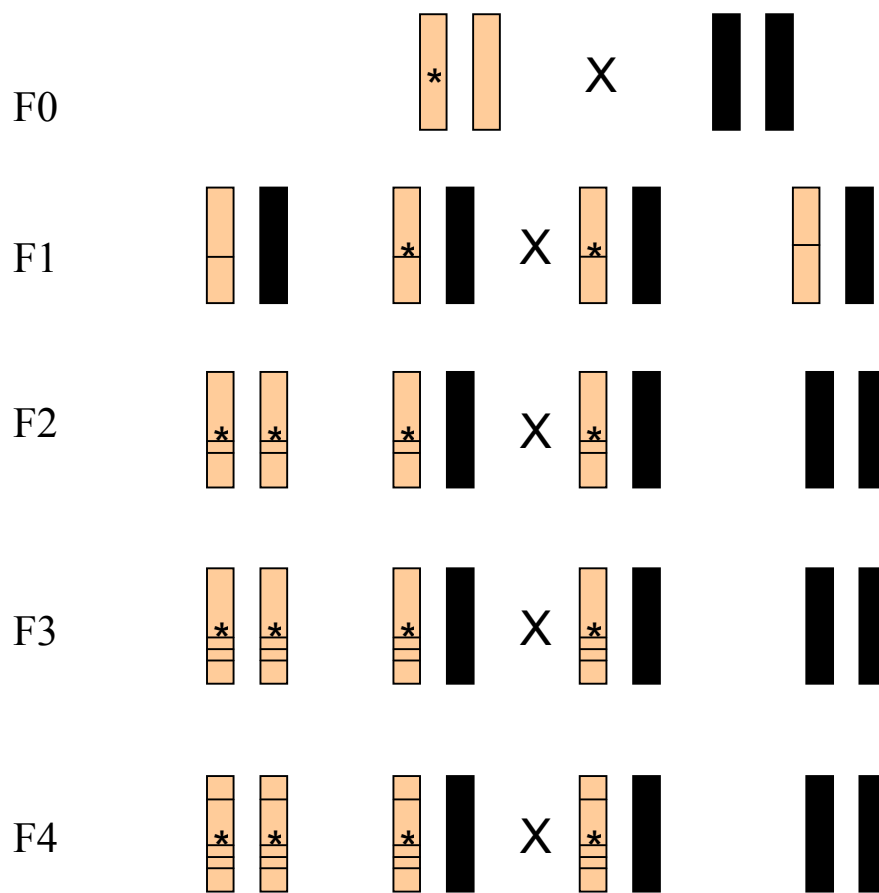


Figure 2.1: Schematic representation of breeding strategy undertaken for linkage analysis. Here brown block represent 129X1/SvJ genotype and black block depicts C57BL/6J genotype, asterisk means a mutation and horizontal line over block was used for meiotic recombination breakpoint event.

2.2.23.2 Framework linkage map: evaluation of swept radius

The first step of this two-stage protocol is to develop a framework map that is “anchored” by well-mapped loci (microsatellite markers) spaced uniformly throughout the entire genome. To accomplish this task most efficiently it is critical to calculate the minimum number of anchor loci required to develop a low-resolution but comprehensive map. This calculation is based on the length of the swept radius that extends on either side of each marker. The swept radius is a measure of the distance over which linkage can be detected between any marker and a test locus. The swept radius can be used in conjunction with the lengths of each individual chromosome to determine the number of anchor loci required to provide complete coverage over the entire genome. Essentially, anchors can be chosen such that their "swept diameters" (twice the swept radius) cover directly adjacent regions that span the length of every

chromosome as illustrated in figure 2.2. With a set number of samples, one can use figure 2.2 to find the corresponding framework swept radius. For example we did our analysis based upon 48 samples, the framework swept radius for such analysis is 15 cM.

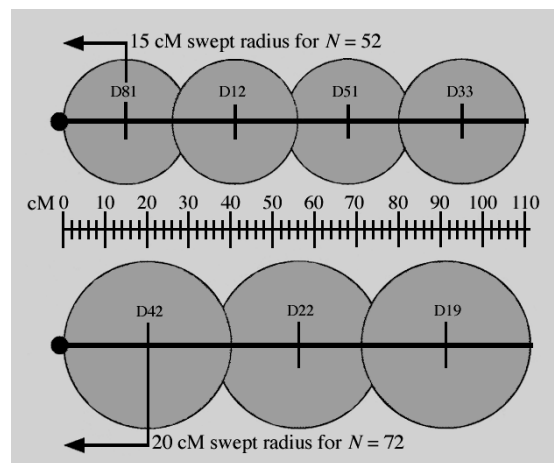


Figure 2.2: Example of the coverage of mouse chromosome 2 with marker loci. The radius of each circle is equivalent to the swept radius around each locus as determined with 52 backcross animals, four marker loci would be required to span the chromosome. With 72, a completely different set of three marker loci would be required

2.2.23.3 A stratified approach to high-resolution mapping

An optimal strategy for high-resolution linkage mapping is to divide the protocol into two separate stages. The goal of the first stage is to link the locus to a defined sub-chromosomal interval. This can be accomplished by typing a relatively small set of markers. Once this first stage is completed, it becomes possible to proceed to the second stage, construction of a high-resolution map just in the vicinity of the locus of interest, with a selected set of markers and a selected number of animal samples. The ultimate goal of this entire protocol is the identification of a handful of markers and recombinant animals that spans a very small interval of loci containing an interesting gene that can be subjected to positional cloning.

2.2.23.4 Genotyping of microsatellite marker

All mapping experiments performed today are based on PCR amplification of target marker. A genome wide screening was performed with fluorescent microsatellite markers at a resolution of 15 cM. The coordinates of the markers were taken from Jackson laboratory web page (<http://www.informatics.jax.org/>) and from Whitehead Institute of Biomedical Research (<http://www.wi.mit.edu/>).

The genomic DNA was prepared from animal tail biopsies as described in section 2.2.1.2.

2.2.23.4.1 Normalisation of genomic DNA

The genomic DNA was diluted by adding 1:1 ratio of DNA sample and water, 50 µl of genomic DNA from each sample was pipetted in a 96 well plate (which was named as grandmother plate) and 50 µl of water was added, which was then let for overnight shaking at 4⁰C to obtain a homogeneous solution. The concentration of DNA was measured by loading 5 µl of sample in 100 µl of HPLC quality water in an optical plate, which was then measured in TECAN UV reader at 260 nm, 280 nm, and 320 nm. A mother plate with uniform concentration of 50 ng/µl was prepared with a computer assisted TECAN robot. The TECAN robot works in conjunction with the optical reader in order to generate a uniform concentration plate by calculating each sample concentration and from that data determines the volume of sample and water to be pipetted in order to achieve a sample concentration of 50 ng/µl. From mother plate a dilution plate of 1 ng/ µl was prepared. From the dilution plate 5 µl of DNA solution was dispensed in each well of fresh 96 well PCR plates with the help of a 96-needle dispenser. These plates were named as daughter plates and were air dried before performing PCR.

2.2.23.4.2 Pooling of PCR reaction for analysis

A panel of marker was identified and characterised before loading into a 370 or 3100 ABI PRISM genetic analyser. Two criteria were considered for formation of a panel:

- 1) Different fluorescence markers
- 2) Different size length of products

A typical panel comprised of 6-10 different markers.

2.2.23.4.3 Analysis of microsatellite markers

5 µl of each marker's PCR reaction was pooled together under one panel, out of this mixture 5 µl was mixed with formamide together with ROX or Tamara size standard, depending upon the filter set used in Genescan run. The reaction was first denatured for 3 min at 95⁰C before loading into 370 or 3100 ABI PRISM genetic analyser.

Filter Set C: TET, FAM, and HEX; size standard Tamara

Filter Set D: NED, FAM, and HEX; size standard ROX

The microsatellite markers were detected depending upon their size under denaturing condition, when exposed to an argon laser and interpreted for scoring. The allelic frequency for each marker where calculated by running the data in GENOTYPER software (ABI PRISM).

22.2.24 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 (CDS) mouse genome database (www.celera.com).

3 Results

3.1 Introduction to result section

In previous studies, our group identified a mouse gene named *Theg* that is expressed only in haploid male germ cells. It was also demonstrated that maintenance/induction of *Theg* gene expression in haploid germ cells requires their interaction with Sertoli cells. The *Theg* encodes a novel protein and no homologous protein was detected in Genebank database (<http://www.ncbi.nlm.nih.gov/>). The complete open reading frame comprised of eight exons spanning 12 kb of genomic DNA (figure 3.1). The gene was mapped in chromosome 10 region B5-C1. The mouse genomic library lambda FIX II (Stratagene) was screened using *Theg* cDNA fragments (U3 and O1) as hybridisation probes. Four positive clones were isolated and were found to contain complete genomic sequence of *Theg*.

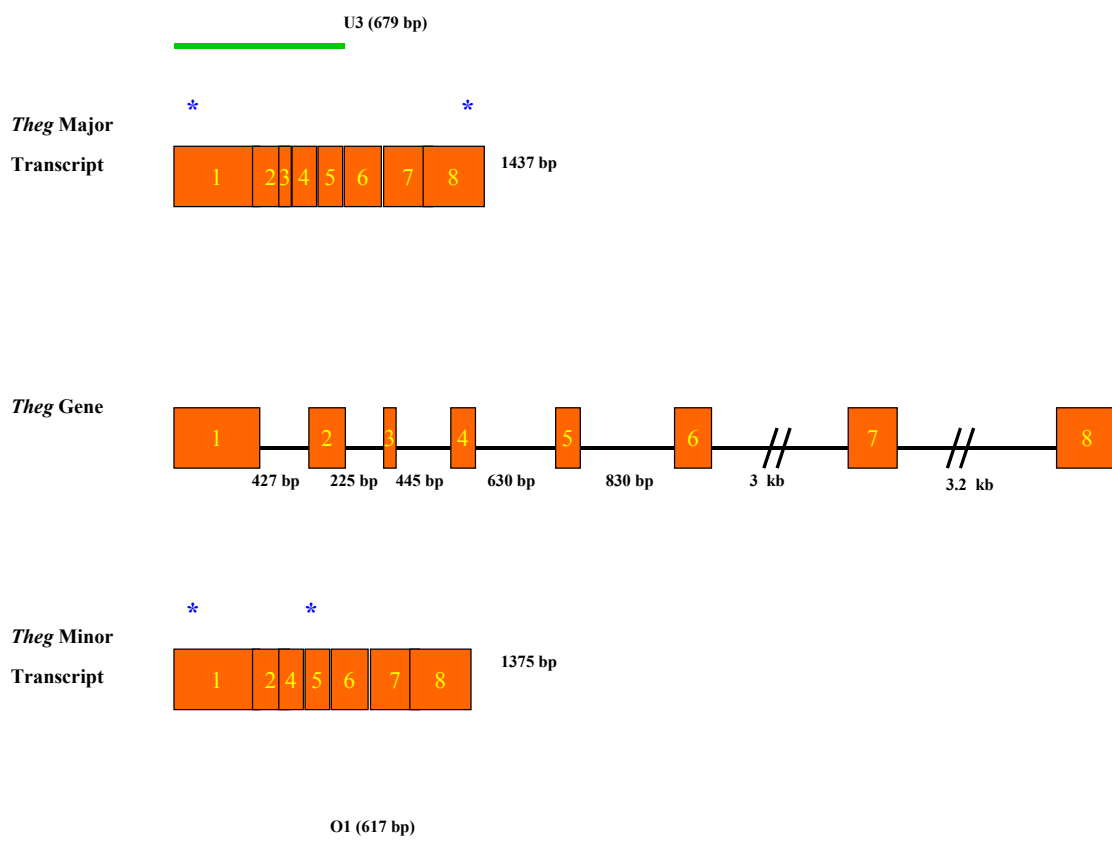


Figure 3.1: Schematic representation of the exon-intron structure of *Theg*. The two alternative spliced products, *Theg* major and *Theg* minor, and two cDNA products (U3 and O1) are shown. Translational start and stop sites are marked as asterisks

3.2 Alternative splicing of *Theg*, cDNA and deduced amino acid sequence analysis

The tissue-specific expression of the mouse *Theg* gene was first determined by Northern Blot analysis with total RNA prepared from various tissues. A single strong transcript of size 1.4 kb was observed only in testicular RNA, but not in RNA from any other tissue examined



Figure 3.2: RT-PCR analysis on total RNA isolated from mouse testis using *Theg*-specific primers mTHEG1 and mTHEG2. Both *Theg* major (352 bp) and *Theg* minor (280 bp) transcripts can be detected by RT-PCR analysis. M, standard molecular weight marker; 1, C57BL/6J testicular RNA; 2, 129X1/SvJ testicular RNA; 3 control RT-PCR for transition protein2 transcript.

(Nayernia et al., 1999). However, no spliced variant was detected in Northern blot. In order to determine if alternatively spliced *Theg* transcripts are expressed in mouse testis, we performed RT-PCR on total mouse testicular RNA using primers mTHEG1 and mTHEG2 located in exon 1-2 and exon 5-6, respectively (figure 3.3). Two RT-PCR products of differing length (352 bp and 280 bp, respectively) were detected in 1.6% agarose gel (figure 3.2). Both RT-PCR fragments were purified from the gel and cloned in pGEMT-easy vector. About 10 clones from each product were sequenced. Sequence analysis of these clones led us to identify two different splice variants, which were named *Theg* major and *Theg* minor. *Theg* major was a full-length transcript, comprising of eight exons, *Theg* minor was identical to *Theg* major

except it differed by 72 bp in length, (from nucleotide position 663 to nucleotide position 735), which constitute the exon 3 of *Theg* transcript and was lacking in *Theg* minor transcript (figure 3.3).

```

gcacagatggcgggtctcggcactcaggag 30
      *          *
gttgaggccagcctggctacatagtgagaccctgtcgcctcttgt 75
ctcccactccaaaagaaagacacgaaggggtggagtcctgagctgg 120
tgcacctgctccaacactgggatggaatggggcctcggaatccat 165
gggtgggcattcttctgttggctgtggcaaccataggtccatttggg 210
tgctgaggagggagtgagggttaaagggtaggaggggcagtacac 255
atggggggaacttgggtgaacaccgggccagtttgctcagtaaccca 300
- G E L G E H R A S L L S N P 15
atccccggaggtcaagactttgggagagttaaagcagggacagaac 345
I P E V K T L G E L K Q G Q N 30
aatggcaatttggaccttgagagcgagccatttgggagccattgg 390
N G N L D L E S E P F G S H W 45
ttgcagggctctaagggcactacagggcgaacctcagaagaacca 435
L Q G S K A T T G R T S E E P 60
Peptide1
gaggaggaaatccccccagaggagatggctgggggaagagctcccg 480
E E E I P P E E M A G E E L P 75
gagacctccaatctggatggctcctcttcaacaagacctggaggtg 525
E T S N L D G P L Q Q D L E V 90
mTHEG1
gaagtggttgaaatgagccatctgtccatcacttgagcggactccc 570
E V V E M S H L S I T E R T P 105
agtgtctccacagccaagggcaggaaaaagaggagccggaggctg 615
S V S T A K G R K K R S R R L 120
ctggagctggcgaagcctaagaccaactggcaatgtctgagagac 660
L E L A K P K T N W Q C L R D 135
aggacggggcgctggttgaagggctatgcoctggatttccccacgc 705
R T G R C C K G Y A W I S P R 150
aagacaacttgcagttctgcoctgtattggccttctgtatactgg 750
K T N L Q F C L Y W P S V Y W 165
accgagcgatattatcgaggacaccacggtgaccatcacggtcct 795
T E R F I E D T T L T I T V P 180
gtgggtgtcccagcgcagtgaggagcttttcgaggcccaagcggttc 840
V V S Q R M E E L S R P K R F 195
Peptide2
taccaggaatattacaacaataacaggacaacgcccatctggtcc 885
Y Q E Y Y N N N R T T P I W S 210
mTHEG2
atccctcgggtccactctggagtaccaagcttctaaccggctgaag 930
I P R S T L E Y Q A S N R L K 225
caactggctaccccgaaggttcgcaataacatctggagcattaac 975
Q L A T P K V R N N I W S I N 240
atgtctgaggtttcccaagtctccagagcagcccagatggcagtc 1020
M S E V S Q V S R A A Q M A V 255
ccgaggccacggacgcttcggctgctggcaaagcccaggccccca 1065
P R P R T L R L L A K P R P P 270
gccactctgttgaagagtgaggaccctatgccaaaacccaagcct 1110
A T L L E E W D P M P K P K P 285
tacgtgtcagactataaccgccttcttcagctggcaacacccaag 1155
Y V S D Y N R L L Q L A T P K 300

```

```

gccctgtcagaaaagtgcggttctctgatcgcagtcctcagtgggag 1200
A L S E K C V P D R S P Q W E 315
gtcctggatgtcaccaagaacgcggtggccagttcacggatcatc 1245
V L D V T K N A V A S S R I I 330
tccctggccccagccccaaaattcgaaaggacctcaacgagggctac 1290
S L A Q P K I R K D L N E G Y 345
aaccatactacatctccccagcctctctggtggctcaggcattct 1335
N P Y Y I S P A S L V A Q A S 360
cctcgaatttatgagcttgccacccccaaatacatcaccaagaaa 1380
P R I Y E L A T P K Y I T K K 375
gtgtgacagccaagccttctctccagctcacccccagtaaacacc 1425
V * 376
aagtgcattcctc 1437

```

Figure 3.3: Nucleotide sequence and deduced amino acid sequence of *Theg* cDNA. Numbering of nucleotide and amino acid sequences is given on the right, transcription start sites are indicated by asterisks, end of an exonic sequence is marked with red underline and polyadenylation signal is marked with blue underline. In deduced amino acid sequence nuclear targeting signal sequences (NLS) and RNA binding signal sequences are marked with bold font. The amino acid sequences of peptide 1 and peptide 2 used for generation of polyclonal antibody are shown in bold blue font. The translation start and stop codon are also marked with asterisks. The exon 3 sequences are marked with green bold font, which is absent in *Theg* minor cDNA. The primer sequences for mTHEG1 and mTHEG2 are marked with blue and the direction of sequence is shown by bold arrow.

3.3 *Theg* protein analysis

3.3.1 Expression pattern of *Theg* protein in adult mouse

Polyclonal antibodies against two *Theg* peptides were generated as described in methods (section 2.2.14). The positions and sequences of the peptides are shown in figure 3.3. In order to determine the specificity and affinity of antiserum, Western blot analysis was performed. Total protein extract from testis and as a control from kidney were separated on SDS-PAGE and transferred onto a PVDF membrane. Western blots were performed with different dilution of antiserum (1:50, 1:100, 1:500, and 1:5000). Several bands were detected in both testicular and kidney protein extracts (figure 3.4). The antiserum derived from rabbits had lots of unspecific antibodies and the affinities of these antibodies against *Theg* protein epitopes were very weak.

In order to obtain high affinity antibody against *Theg* protein, purification of the antiserum was performed. An affinity column (NHS-HiTrapTM Sepharose) conjugated with *Theg* peptides (which were used before for immunisation) was used to purify antibodies against *Theg* peptides. Fractions were eluted after binding of antiserum to the column and dialysed against PBS and then concentrated using centriscan column. Western blot was performed with both testicular and kidney total protein extracts with a dilution of 1:100 of purified antibody. A very strong band was detected only in testicular protein extract but not in kidney protein

extracts. The protein band when compared with pre-stained markers corresponded to a molecular weight of 42 kDa, which was in agreement with the predicted molecular weight of Theg major protein (figure 3.4).

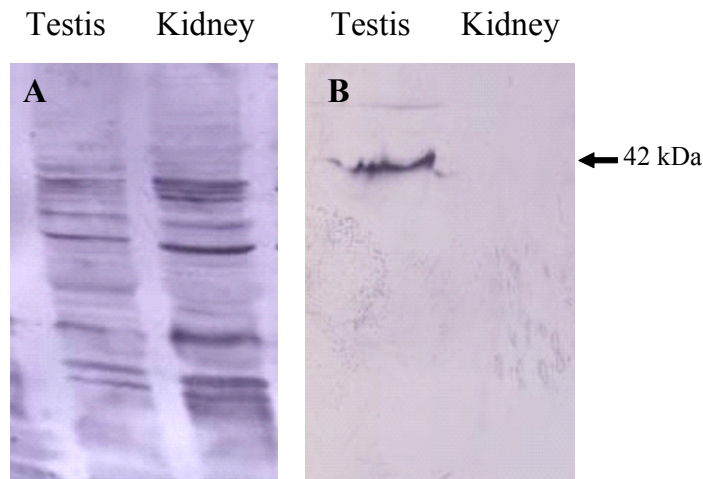


Figure 3.4: Western blot analysis with polyclonal antibody against Theg protein. In this analysis a antibody dilution of 1:100 was used. (A) Unpurified antiserum showed cross-reactivity to lots of testicular and kidney proteins. (B) Affinity purified antibody detected a strong band corresponding to size 42 kDa in testicular protein extract but not in total kidney protein. 20 μ g of protein was loaded in each lane of the gel.

3.3.2 Theg protein expression during different developmental stages of mouse

Total protein extract from testis of mouse from different developmental stages were prepared, namely from stages P18, P20, P23, P27, P30, P35, P40, and from adult mouse. Western blot was performed with purified antibody (1:100 dilutions). Theg expression could be detected from P23 of mouse testicular development. The expression is stronger with ascending developmental stages and the expression is comparable with adult mouse at stage P35. After P35 the expression becomes constant and remains so throughout the adult mouse (figure 3.5). The observation that Theg protein expression initiates around P23 of mouse development also supports Northern analysis result, where expression was first detected at P20 (Nayernia et al., 1999). During mouse testicular development the haploid germ cells appears around P20-P21. Therefore this result confirms that Theg is expressed only in haploid male germ cells.



Figure 3.5: Western blot analysis using affinity purified theg antibody (1:100) on total testicular protein from various developmental stages of mice. 1, P18; 2, P20; 3, P23; 4, P27; 5, P30; 6, P35; 7, P40; 8, adult.

3.3.3 Theg protein expression in different mutant mice

Total testicular protein extracts from different mutant mice (W/W^v , Tfm/y , olt/olt , and qk/qk) were prepared and Western blot was performed. Theg protein could be detected only in olt/olt and qk/qk testicular protein but not in W/W^v or in Tfm/y testicular protein extract (figure 3.6). This result is another added evidence that expression of *Theg* initiates only in haploid male germ cells, as testis of W/W^v lack all stages of germ cells and in Tfm/y mice testis, spermatogenesis is arrested at spermatocyte stage. However in olt/olt mice spermatogenesis proceeds until round spermatids and in qk/qk mice until elongated spermatids.

3.3.4 Theg protein expression in subcellular organelles

To localise Theg protein in subcellular organelles, total testicular protein was fractionated into different subcellular fractions namely nuclear, mitochondrial and cytosolic and separated onto SDS-PAGE and Western blot was performed with affinity-purified antibody. Theg expression was observed in nuclear fraction and also in cytosolic fraction but not in mitochondrial fraction (figure 3.7).

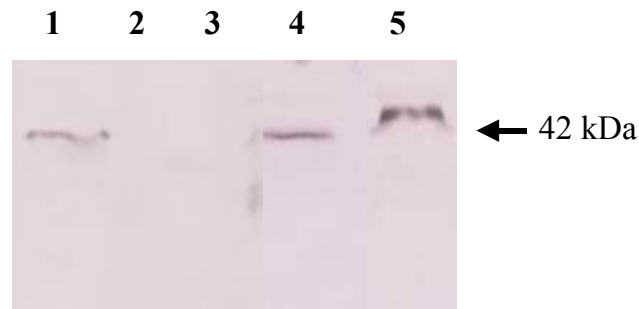


Figure 3.6: Western blot analysis using affinity purified antibody against Theg (1:100) on total testicular protein from different mutants mouse. 1, Wild type; 2, *W/W^v*; 3, *Tfm/y*; 4, *olt/olt*; 5, *qk/qk*. No expression of Theg was observed in *W/W^v* and *Tfm/y* testes. However *olt/olt* and *qk/qk* testes showed a strong expression.

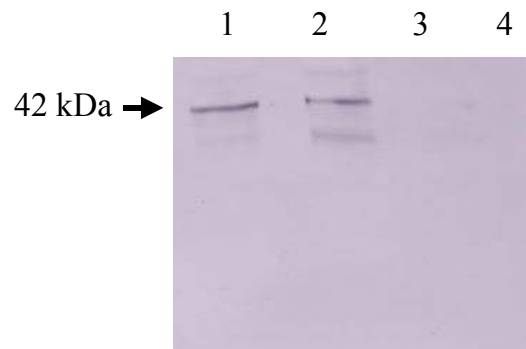


Figure 3.7: Western blot analysis using affinity purified antibody against Theg (1:100) on total testicular protein from different subcellular fractions of testicular cells. 1, nuclear; 2, cytosolic; 3, mitochondrial; 4, total protein extract from kidney. As it can be observed that nuclear fraction shows stronger expression compared to cytosolic fraction, however no expression was detected in mitochondrial fraction.

3.3.5 Theg expression during male germ cell differentiation

To determine the expression pattern of Theg during male germ cell differentiation, paraffin embedded adult mouse testes were cross-sectioned to a thickness of 7-10 μm (which were mounted on a glass slides). These testes sections were then immunostained with antibody against Theg antibody. After an initial blocking step with goat serum in PBS, the slides were incubated with affinity-purified antibody against Theg in a dilution of 1:10. Specific immunostaining as red fluorescence (arising from binding of Cy3 conjugating secondary antibody) was observed only in round spermatids. However, no specific staining was observed in any diploid germ cells, in Sertoli cells or in Leydig cells. It is very interesting to note that staining was also diminished in elongated spermatids and in maturing spermatozoa. When the Theg immunostaining was superimposed with DAPI staining (which stains only the nucleus of the cell), we observed that both labels were overlapping suggesting that Theg is predominately expressed in the nucleus of round spermatids (figure 3.8 A-F).

3.3.6 Intracellular distribution pattern of Theg

Cellular suspension of adult mouse testis was prepared and spread on glass slides. Affinity purified Theg antibody was incubated with the cellular extract on glass slides. Figure 3.8 (panel G-I) illustrates that mouse Theg protein was mainly localised in the nucleus of round spermatids as can be seen by red Cy3 immunostaining of the round spermatids. However other germ cell types and somatic cells failed to stain positive. Cells were counterstained with DAPI and images were superimposed for proper analysis of mouse Theg protein localisation (figure 3.8 G-I), this confirmed a predominant nuclear expression of Theg protein in round spermatids.

In another approach, the intracellular distribution pattern of the mouse Theg protein was examined by transfection of NIH3T3 cells with a *Gfp*-tagged *Theg* construct. Figure 3.8 (panel J-L) shows both a phase contrast microscope image and a microscope image of an NIH3T3 cell expressing the Gfp-tagged mouse Theg protein. The green colour represents Gfp fluorescence from expressed Theg protein and again the blue colour represents fluorescence from DAPI that was used to identify nuclei in these cells. To confirm any apparent nuclear localisation of expressed Theg protein by visualisation of Gfp fluorescence alone, an overlay image of the Gfp and DAPI fluorescence is presented. In this overlay image, the bluish-green

colour represents Theg protein that is localised in the nucleus and as shown Theg protein is localised predominantly in the nucleus of NIH3T3 transfectants (figure 3.8 panel L).

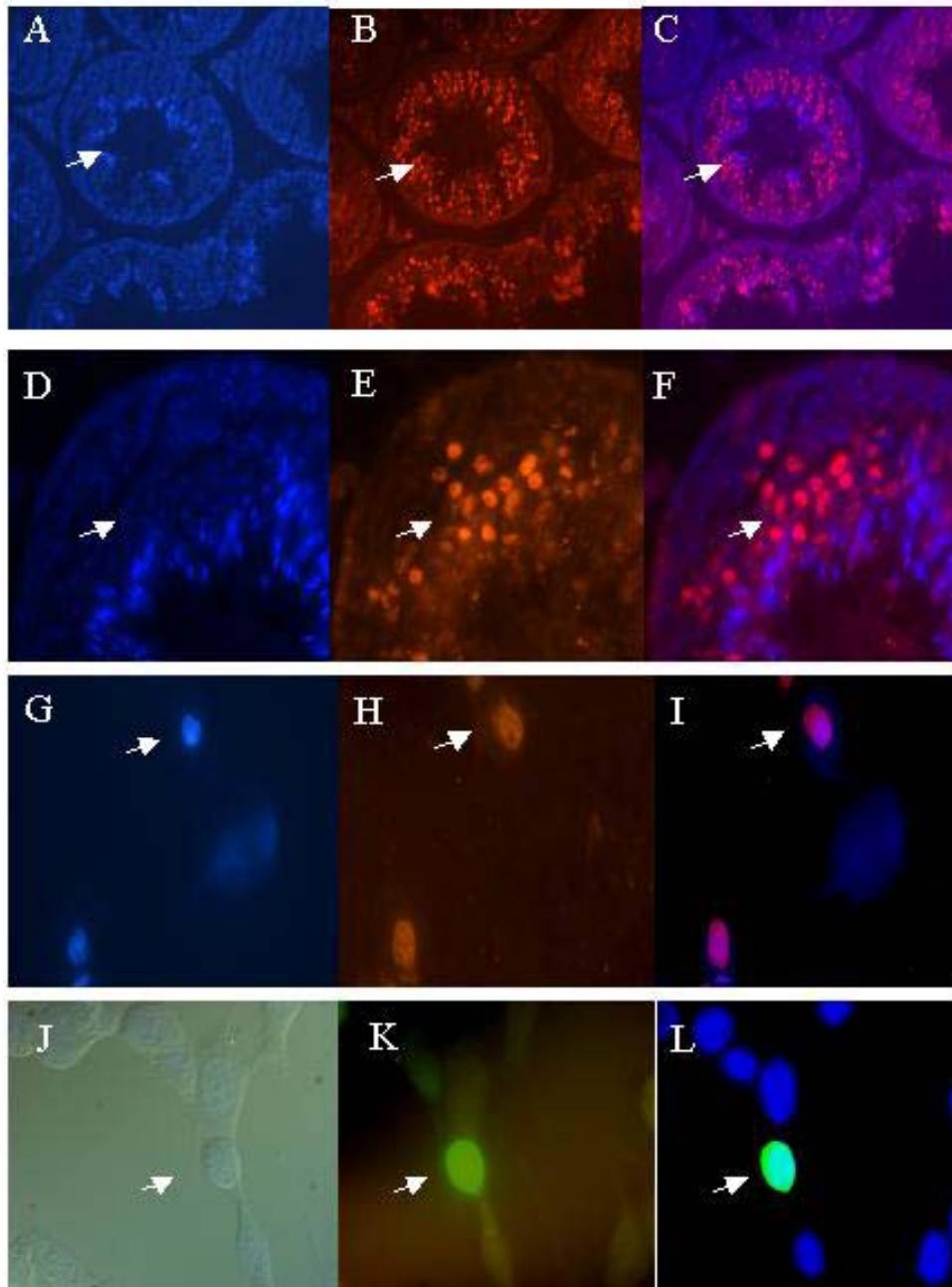


Figure 3.8: Tissue distribution and subcellular localisation of mouse Theg protein. (A) Microscopic image (magnification 200X) of DAPI counterstaining of testicular sections from adult mice. (B) Immunostaining of the same section with an antibody directed against mouse Theg showing predominant nuclear staining in round spermatids. (C) Superimposition of the two labels. (D–F) Higher magnification (600X) of a different area of the

testicular section (staining are identical to the images as shown in the upper panel). (G–I) DAPI counterstaining and immunostaining with anti-mouse *Theg* antibody on cellular suspensions isolated from adult testis showing strong signals in the nucleus of round spermatids. Both images were superimposed and arrows point to the labeled nucleus (magnification 600X). (J) Phase contrast microscopic image of transfected NIH3T3 cells with a Gfp-tagged mouse *Theg* expression vector. (K–L) Microscopic image and overlay image of the same transfected NIH3T3 cell expressing Gfp -tagged mouse *Theg* protein. Green represents Gfp fluorescence from expressed mouse *Theg* protein, blue represents DAPI-stained cell nuclei and bluish-green (in the overlay image) shows overlapping green and blue fluorescence. Arrows point to the positively stained nucleus of the transfected cell (all magnifications 600X).

3.4 Targeted inactivation of mouse *Theg* gene

3.4.1 Genomic structure and restriction digestion analysis of the *Theg* Gene

The mouse genomic library lambda FIX II (Stratagene) was screened using *Theg* cDNA fragments as hybridisation probes. Four positive clones were isolated and were found to cover about 12 kb of continuous genomic DNA (Fig 3.9). The deduced genomic structure was confirmed by genomic Southern blot analyses. The complete cDNA sequence was found to be located within these clones. As shown in figure 3.1, the *Theg* cDNA is divided into 8 exons within a 12 kb segment of genomic DNA. Comparison of the cDNA and genomic sequences indicates that the *Theg* minor cDNA lacks exon 3. In order to generate a complete restriction map of *Theg* gene, genomic clones were digested with different restriction enzymes and separated on 0.8% agarose gel and transferred onto Hybond N membrane and Southern blot was performed with different fragments of *Theg* cDNA namely U3 and O1.

3.4.2 Construction of *Theg* gene targeting vector

A knock-out construct was generated, in which a 4.5 kb *Xba*I fragment named as fragment A (figure 3.9 and 3.10) containing 5' end of *Theg* gene including exons 1, 2, and 3 and 2 kb 5' flanking sequence was cloned into *Sal*I/*Cla*I site of pTKneo vector and 4 kb *Bam*HI fragment named fragment B (figure 3.9 and 3.10) containing exon 8 and 3 kb 3' flanking sequence of *Theg* gene was cloned into *Bam*HI site of pTKneo vector. This construct was designed to delete about 7 kb of *Theg* gene, which included exons 4, 5, 6, and 7, which will be replaced by the *Neomycin* (*Neo*) gene under the control of *Phosphoglycerate kinase* promoter (figure 3.10). *Neomycin* served as a positive selection marker, the construct also contained two copies of herpes virus *Thymidine kinase* gene (*Tk*), at the 3' end of the construct, which served as a negative selection. The orientation of both 5' wing and 3' wing was confirmed by sequencing of the construct with vector specific primers pTKNf, pTKNr, and pTKR.

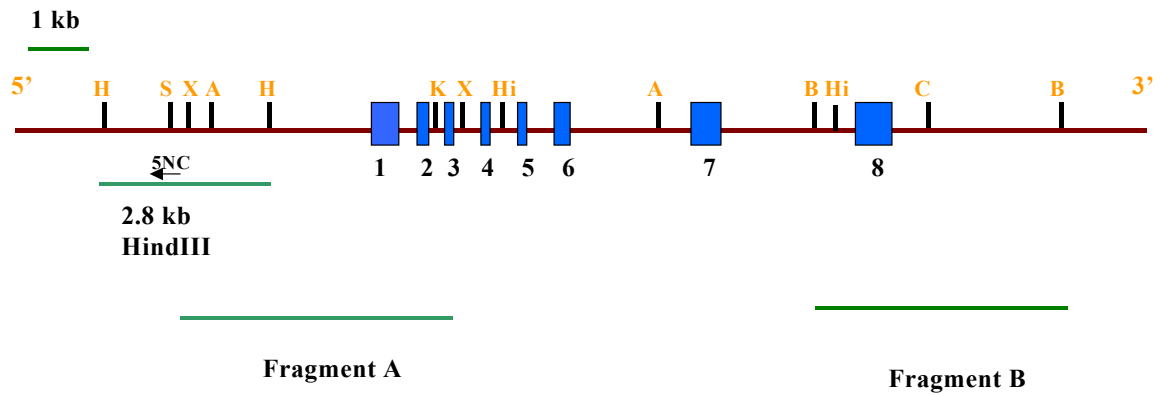


Figure 3.9: Restriction digestion map of theg genomic DNA and fragment which were cloned abbreviations are: A, *AccI*; B, *BamHI*; C, *ClaI*; H, *HindIII*; Hi, *HincII*; K, *KpnI*; S, *SpeI*; X,

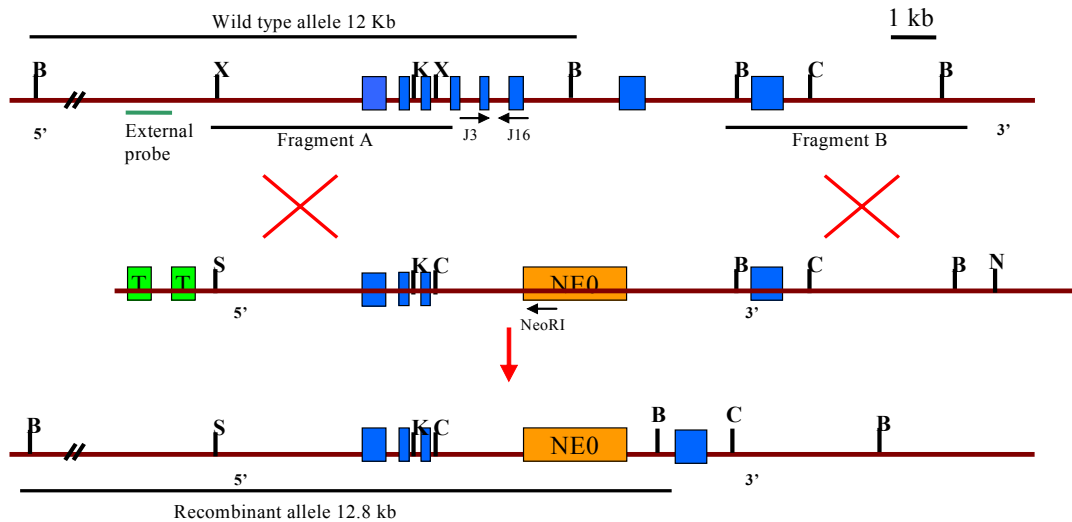


Figure 3.10: Schematic representation of the targeting strategy. Wild type *Theg* locus (top), targeting vector (middle) and mutated allele (bottom). Exons 4, 5, 6, and 7 were replaced by the *Neomycin* resistance gene cassette. External probe is marked as green block, which will be able to recognise a 12 kb *BamHI* fragment in wild type DNA and a 12.8 kb fragment in recombinant DNA. The blue boxes represent exons of *Theg* gene, green boxes stand for *Thymidine kinase* and orange box stands for *Neomycin*. J3, J16 and NeoRI were the primers used for genotyping. The restriction site abbreviations are: B, *BamHI*; C, *ClaI*; K, *KpnI*; N, *NotI*; S, *SalI*; X, *XbaI*.

3.4.3 Identification of a 5' upstream external probe

The lambda genomic clone containing *Theg* genomic fragment was digested with various restriction enzymes, which included *Bam*HI, *Eco*RI, *Hind*III, and *Acc*I and blotted onto Hybond C membrane. A 500 bp *Acc*I fragment from the 5' end of fragment A (figure 3.9) was subcloned in pBluescript and used as a probe in Southern blot analysis in order to identify a genomic fragment upstream of 5' wing of *Theg* knock-out construct (figure 3.10). A 2.8 kb *Hind*III genomic fragment identified through this screen was cloned in *Hind*III site of pZERO vector (figure 3.11). This clone was sequenced, which confirmed that it contained 5' flanking sequence of *Theg* genomic DNA. A 1.2 kb fragment upstream of 5' wing was amplified by PCR using one gene specific primer 5NC (figure 3.9) and one vector specific primer T7 and cloned in pGEMT-easy vector.

Genomic DNA from 129X1/SvJ mice was digested with *Bam*HI and *Kpn*I restriction enzymes and transferred onto Hybond C membrane and Southern blot analysis was done using this 1.2 kb external probe. A 12 kb band was observed in *Bam*HI digestion of genomic DNA (figure 3.12).

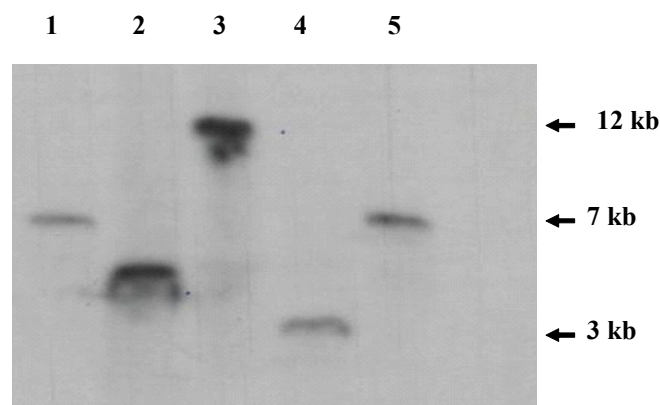


Figure 3.11: Screening for external probe. Genomic lambda clone DNA was digested with various restriction enzymes and blotted with radioactive labelled -500 bp *Xba*I-*Acc*I fragment. Different genomic fragments upstream of exon 1 of *Theg* gene were identified. Lane 1, *Eco*RI digestion shows a 7.5 kb genomic fragment; 2, *Sst*I digestion shows a 4.5 kb fragment; 3, *Bam*HI digestion shows a 12 kb genomic fragment; 4, *Hind*III digestion shows a 3 kb genomic fragment; 5, *Acc*I digestion shows a 7 kb genomic fragment.

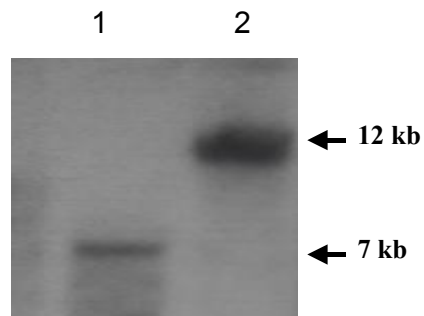


Figure 3.12: Genomic Southern analysis with 1.2 kb external probe. Lane 1, genomic DNA from 129X1/SvJ strain digested with *KpnI*; 2, genomic DNA from 129X1/SvJ strain digested with *BamHI*. A 12 Kb single band was identified on *BamHI* digested DNA.

3.4.4 Electroporation of *Theg* targeting construct into ES cells and screening of ES clones for homologous recombination events

The *Theg* targeting vector was linearised by digesting with *NotI* restriction enzyme and 50 µg of linearised DNA was electroporated into RI ES cell line as described in section 2.2.19.1.3. The cells were plated on fibroblast feeder layer and after 10 days of selection 56 individual *Neomycin* resistant clones were picked in 24 well plates and replicated. Genomic DNA was prepared from these ES clones for Southern blot analysis. DNA of individual ES clones were digested with *BamHI* and electrophoresed into 0.6% agarose gels and blotted onto Hybond C membranes. The blots were hybridised with P³²-labelled 1.2 kb 5' external probe. In the case of homologous recombination event two bands were expected, a wild type allele of 12 kb and recombinant allele of 12.8 kb. However in the event of a random integration, only wild type allele will be detected (figure 3.10). After screening of all the ES clones, two recombinant clones were identified, namely Th13 and Th14. Southern blot for recombinant clone Th14 is shown in figure 3.13.

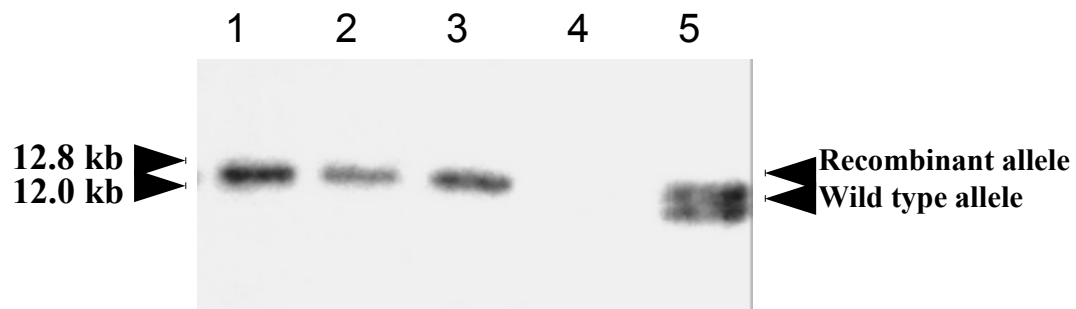


Figure 3.13: Genomic Southern analysis. ES clones genomic DNA were digested with *Bam*HI and separated on 0.6 % agarose gel and transferred onto Hybond C membrane. The blot was hybridised with radioactive labelled 1.2 kb external probe. In case of wild type allele (ES clones 3, 4, and 5) one band of 12 kb was observed and in case of recombinant allele (ES clone 1, which was later name as Th14) two bands were observed, in addition to wild type allele, a mutant allele of 12.8 kb.

3.4.5 Generation of chimeric mice

ES cells from the recombinant clones were injected into 3.5 dpc blastocyst derived from C57BL/6J mice. The blastocysts were implanted into pseudopregnant CD1 mice to generate chimeric mice. Thirty-five chimeras were obtained after four independent injection of two recombinant ES clones named as Th13 and Th14. The chimeras were scored according to coat colour (in percentage). Two chimeras from each line with 90-95 % of chimerism were bred with C57BL/6J and 129X1/SvJ mice to obtain F1 animals in respective background namely C57BL/6J x 129X1/SvJ and in 129X1/SvJ. The germline transmission of *Theg* deleted allele was checked by genomic PCR with J3, J16, and NeoRI primers (figure 3.10) with genomic DNA isolated from tail biopsies of the mice (figure 3.14).

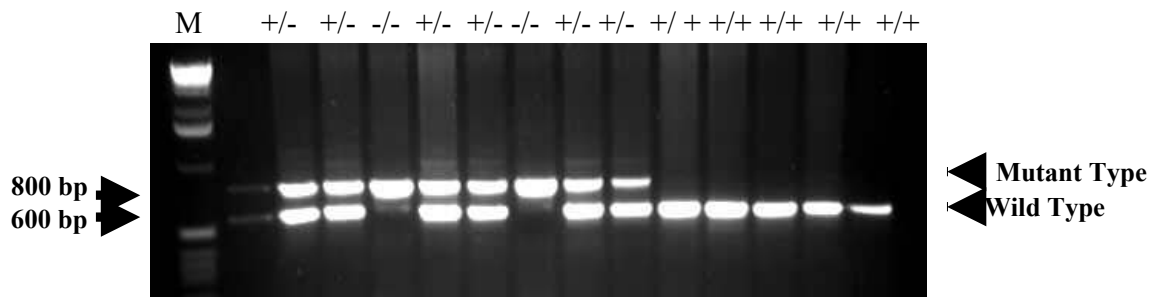


Figure 3.14: PCR genotyping of mice. The wild type allele for *Theg* yields a PCR product of 600 bp using the J3 and J16 primers. The mutated *Theg* allele is expected to generate a fragment of 800 bp using the primers J3 and NeoRI. M, Molecular weight marker; +/+, wild type; +/-, heterozygous; -/-, homozygous mice.

3.4.6 Generation of *Theg* deficient mice

F1 animals, which transmitted the *Theg* deleted allele (that means they were heterozygous at *Theg* locus) were intercrossed to obtain F2 animals. The breeding strategy was undertaken in such a way that two independent lines (Th13 and Th14) were established in both C57BL/6J x 129X1/SvJ and in 129X1/SvJ genetic background. It is noteworthy to mention that heterozygous animals were healthy, appeared phenotypically normal and were fertile. Homozygous mice for *Theg* deleted allele also appeared normal and were fertile. The average litter size derived from their matings was comparable to that of wild type matings (table 3.1 on page 75).

3.5 Analysis of *Theg* expression in knock-out mice

Total testicular RNA was prepared from wild type, heterozygous and homozygous mice. Northern blot analysis was performed with 3' end of *Theg* cDNA (exons 4,5,6, and 7), complete cDNA and *Neomycin* cDNA probes. The blot was re-hybridised with β -actin to show the integrity of the RNA.

Northern blot analysis with 3' cDNA probe showed expression of *Theg* in wild type and heterozygous mice but not in homozygous mice. However when complete *Theg* cDNA was used as a probe a weak band was also detected in homozygous mice testicular RNA. When

the blot was hybridised with *Neomycin* cDNA probe, no transcript was observed in wild type mice but two transcripts of size ~1.2 kb and ~1.4 kb were observed in both heterozygous and homozygous mice (figure 3.15). The size of endogenous *Neomycin* transcript initiated from *Phosphoglycerate* promoter is 1.2 kb, thus the additional transcript detected by *Neomycin* cDNA probe could be a fusion transcript between 5' end of *Theg* and *Neomycin* derived through promoter of *Theg*.

In order to check whether in fact a fusion transcript exists in homozygous mice, RT-PCR was performed with mTHEGf/mTHEFr primer pair and with mTHEGf/NeoRI primer pair (figure 4.1, page 105). PCR product in wild type and heterozygous mice was observed with mTHEGf-mTHEFr primer pair but not with RNA of homozygous mice (figure 3.16). However when mTHEGf-NeoRI primer pair was used PCR product was observed only with RNA of heterozygous and homozygous mice. The fusion transcript was purified from 1% agarose gel and cloned into pGEMT-easy vector. Sequence analysis of this clone showed that exon 1 and 2 of *Theg* transcript were fused with *Neomycin* transcript thus producing a fusion transcript.

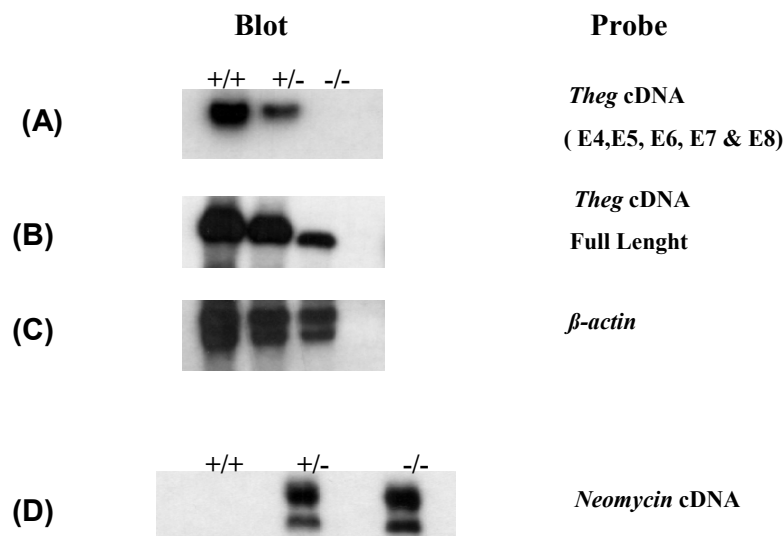


Figure 3.15: Northern blot analysis, testicular total RNA of wild type (+/+), heterozygous (+/-), and homozygous (-/-) mice were hybridised with (A) 3' *Theg* cDNA, which include exon 4, 5, 6, and 7. *Theg* expression can be seen only in +/+ and +/- but not in -/- mice. (B) With full length cDNA, a weaker band was also observed in -/-. (C) *β-actin* re-hybridisation. (D) *Neomycin* cDNA probe showed two transcript in -/- and +/- but not in +/+ mice.

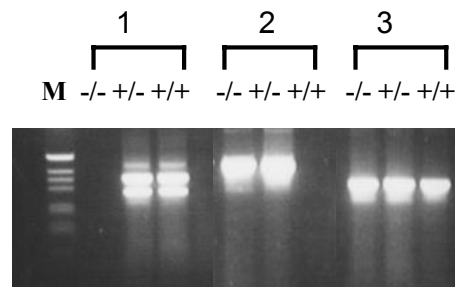


Figure 3.16: RT-PCR analysis for the identification of truncated *Theg* transcript in knock-out mice. Total testicular RNA from wild type, heterozygous and homozygous mice was isolated. 1) RT-PCR using mTHEG1 and mTHEG2 primers, no product was observed in -/- (mice) RNA, however both spliced variants could be seen in +/- (mice) and +/+ (mice) RNA, 2) mTHEG1 and NeoRI primers were used. The fusion transcript can be observed in -/- (mice) and +/- (mice) RNA, but not in +/+ (mice) RNA, 3) control PCR with *Tnp2* primers.

3.6 Phenotypic analysis of *Theg* knock-out mice

3.6.1 Statistical Analysis

In order to determine the mode of inheritance of *Theg* deleted allele in mice, F1 heterozygous mice from lines Th14 and Th13 were intercrossed to obtain F2 generation. The mice were genotyped by genomic PCR on DNA isolated from tail biopsies. The result of this statistical analysis are summarised in figure 3.17.

Line Th14 in genetic background C57BL/6J x 129X1/SvJ showed deviation from Mendelian inheritance in the ratio of wild type (+/+) and homozygous (-/-) progenies; with a chi square value 38.06, $p < 0.000001$. However in genetic background 129X1/SvJ, line Th14 was in agreement with Mendelian inheritance with chi square value of 2.6, $p < 0.27$.

A						B					
Th14/ C57BL/6J x 129X1/SvJ						Th14/ 129X1/SvJ					
+/+		+/-		-/-		+/+		+/-		-/-	
M	F	M	F	M	F	M	F	M	F	M	F
47 (39.5%)	55 (44%)	52 (43.7%)	50 (40%)	20 (16.8%)	20 (16%)	28 (30%)	20 (21%)	41 (44%)	51 (52%)	24 (26%)	26 (27%)
102 (41.8%)		102 (41.8%)		40 (16.4%)		48 (25%)		92 (49%)		50 (26%)	

Figure 3.17: Statistical analysis of genotype for *Theg* locus of mice from F2 generation. (A) Line Th14/ C57BL/6J x 129X1/SvJ showed deviation from Mendelian inheritance, there were more wild type animals than expected, in total 244 animals were genotyped. (B) Line Th14/129X1/SvJ was in accordance with Mendelian segregation, here 190 animals were genotyped. M, male; F, female.

3.6.2 Growth curve of testes during mouse development

Testes of *Theg* knock-out mice appeared smaller in size than wild type mice therefore a growth curve of testes during mouse development was plotted for wild type, heterozygous and homozygous mice. Testes from different developmental stages were dissected under aseptic condition and weighed. For each developmental stage, testes of three mice were weighed and the mean weight was taken. The result of growth curve is summarised in figure 3.18.

The growth curve shows a reduced testis weight of homozygous mice as compared with wild type and heterozygous mice.

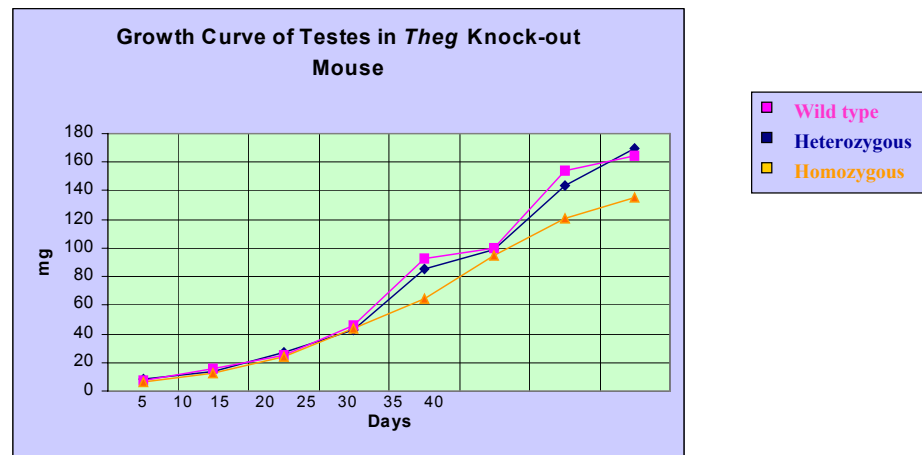


Figure 3.18: Growth curve analysis of testes development. Adult homozygous mice testis are smaller than wild type and heterozygous mice testes.

3.6.3 Sperm motility analysis

To characterise sperm motility, sperms were prepared from wild type, heterozygous, and homozygous mice as described in section 2.2.20. A computer assisted sperm analyser was used and following parameters were evaluated; path velocity (which represents the total distance passed through by a sperm in a unit time), track speed, progressive velocity (which represents the straight line progressive movement of the sperm between the beginning and the end of the measurement divided by the time elapsed), straightness, linearity, lateral amplitude and beat frequency. The result of sperm motility analysis is summarised in figure 3.19.

Homozygous mice showed a reduced motility compared to wild type mice in total path velocity, progressive velocity and track speed. However the reductions were subtle and did not affect sperms ability to fertilise oocytes.

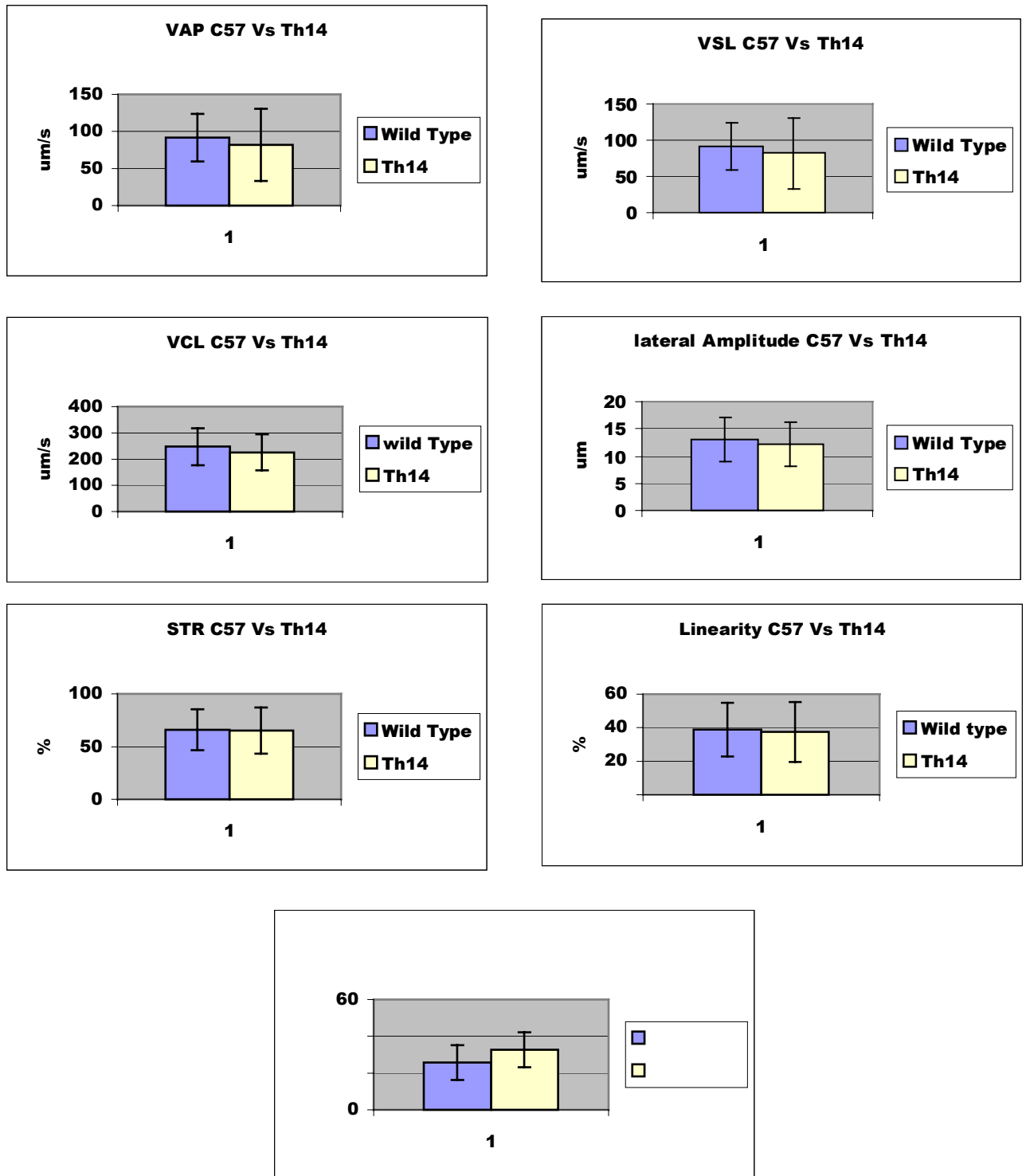


Figure 3.19: Results of computer assisted analysis of sperm motility. Different parameters of sperm motility were measured, which includes: total path velocity (VAP), progressive velocity (VSL), track speed (VSL), lateral amplitude, straightness (STR), linearity and beat frequency. Th14 knock-out mice sperms showed only minor deviations from wild type sperms.

3.6.4 Subcellular localisation of mutant *Theg* protein

In order to determine the subcellular localisation of mutant *Theg* protein, which arose due to fusion of exon 1 and exon 2 of *Theg* with *Neomycin* gene, we undertook two different approaches; firstly we performed immunostaining of testicular suspension from *Theg* homozygous mutant mouse and as a control wild type mouse with polyclonal antibody specific for epitope present in exon 2 of *Theg*. As can be seen in figure 3.20 (D-F) that in round spermatids from homozygous mice *Theg* fusion protein was localised mostly in cytosol in contrast to wild type mice (figure 3.20 A-C), where *Theg* was localised predominantly in nucleus. In a second approach, we generated a *Gfp* tagged mutant *Theg* construct and transfected it into NIH3T3 cell line. The mutated *Theg* showed a predominant cytosolic localisation in the NIH3T3 cell line (figure 3.20 G-I); in contrast *Gfp* tagged wild type *Theg* showed a nuclear localisation (figure 3.20 J-L). These results suggests that *Theg* fusion protein is predominantly localised in cytosol in contrast to wild type *Theg* protein, which is mainly present in the nucleus of spermatid cells.

3.7 Generation of *Theg* knock-out mice (Th14) in C3H/J background

From the results presented above, we concluded that *Theg* knock-out mice are fertile. However Yanaka et al. 2000, reported that an insertional mutation of *Theg* in C3H/J background resulted in infertility due to arrest in spermatogenesis at elongated spermatid stage. To examine the fertility of Th14 mice in C3H/J background, homozygous mice for Th14 locus in 129X1/SvJ background were bred with C3H/J mice to obtain heterozygous mice, which were intercrossed to obtain F2 animals. The homozygous mice obtained in F2 generation were backcrossed to C3H/J mice. This breeding strategy was pursued for seven generation to obtain high level of genetic homozygosity for C3H/J strain except for Th14 locus. Th14 knock-out mice in C3H/J background were found to be fertile and average litter size was normal (table 3.1).

Strain	Average litter size
129X1/SvJ	5.9
Th14 (129X1/SvJ)	4.1
Th14 (C57BL/6J)	8.7
Th14 (C3H/J)	7.1

Table 3.1: Summary of average litter size of Th14 mice in different genetic background.

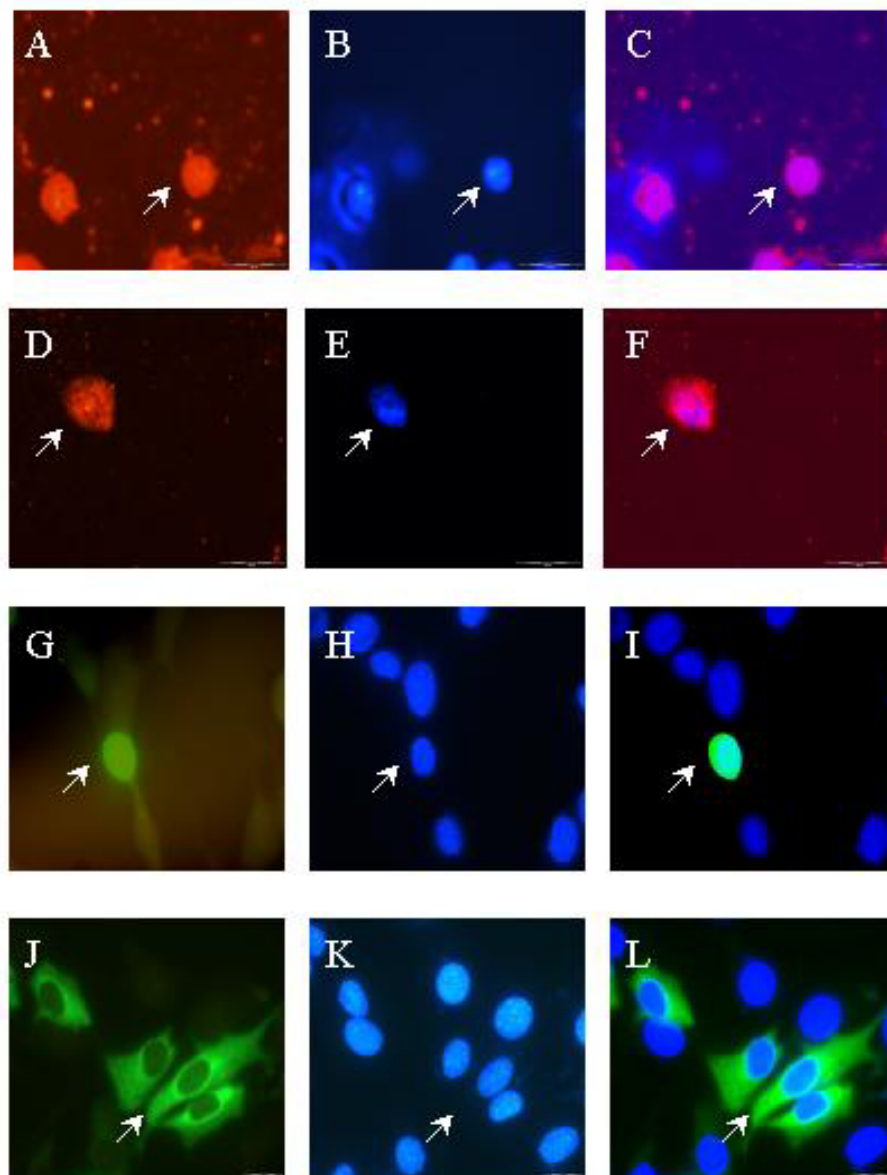


Figure 3.20: Subcellular localisation of mutated Theg protein. (A) Immunostaining with anti-mouse Theg antibody on cellular suspensions isolated from adult testis showing strong signals in the nucleus of round spermatids. (B) DAPI counterstaining of image (A). (C) Superimposition of the both labels. (D) Immunostaining on cellular suspensions isolated from adult testis of knock-out mice showing mainly cytosolic signal. (E) DAPI counterstaining of image (D). (F) Superimposition of the both labels. (G-I) Microscopic image, DAPI counterstaining and overlay image of the same transfected NIH3T3 cell expressing Gfp-tagged mouse Theg protein showing a predominant nuclear expression. (J-L) Microscopic image and overlay image of the same transfected NIH3T3 cell expressing Gfp-tagged mutated Theg protein showing a cytosolic expression.

3.8 Generation of new *Theg* knock-out mice deleting 5' end of the gene

In Th14 knock-out mice, the N-terminal of *Theg* protein (1-136 amino acids) was not deleted and a fusion protein between N-terminal *Theg* and *Neomycin* was generated. In order to check whether this N-terminal domain of *Theg* plays any functional role, we decided to delete this domain in a new knock-out mice.

3.8.1 Restriction digestion analysis

A 17 kb *Theg* genomic DNA sequence and its flanking region was downloaded from the Celera database using Celera Discovery System (CDS) mouse genome (www.celera.com). The genomic structure of the *Theg* gene was confirmed by comparing *Theg* cDNA sequence with the downloaded genomic sequence using BLAST 2 program (www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). A restriction map was created using NEB cutter program (<http://tools.neb.com/NEBcutter/index.php3>) (figure 3.21).

3.8.2 Cloning of 5' wing of the new knock-out construct

A 2.8 kb *Hind*III fragment upstream of *Theg* transcription start point (fragment A in figure 3.21) was chosen as 5' wing of new construct. The lambda genomic clone containing *Theg* genomic fragment was digested with *Hind*III and separated on 0.8 % agarose gel and this 2.8 kb *Hind*III fragment was purified and cloned into corresponding site of pZERO vector. The clone 2.8 kb *Hind*III-pZERO was further digested with *Spe*I to obtained a 1.6 kb *Theg* genomic fragment, which was then cloned into *Spe*I site of pTKneo vector (figure 3.22). The orientation of the genomic fragment in pTKneo vector was checked by sequencing with PTKNf primer.

3.8.3 Cloning of 3' wing of the new knock-out construct

A 5 kb *Hinc*II genomic fragment, which contained exonic sequences of exons 5,6,7, and 8 was chosen as 3' wing of the targeting construct (fragment B in figure 3.21). This *Hinc*II fragment was isolated from a *Hinc*II digested lambda *Theg* genomic clone. After purification of the fragment from 0.8% agarose gel it was cloned into *Hinc*II site of pBluescript (SK) vector. The orientation of the clones were determined by sequencing and clone with correct

orientation was digested with *ClaI* and *XhoI* enzyme and the resulting genomic fragment was cloned in a directed manner into pTKneo vector which was digested with *ClaI* and *SalI* restriction enzymes. It is important to mention that *XhoI* and *SalI* digested DNA have compatible cohesive ends (figure 3.21).

3.8.4 Identification of 5' external probe

In order to generate an external probe, primers expF and expR were designed on the basis of *Theg* genomic sequence to amplify a 900 bp fragment upstream of 2.8 kb *HindIII* fragment. The locations of primers are shown in figure 3.21. After PCR amplification the 900 bp fragment (external probe) was purified and cloned into pGEMT-easy vector. Genomic DNA from 129X1/SvJ and C57BL/6J mice was digested with *KpnI* enzyme and a Southern blot analysis was performed using P³²-labelled external probe. A 7 kb fragment was observed in the digested DNA from both strains (figure 3.22), which was also expected from restriction digestion analysis. This Southern blot also confirmed that there is no RFLP between 129X1/SvJ and C57BL/6J mice at this locus.

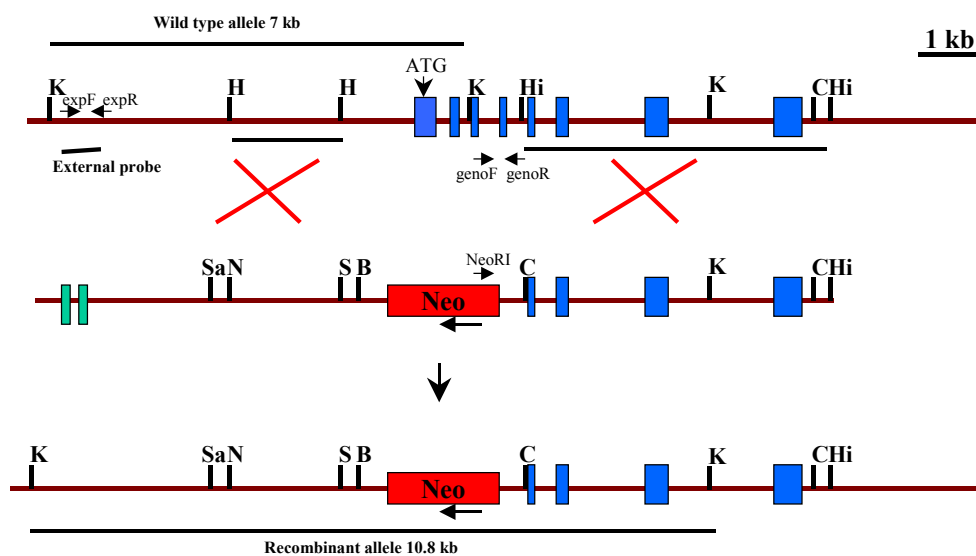


Figure 3.21: Schematic representation of the new *Theg* targeting strategy. Wild type locus (top), targeting vector (middle), and mutated allele (bottom) are shown. A 2.8 kb *HindIII*

HindIII; Hi, *HincII*; K, *KpnI*; Sa, *SstII*; S, *SpeI*

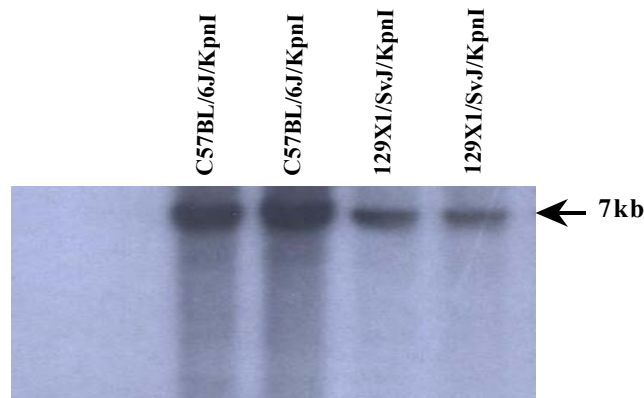


Figure 3.22: Identification of external probe. Southern blot was performed with a 900 bp fragment on *KpnI* digested genomic DNA from C57BL/6J and 129X1/SvJ mice. A single band of 7 kb was observed, which was in agreement with restriction digestion pattern

3.8.5 Electroporation of new *Theg* targeting construct into ES cells and screening of ES clones for homologous recombination events

The new *Theg* targeting construct which was designed to delete 5' end of *Theg* gene was linearised by digesting with *SstII* restriction enzyme and 50 μ g of linearised DNA was electroporated into RI ES cell line as described in section 2.2.19.1.3. The cells were plated on fibroblast feeder layer and after 10 days of selection 96 individual *Neomycin* resistant clones were picked in 24 well plates and replicated. DNA was prepared from these ES clones for Southern blot analysis. Individual ES clones were digested with *KpnI* and electrophoresed into 0.6% agarose gels and blotted onto Hybond C membranes. The blots were hybridised with P^{32} -labelled 900 bp external probe. In the case of homologous recombination event two bands are expected, a wild type allele of 7 kb and a recombinant allele of 10.8 kb (figure 3.21 and 3.23). However in the event of a random integration, only the wild type allele will be detected (figure 3.23). After screening of 48 ES clones, 17 recombinant clones were identified. The blots were re-hybridised with P^{32} -labelled *Neomycin* probe to rule out any multiple integration of targeting vector into the genome.

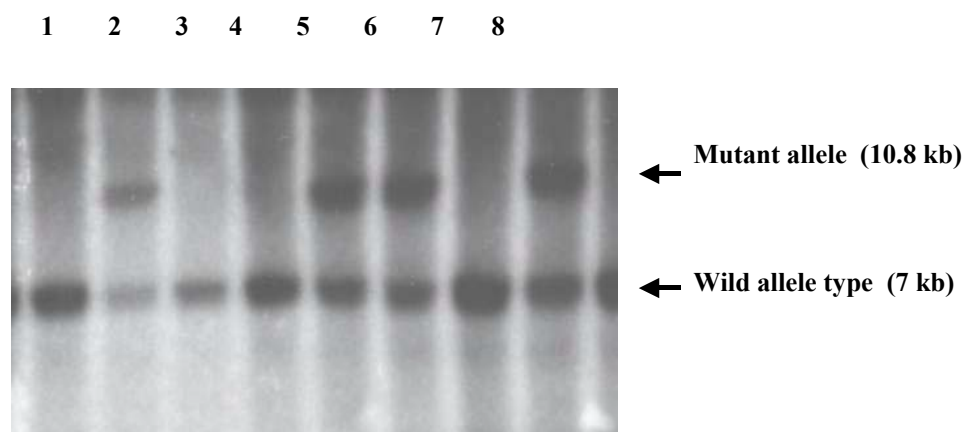
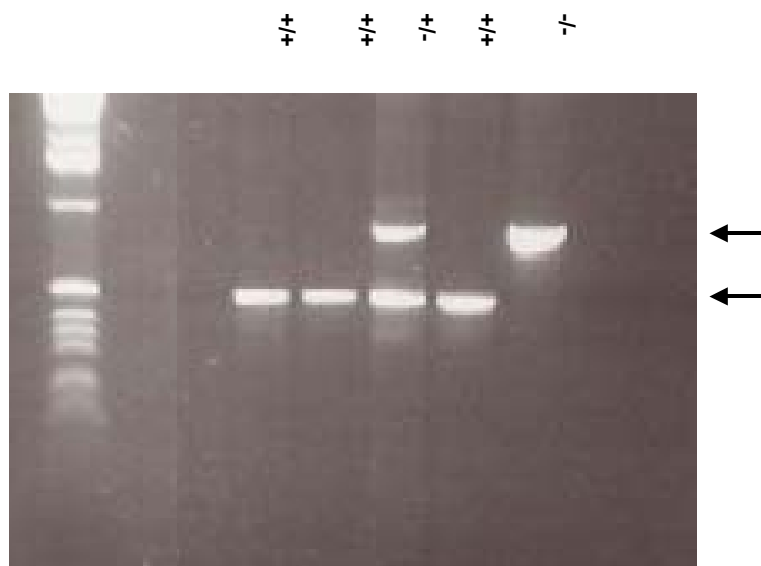


Figure 3.23: Screening of ES clones for homologous recombination. Southern blot was performed with 900 bp external probe on *KpnI* digested genomic DNA of ES clones. A 7 kb band can be detected in wild type allele and mutant allele will give rise to 10.8 kb fragment. ES clones 2, 5, 6 and 8 are recombinant clones.



3.8.6 Generation of chimeric mice

ES cells from one of the recombinant clones (clone 39) were injected into 3.5 dpc blastocysts derived from C57BL/6J mice. The blastocysts were implanted into pseudopregnant CD1 mice to generate chimeric mice. 7 chimeras were obtained after two independent injections of the recombinant ES clone 39. The chimeras were scored according to the coat colour (in percentage). Two chimeras with 90-95% of chimerism were bred with C57BL/6J and 129X1/SvJ mice to obtain F1 animals in each background namely C57BL/6J x 129X1/SvJ and in 129X1/SvJ. The germline transmission of *Theg*-deleted allele was checked by genomic PCR with genoF, genoR, and NeoRI primers (figure 3.21) with genomic DNA isolated from tail biopsies of the mice (figure 3.24).

3.8.7 Generation of *Theg* deficient mice

F1 animals, which transmitted the *Theg*-deleted allele (heterozygous at *Theg* locus) were intercrossed to obtain F2 animals. The breeding strategy was undertaken in such a way that new *Theg* deleted locus was established in both C57BL/6J x 129X1/SvJ and in 129X1/SvJ genetic background. It is noteworthy to mention that heterozygous animals were healthy, appeared phenotypically normal and were fertile. Homozygous mice for *Theg*-deleted allele also appeared normal and fertile. Detail phenotypic and physiological analyses to identify any subtle abnormalities are awaited.

3.9 Characterisation of human *THEG* gene

3.9.1 Cloning of full-length *THEG* cDNA

To clone the human homologue of mouse *Theg*, we first searched the Genbank database for the presence of human DNA sequences similar to mouse *Theg*. By using this method, a human EST clone with high homology to mouse *Theg* (accession No. AJ007421) was identified. This human EST clone was sequenced in its entirety and sequence comparison of the human EST clone with mouse *Theg* cDNA showed an overall similarity to the 3' coding region of *Theg*. To obtain the full-length cDNA 5'-RACE was performed on total RNA isolated from human adult testis. A cDNA fragment of 813 bp was amplified and sequence analysis of this PCR product revealed that the cDNA fragment contains additional 613 bp upstream of the human EST clone (figure 3.25).

```

gaccggtgt 9
atggggggacagcaggcgaaggtcactcgggaaccagcccagctct 63
  M G D S R R R S L G N Q P S S 15
gaggctgcgggcaggtcggaaagggagcaggacggcgacccccgt 108
  E A A G R S E R E Q D G D P R 30
ggcctccagagctctgtgtacgagagccggcggtcacagacccc 153
  G L Q S S V Y E S R R V T D P 45
gaacgccaggacctggacaatgcagagctgggaccagaagaccca 198
  E R Q D L D N A E L G P E D P 60
gaagaggagcttccccccgaggaggtggccggggaggagttcccg 243
  E E E L P P E E V A G E E F P 75
gagaccctggatcccaaagaggcactttctgagttggagagagtc 288
  E T L D P K E A L S E L E R V 90
ctggacaaggacttgggaaggacattcctgaaatcagcgggctg 333
  L D K D L E E D I P E I S R L 105
tccatcagccagaagctccccagcaccacatgaccaaagcaagg 378
  S I S Q K L P S T T M T K A R 120

aagaggaggagggcggaggaggctcatggagctggcagagcccaag 423
K R R R R R R L M E L A E P K 135
JH3
ataaaactggcaagtcctgaaagacaggaagggacgctgtggtaag 468
  I N W Q V L K D R K G R C G K 150
gggtatgcctggatctcccatgtaagatgagcttgcaacttctgt 513
  G Y A W I S P C K M S L H F C 165
ctctgctggccctctgtgtactggaccgagcggttccttgaggac 558
  L C W P S V Y W T E R F L E D 180
accaccctcaccatcacagtgcccgcggtgtcccgcgctggag 603
  T T L T I T V P A V S R R V E 195
gaactgtctcggccaagagattctacctggaatattacaacaac 648
  E L S R P K R F Y L E Y Y N N 210
aacaggagcactcctgtctggccattcctcggctcctccctggaa 693
  N R T T P V W P I P R S S L E 225
tacagagcgtcgagtcgcctgaaggaactggccgccccgaagatt 738

Y R A S S R L K E L A A P K I 240
cgtgataacttctggagcatgcccatgtctgaggtgtcccaggta 783
JH4
R D N F W S M P M S E V S Q V 255
tccagggcagcccaatggcagtcctccagctcgcggatcctccag 828
  S R A A Q M A V P S S R I L Q 270
ttgtcaaagccgaaggccccagccaccctcttggagagtgaggac 873
  L S K P K A P A T L L E E W D 285
cccgtgccaaaaccaagccacatgtgtcagaccataaccgcctc 918
  P V P K P K P H V S D H N R L 300
cttcacttggccagggcccaaagctcagtcggacaagtgcgcttct 963
  L H L A R P K A Q S D K C V P 315
gaccgagatcctcgtctgggaggtgctggatgtcaccaagaagtg 1008
  D R D P R W E V L D V T K K V 330
gtggccagcccccgatcatctcctggccaagcccaaagtgcgc 1053
  V A S P R I I S L A K P K V R 345
aagggcctcaacgagggatagcagggcgtcccctcgcctctatg 1098
  K G L N E G Y D R R P L A S M 360
agcttgccacccccaaaagcatcaccagaaaagtgtgatcaacc 1143
  S L P P P K A S P E K C D Q P 375

```

```

aggcctggcctctaagacctccgctcccagtaaacaccctcaggc 1188
  R P G L *                                     379
accctaaccctgtgtatgtgattattctgagctttttggcctgga 1233
gaagggagggcgggctagaaggcctaaaagaaggatgaggaggaca 1278
ataatatttttttttgcattcc                       1291

```

Figure 3.25: Nucleotide sequence and deduced amino acid sequence of both human *THEG* major and *THEG* minor cDNAs. Numbering of nucleotide and amino acid sequences are given on the right. The start of exon sequences and the modified putative polyadenylation signal are marked in bold type and are underlined. The spliced out exon 3 sequences in the *THEG* minor cDNA sequence is marked by blue bold font. The translational start codon is marked with bold letter and termination codon is marked with asterisk. The positions of primers used for RT-PCR are shown by arrows. Putative nuclear targeting signals are marked in the amino acid sequence with bold type.

3.9.2 Chromosome localisation and genomic organization of *THEG*

To investigate the chromosome location of *THEG*, we first amplified a 1.0 kb genomic fragment by PCR on human genomic DNA using primers corresponding to the sequence of clone 19pTEL009 (Ning et al, 1996). This PCR product was used as a probe to screen a human PAC library (Ioannou et al., 1994) and a single hybridising PAC clone (PAC RPCIP704D04835Q3) was isolated. The PAC DNA was subjected to different restriction digestions and hybridised with a 240 bp *Bam*HI fragment generated from the 5' end of the *THEG* cDNA. After cloning and sequence analysis of different genomic fragments, we could verify that the isolated PAC clone contained identical sequences as compared to the *THEG* cDNA including the entire exon 1, part of intron 1 and about 1 kb of the 5' untranslated region. Using this PAC clone as a probe for FISH, we obtained twin signals specifically within the chromosome region 19ptel-p13; no signals could be detected on any other chromosome among 100 metaphase cells examined (figure 3.26). Due to the release of the working draft sequences of the human genome in the Genbank database, we were able to identify two genomic clones on chromosome 19 (clone CTD-3113P16, accession No. AC016588 and clone LLNLFOS-22E10, accession No. AC010641), which contain the complete *THEG* cDNA sequence. As shown in figure 3.27, the *THEG* cDNA is divided into eight exons within a 14 kb segment of genomic DNA.

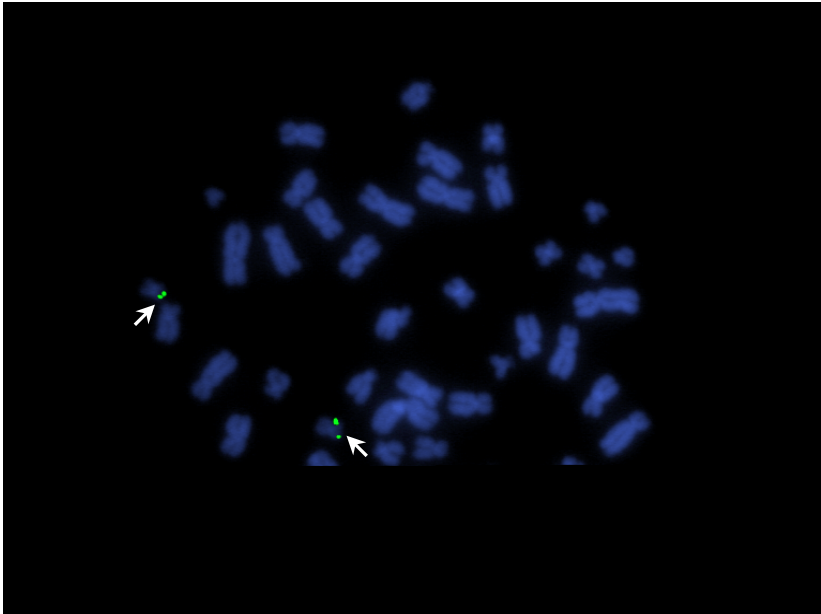


Figure 3.26: Chromosomal localisation of the *THEG* gene by fluorescence in situ hybridisation (FISH) to human metaphase chromosomes. Arrows point to specific *THEG* signals on chromosome 19ptel-p13.

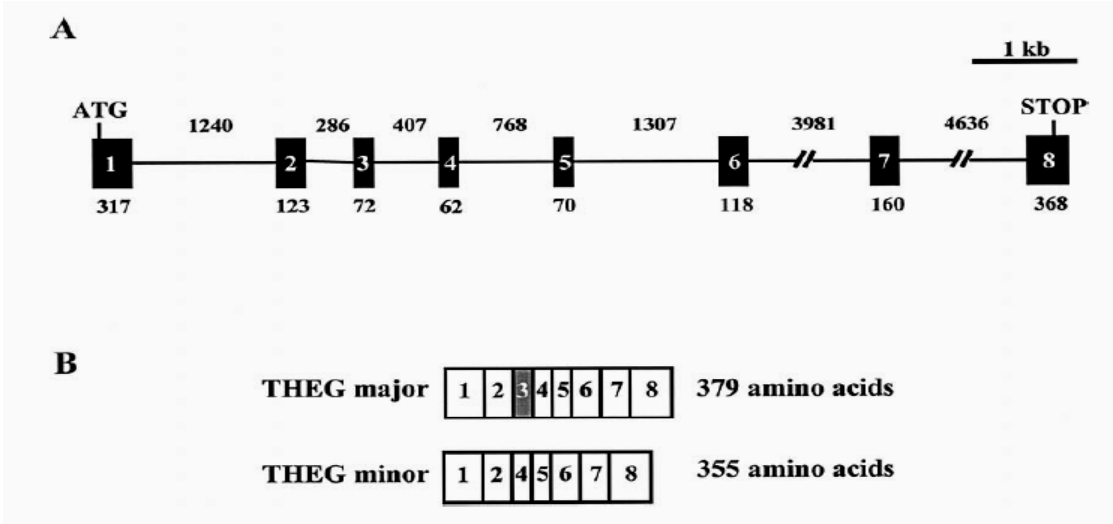


Figure 3.27: (A) Schematic representation of the exon-intron structure of *THEG*. Exons and introns numbers inside the exons refer to the order of exons. Numbers directly beneath the exons and directly above the introns refer to their exact lengths (in base pairs). Translational start and stop sites are indicated. (B) Schematic illustration of the deduced human *THEG* major and *THEG* minor proteins. As can be seen exon 3 is spliced out of the *THEG* minor transcript producing a 24 amino acid smaller human *THEG* minor protein.

3.9.3 Expression and alternative splicing of the *THEG* gene

The tissue-specific expression of the *THEG* gene was evaluated in a first experiment by RNA dot blot analysis on a large panel of human adult and fetal tissues, as well as tumour cell lines (Multiple Tissue Expression array, Clontech) using the *THEG* EST clone as a probe. The blot also included various control RNAs and DNAs (A in figure 3.28). In order to determine if alternatively spliced *THEG* transcripts are expressed in human testis, we performed RT-PCR on total human testicular RNA using primers JH3 and JH4 located in exon 2 and exon 6, respectively. Two RT-PCR products, which were of differing lengths could be detected in 1.6% agarose gel (352 bp and 280 bp, respectively), which were named *THEG* major and *THEG* minor (B in figure. 3.28). Sequence analyses of *THEG* major and *THEG* minor cDNA revealed that their DNA sequences were identical, except that they differed by 72 bp in length (from nucleotide position 440 to nucleotide position 512 in figure 3.25). The human *THEG* major cDNA sequence is 1291 bp in size (without the poly (A)⁺ tail) and was registered in the Genbank database (accession No. AF268610).

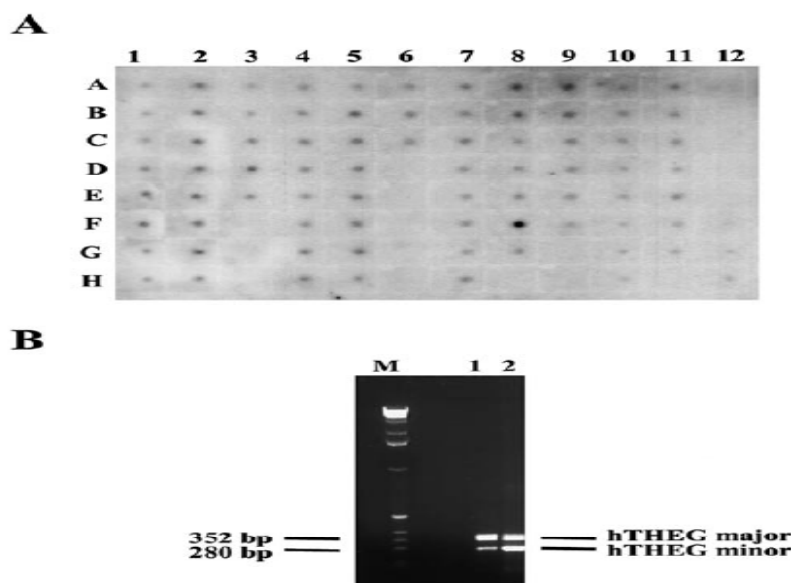


Figure 3.28: (A) RNA dot-blot (Multiple Tissue Expression Array, Clontech) analysis to detect *THEG* mRNA expression in a panel of normal human tissues (adult and embryonic) and cancer cell lines. The position of the samples on the membrane is indicated by numbers and letters. A strong expression signal was observed exclusively on RNA from adult testis (position F8 on the membrane) by using human *THEG* cDNA as a probe. (B) RT-PCR analysis on total RNA isolated from human testis using *THEG*-specific primers JH3 and JH4. Both *THEG* major (352 bp) and *THEG* minor (280 bp) transcripts can be detected by RT-PCR analysis. M, standard molecular weight marker. 1, patient1; 2, patient 2.

3.10 Identification and characterisation of *nax* mutant mice

3.10.1 About *nax* mutation

nax is an autosomal recessive mutation, which arose in mice (C57BL/6J x 129X1/SvJ) during the generation of *Theg* (Th14) knock-out mice. The mutation could have arose because of two reasons:

- 1) A random insertion of partial fragment of *Theg* targeting construct into mouse genome.
- 2) A spontaneous mutation in the ES cell line.

3.10.2 Identification of *nax* locus in mouse genome

As discussed before in section 3.11.1 there could be only two events, which could give rise to *nax* mutation. It is either a spontaneous mutation or a partial insertion of *Theg* targeting vector into *nax* mouse genome. To address the latter possibility, i.e. a random integration of *Theg* construct, genomic DNA from *nax* mice that were wild type for *Theg* locus and wild type DNA were digested with various restriction enzymes and separated on 0.6% agarose gel. The digested genomic DNAs were transferred onto Hybond C membrane and blotted with P³²-labelled probe.

For above Southern analysis, *Theg* targeting construct was fragmented into five different parts using restriction enzymes and used as probes for Southern blot (figure 3.29).

1. 4.5 kb *Xba*I fragment, used before as 5' wing of construct
2. 4.0 kb *Bam*HI fragment, used before as 3' wing of construct
3. *Neomycin*
4. *Thymidine kinase*
5. pBluescript without *lacZ*

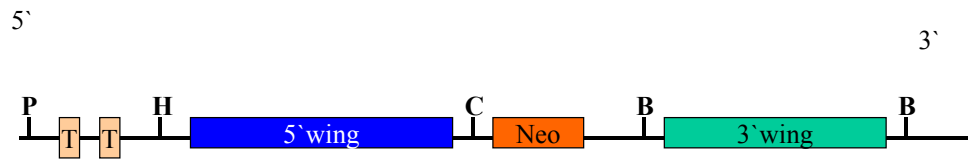


Figure 3.29: Schematic diagram showing five different parts of *Theg* knock-out construct, which were used for Southern blot analysis for identification of *nax* mutation. B, *Bam*HI; C, *Cla*I; H, *Hind*III; P, *Pst*I.

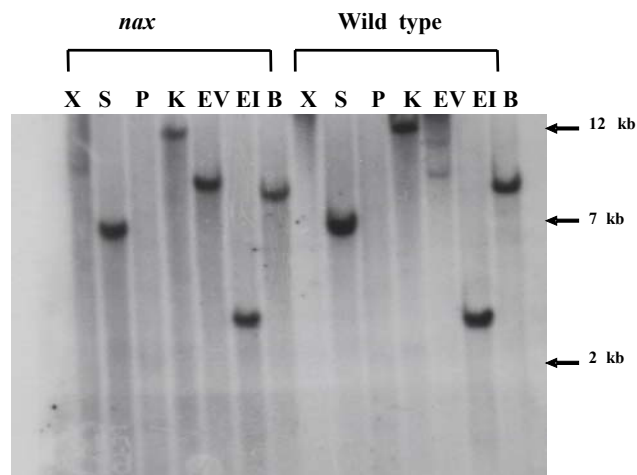


Figure 3.30: Southern blot analysis using 5' wing of *Theg* knock-out construct as a probe on *nax* mice and wild type mice genomic DNA, which were digested with different restriction enzymes. No difference in pattern between *nax* DNA and wild type DNA can be observed. B, *Bam*HI; EI, *Eco*RI; EV, *Eco*RV; K, *Kpn*I; P, *Pst*I; S, *Sst*I; X, *Xho*I.

As can be seen in figure 3.30, no difference in Southern blot pattern between *nax* mice and wild type mice was observed, when 5' wing of *Theg* knock-out construct was used as a probe. Similar results were observed when genomic Southern blot were performed with other parts of *Theg* construct, thus ruling out any integration of *Theg* construct into *nax* mouse genome.

3.10.3 Phenotypic description of *nax* mutant mice

After birth *nax* mice appear comparatively smaller than their wild type littermates. They showed a much delayed hair development. Normally in mice hair appearance starts around stage P10. However *nax* mice are completely naked till late stages (P35-P40) of development. Though hairs do appear in adult *nax* mice but only in the anterior part of the body (figure 3.31). However the most interesting observation is that *nax* mice show characteristic symptoms of ataxia around stage P10-P12, they walk incessantly inside the cage and fall down frequently. Apparently they show clear signs of motor imbalance and lack of coordination of body movements.

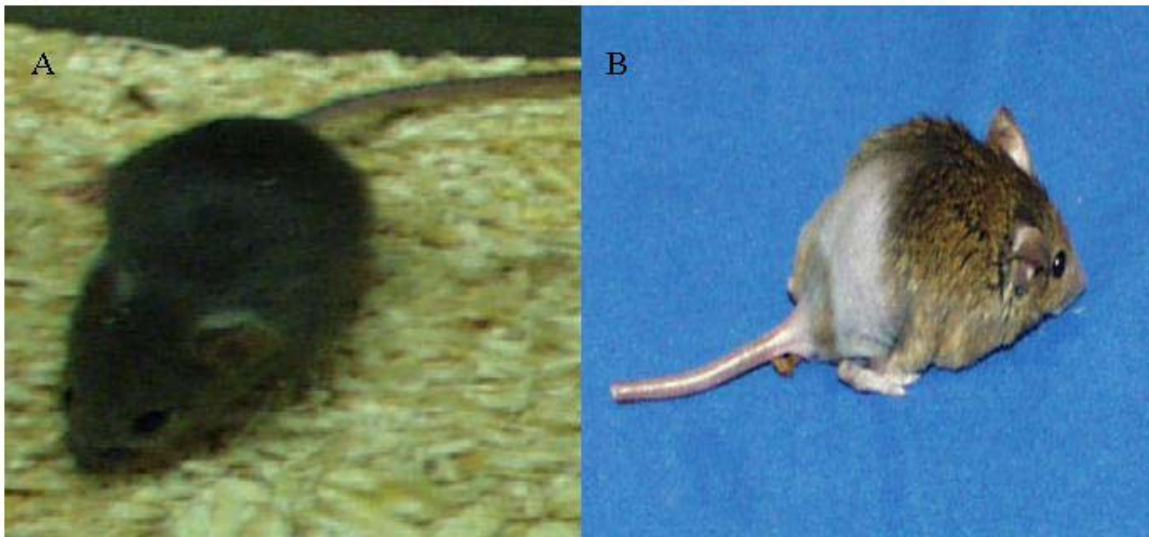


Figure 3.31: (A) Adult wild type mouse (strain 129X1/SvJ). (B) Adult *nax* mouse lacks hair in the posterior part of their body. They walk incessantly and have motor imbalance.

3.10.4 Growth curve analysis

As *nax* mice show apparent growth retardation, we investigated the growth curve of *nax* mice. *nax* mice from different developmental stages and as control normal littermates were weighed. Three *nax* mice and three normal mice were weighed with a period size of 3 days for duration of 45 days (from the same litter). The mean weight was taken for both *nax* and normal mice and plotted against period days. The result of growth curve is summarised in figure 3.32 and it is evident that *nax* mice show dramatic growth retardation and adult *nax* mice showed more than 50% reduction in weight compared to wild type adult mice.

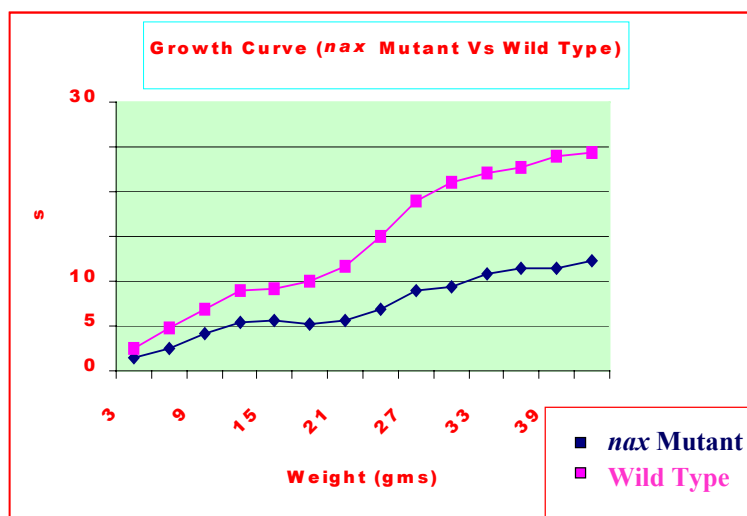


Figure 3.32: Growth curves. *nax* mice show a severe growth retardation, adult *nax* mice are less than half the weight compared to their wild type littermates.

3.10.5 Analysis of brain of wild type and *nax* mice

Macroscopic inspection of *nax* mice brain showed a pronounced reduction in the brain size. When paraffin embedded brain sagittal sections were stained with crystal violet and observed under microscope, all main morphological subdivisions of brain were observed, but subtle structural disturbances in different subdivisions were observed. However, the most affected structure was cerebellum. Though cerebellum is still present in mutant mice, however cerebellum cortical structures were severely impaired (figure 3.33). More detailed analysis with respective markers, for the granular layer, Purkinje cell layer would be necessary to pinpoint the exact phenotype.

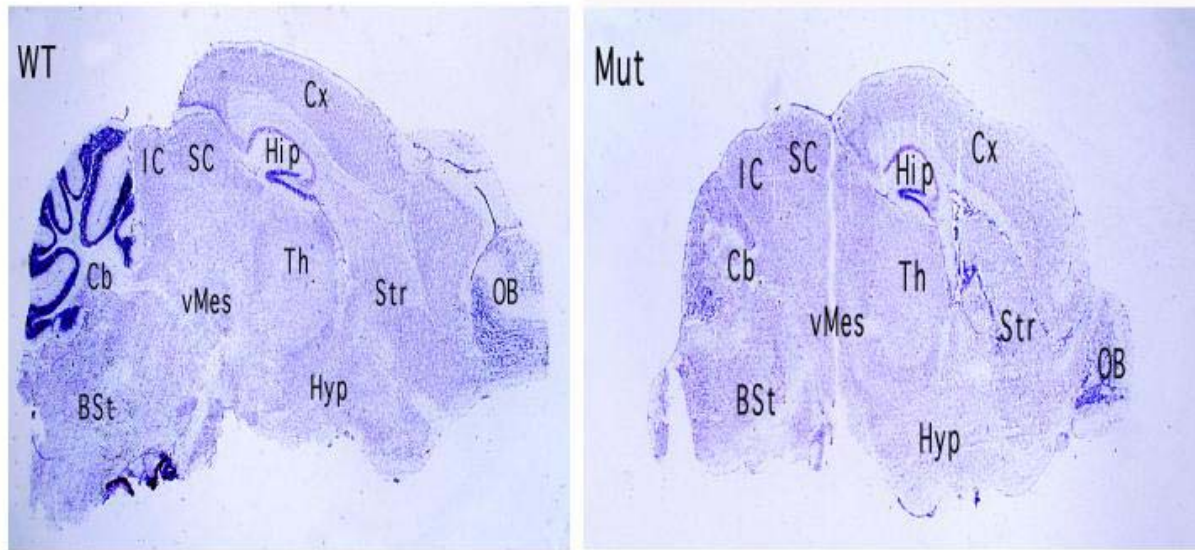


Figure 3.33: Nissl staining of mouse brain section. The brain of *nax* mutant mice is overall smaller compared to wild type, because the animals are smaller. All main morphological subdivisions are present, but that does not exclude small disturbances in the overall structure of brain. The most affected structure is the Cb. BSt=Brain stem, Cb=Cerebellum, Cx=Cortex, Hip=Hippocampus, Hyp=Hypothalamus, IC=Inferior Colliculi, OB=Olfactory bulb, SC= Superior colliculi, Str=Striatum, vMes=Ventral mesencephalon

3.10.6 Analysis of cerebellum sagittal section of developing *nax* mice

To understand the aetiology of the cerebellum abnormalities in *nax* mutant mice, we sought to determine the developmental stage at which the abnormalities became apparent. To study the cerebellar cortical structures, staining of sagittal sections from different stages of mice with respective markers e.g. for the granular layer and Purkinje cell layer were performed to pinpoint the exact phenotype.

The cerebellum is a highly ordered structure with a tight regulation of the ratio of Purkinje cells to granule cells. Neurons in the cerebellar cortex are organised into three layers. The outermost layer is called molecular layer and contains the stellate and basket cells. Beneath the molecular layer is the Purkinje cell layer (PCL) consisting of a single layer of Purkinje cell bodies. The innermost or the granule layer contains a vast number (estimated 10^{11}) of granule cells and a few larger Golgi interneurons.

When sagittal sections of adult *nax* mice cerebellum were stained with crystal violet, the cellular organisation of cerebellum appeared to be completely disorganised. Virtually the

complete inner granule layer (IGL) was lacking. The folia of cerebellum were also much smaller than wild type cerebellum (figure 3.34 A and B). When the sections were immunostained with Calbindin, which specifically stains Purkinje cell bodies, it was apparent that Purkinje cells also failed to form a single layer and were dispersed in a much broader layer. When immunostaining with glial filament acid protein (GFAP) was performed (which specifically stains Bergmann fibers), the Bergmann fibers also appeared to be in complete disarray (figure 3.34 C and D).

Nissl staining of sagittal sections of cerebellum of *nax* mouse from stage P20 suggested abnormalities in radial migration of granule cells from external granule layer (EGL) towards IGL. Calbindin staining showed that in stage P20 also Purkinje cells fails to form PCL (figure 3.35 A and B).

Nissl staining showed that newborn *nax* mice (P0) do have EGL though much thinner compare to wild type mice of the same stage. Purkinje cells still appeared to disperse thus suggesting that onset of phenotype is prenatal (figure 3.35 C and D).

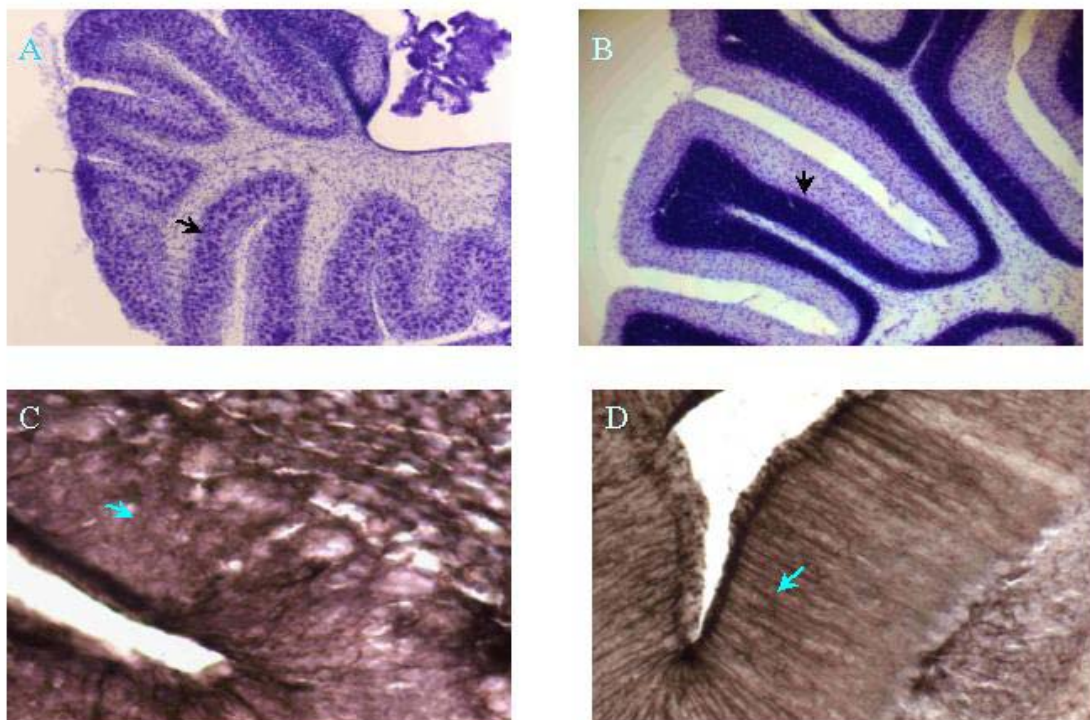


Figure 3.34: Sagittal sections of adult mouse cerebellum. (A) Nissl staining of *nax* mouse cerebellum, showed complete lack of inner granule layer (IGL). (B) Nissl staining of wild type cerebellum. (C) Immunostaining of *nax* cerebellum for glial filament acid protein (GFAP), Bergmann fiber appears to be disorganized as compared to wild type (D).

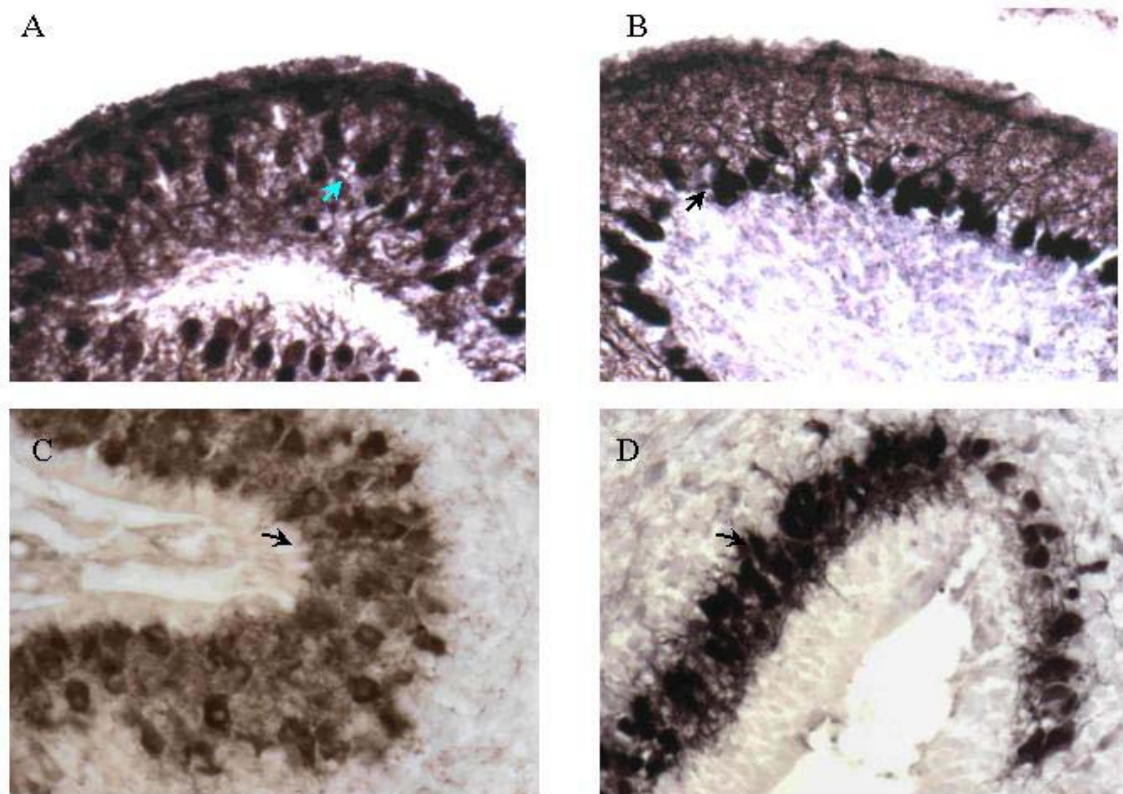


Figure 3.35: Sagittal sections of cerebellum from developing mice. (A) Calbindin staining of *nax* mouse at stage P20, Purkinje cells fail to form a single Purkinje cell layer (PCL), however in wild type mouse section (B) a thin PCL can be seen. (C) At stage P0, *nax* mouse have a disrupted PCL, but in case of wild type mouse at stage P0, PCL starts forming.

3.10.7 Linkage analysis

Linkage maps can only be constructed for loci that occur in two or more heritable forms, or alleles. Linkage maps are generated by counting the number of offspring that receive either parental or recombinant allele combinations from a parent that carries two different alleles at two or more loci. Analyses of this type of data allow one to determine whether loci are "linked" to each other and, if they are, their relative order and the relative distances that separates them.

3.10.7.1 Breeding strategy

The original *nax* mutation arose in embryonic stem cell DNA (129X1/SvJ). The mutation was introduced into C57BL/6J x 129X1/SvJ background. F1 animals were intercrossed to obtain F2 animals. Heterozygous F2 animals were intercrossed to obtain F3 animals. In a similar manner F4 animals were generated (a schematic diagram is shown in figure 2.1, page 50).

This breeding scheme was undertaken in order to maximise the number of meioses per animal in genome scan analysis, which enabled us to increase the resolution of genome scan.

3.10.7.2 Genome scan analysis

In order to identify the *nax* locus, a genome scan analysis was performed using microsatellite markers. For our initial analysis 48 animals were used, which included 42 affected animals and 6 control animals as shown in table 3.2. In total, this genome scan included 218 meioses and had a resolution of 15 cM. The coordinates of microsatellite markers were taken from Jackson Laboratory database (<http://www.informatics.jax.org>) (figure 3.36). The primers for microsatellite markers were conjugated with 5' end flurochrome and PCR amplification on various mice DNA was performed as described in section 2.2.8.3. The genotyping for each microsatellite marker for all animals was done by using GENOTYPER software and allelic frequency was scored. After initial genome scan analysis most of the microsatellite markers showed an even allelic frequency for both 129X1/SvJ and C57BL/6J parents, however marker D2mit206 showed an allelic frequency of 0.95 for 129X1/SvJ genotype and 0.05 for C57BL/6J. The coordinate of marker D2mit206 corresponds to chromosome 2 (51.6 cM), which shows a very strong linkage to *nax* locus (table 3.3).

	1	2	3	4	5	6
A	129/Sv	C57BL	ES cell	F1-1	F1-2	F1-3
B	F2-1	F2-2	F2-3	F2-4	F2-5	F2-6
C	F2-7	F2-8	F2-9	F2-10	F2-11	F2-12
D	F2-13	F2-14	F2-15	F2-16	F2-17	F3-1
E	F3-2	F3-3	F3-4	F3-5	F3-6	F3-7
F	F3-8	F3-9	F3-10	F3-11	F3-12	F3-13
G	F3-14	F3-15	F3-16	F3-17	F3-18	F3-19
H	F3-20	F3-21	F3-22	F3-23	F3-24	F3-25

Table 3.2: Genomic DNA array used for initial genome scan analysis. The identity of each animal used in this analysis is shown with a name, for example: F3-15 means the animal was from third filial generation and was fifteenth animal in this generation. After normalization of DNA concentration, 5 ng of each sample DNA was dispensed into 96 well PCR plates and above table is a pictorial representation of this 96 plate. 129/Sv, 129X1/SvJ; C57BL, C57BL/6J.

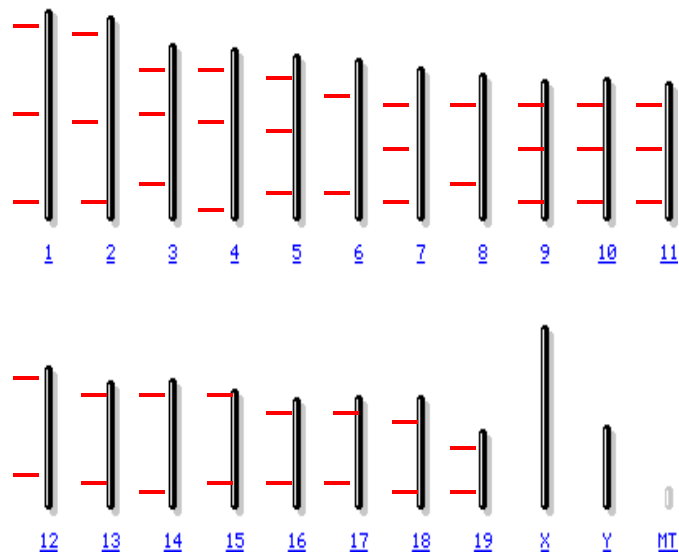


Figure 3.36: Genome scan analysis. The coordinates of the microsatellite markers were taken from Jackson Laboratory's database. Each vertical bar represents mouse chromosome and the horizontal bars are the coordinates of microsatellite marker. The resolutions of this genome scan was 15 cM.

3.10.7.3 Haplotype analysis

To narrow down the *nax* locus, we undertook haplotype analysis for chromosome 2 with a set of microsatellite markers present in chromosome 2 as listed in figure 3.37. We selected 32 animals based upon critical recombination events in chromosome 2 in their genome. After scoring for a series of markers placed on chromosome 2, *nax* locus could be localised within a genetic distance of 2 cM between marker D2mit158 and D2mit100. *nax* locus was located between marker D2mit158 (45.5 cM) and marker D2mit100 (47.5 cM). In our haplotype analysis one animal showed critical recombination in one arm of chromosome 2 from proximal position at D2mit158 marker and three animals where recombinant from distal position at marker D2mit100. The result of complete haplotype analysis is summarised in figure 3.37 and in figure 3.38.

Chr	Marker	Label	L (cM)	129X1/SvJ	C57BL/6J
1	D1mit211	FAM	12	-	-
1	D1mit216	HEX	49	.49	.51
1	D1mit206	TET	94	.55	.45
2	D2mit117	TET		.71	.29
2	D2mit206	FAM	51	.95	.05
2	D2mit148	HEX	92	.39	.61
3	D3mit310	FAM		-	-
3	D3mit278	TET	25	.70	.30
3	D3mit87	FAM	63	-	-
4	D4mit211	FAM	5	.42	.58
4	D4mit178	TET	59	-	-
4	D4mit310	HEX	71	.67	.33
5	D5mit388	FAM	14	.42	.58
5	D5mit25	FAM	49	-	-
6	D6mit14	HEX	63	.48	.52
6	D6mit102	HEX	12	.48	.52
7	D7mit145	TET	24	.43	.57
8	D8mit121	TET	35	-	-
9	D9mit269	TET	41	-	-
10	D10mit183	FAM	12	-	-
10	D10mit266	TET	63	.45	.55
11	D11mit36	TET	43	-	-
11	D11mit224	TET	70	.57	.43
12	D12mit263	HEX	56	.31	.69
13	D13mit14	FAM	15	-	-
13	D13mit67	TET	24	-	-
14	D14mit97	HEX	67	.49	.51
14	D14mit160	HEX	47	.38	.62
14	D14mit109	FAM	3	.51	.49
15	D15mit171	HEX	52	.45	.55
16	D16mit189	FAM	55	.42	.58
17	D17mit133	FAM	10	.54	.46
18	D18mit60	FAM	16	.46	.54
19	D19mit6	FAM	4	.44	.56
19	D19mit8	TET	34	.48	.52
19	D19mit111	FAM	55	.5	.5

Table 3.3: Allelic frequencies of microsatellite markers used in genome scan analysis. Microsatellite marker D2mit206, whose coordinate is chromosome 2 (51.6 cM) scored an allelic frequency of 0.95 for 129X1/SvJ genotype thus suggesting a very strong linkage for *nax* locus. For some microsatellite markers, allelic frequency could not be scored because the markers were not informative; they are indicated as hyphen (-) in allelic frequency section of the corresponding marker. Chr, chromosome; Label, fluorochrome conjugated label to marker; L, locus.

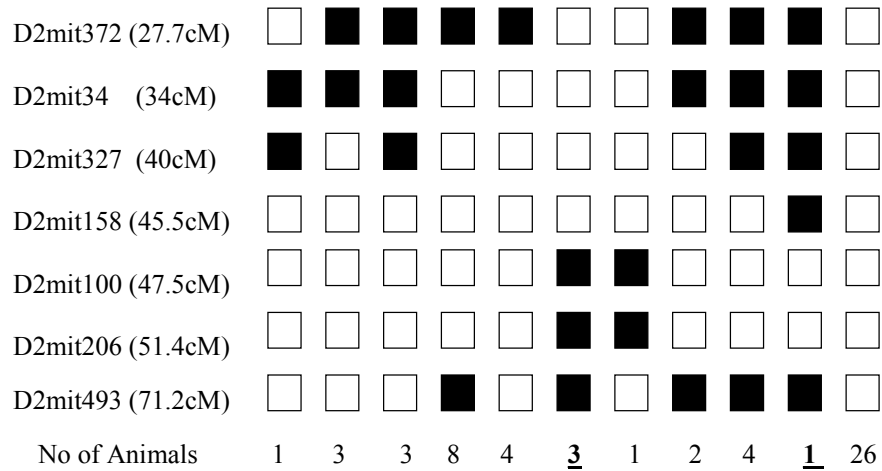


Figure 3.37: Haplotype analysis. Microsatellite markers with their respective coordinate which were used in haplotype analysis are listed in first column. The white block represents the genotype of 129X1/SvJ strain and the black box represent a recombinant genotype, which means one of the sister chromatids had undergone recombination and represents a genotype of C57BL/6J strain. In summary, one animal haplotype suggest a critical recombination from proximal position of chromosome 2 at position 45.5 cM (D2mit158) and three animals had a critical recombination at position 47.5 (D2mit100) from distal position.

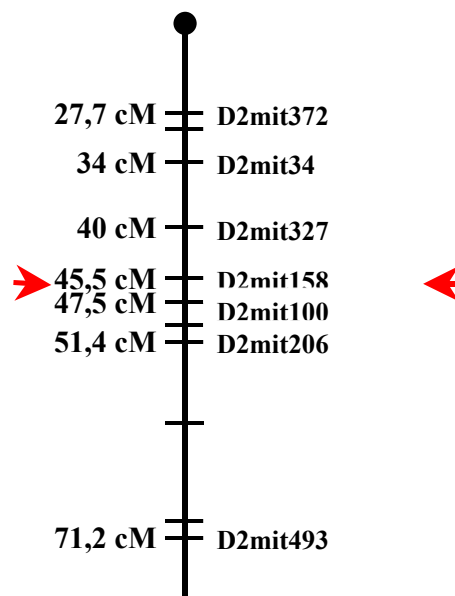


Figure 3.38: Schematic representation of chromosome 2. The *nax* locus (indicated by two red arrows) was placed between marker D2mit158 (45.5 cM) and marker D2mit100 (47.5 cM) by linkage analysis.

3.10.7.4 Characterisation of *nax* locus

The genomic sequence between marker D2mit158 and D2mit 100 was downloaded from Celera database (www.celera.com). To our surprise the physical distance between these two markers was more than 14 Mb. When the coordinates of above markers were checked in Whitehead Institute of Biomedical Research database, we discovered that the coordinates for above marker were different so D2mit158 was placed at 44.8 cM instead of 45.5 cM and D2mit100 was placed at 50.3 cM instead of 47.5 cM. Therefore actual genetic distance of characterised *nax* locus was 5.5 cM and coordinates for above markers in Jackson Laboratory database was incorrect. The *nax* locus spanned into five different scaffolds of Celera Discovery System's (CDS) mouse genome as shown in figure 3.39. We checked few strong candidate genes present in this regions for presence of *nax* mutation by using Northern blot, RT-PCR, genomic Southern and genomic sequencing, which included 1) *Neurod1*, a helix-loop-helix transcription factor, a knock-out of this gene in mouse had a similar phenotype like *nax* mouse. 2) γ -*catenin*, a putative member of WNT signalling pathway. However no mutation can be observed in above genes in *nax* mouse.

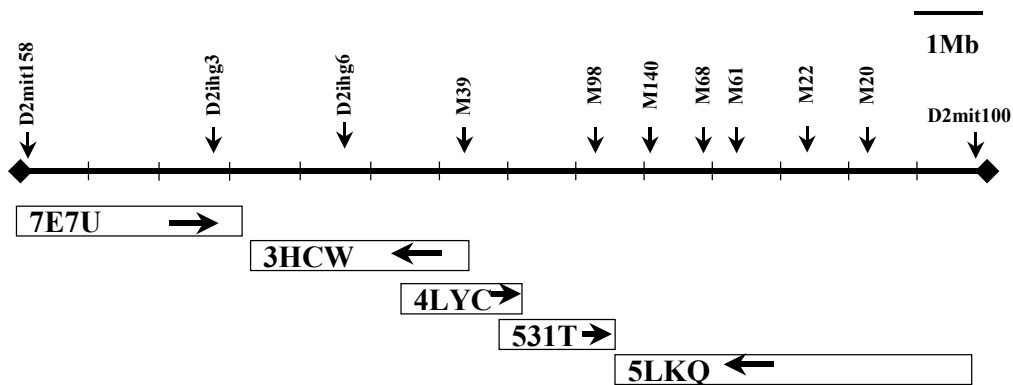


Figure 3.39: Schematic representation of *nax* locus, showing the relative position of five scaffold identified from Celera database. The *nax* locus (between marker D2mit158 and D2mit100) span about 14 Mb length of genomic region of chromosome 2. D2ihg3, D2ihg6, M39, M98, M140, M68, M61, M22, and M20 are the putative microsatellite markers identified from the genomic sequence downloaded from Celera database.

3.10.7.5 Identification of putative polymorphic markers

As the region characterised for *nax* locus was very large containing more than 200 genes, we decided to narrow down the *nax* locus further. No polymorphic markers from any public database were available which lie within this region therefore we searched for putative di, tri, tetra, and penta repetitive sequence within this locus with tandem repeat finder software (<http://tandem.biomath.mssm.edu/trf/trf.html>). Primers were designed for PCR amplification of each putative microsatellite marker and PCR reactions were performed on genomic DNA from 129X1/SvJ and C57BL/6J strain. The PCR products were separated on 10 % PAGE gel in order to identify any polymorphism. Nine polymorphic markers placed at regular interval within this 14 Mb region were identified after screening more than 400 putative markers for polymorphism.

3.10.7.6 New haplotype analysis

After identification of new polymorphic markers, we initiated a new haplotype analysis on chromosome 2 to narrow down the *nax* locus in our analysis. We selected 45 new animals mainly from F4 generation and included three parental controls. All the animals in this new array were genotyped for 9 new microsatellite markers and based on these results, haplotypes for each animal was generated. After the completion of new haplotype analysis, *nax* locus was placed between marker M98 (88.8 Mb) and marker M140 (89.8 Mb), therefore *nax* locus was narrowed down to 0.8 Mb. In our new haplotype analysis one animal showed critical recombination in one arm of chromosome 2 from proximal position at M98 marker and one animal was recombinant from distal position at marker M140. The result of complete new haplotype analysis is summarised in figure 3.40.

D2mit158	■	■	■	□	□	□	□
D2ihg3	□	■	■	□	□	□	□
D2ihg6	□	■	■	□	□	□	□
M39	□	■	■	□	□	□	□
M98	□	□	■	□	□	□	□
M140	□	□	□	■	□	□	□
M68	□	□	□	■	□	□	□
M61	□	□	□	■	■	□	□
M22	□	□	□	■	■	□	□
M20	□	□	□	■	■	□	□
D2ihg9	□	□	□	■	■	□	□
D2mit100	□	□	□	■	■	■	□
No of Animals	2	2	<u>1</u>	<u>1</u>	1	3	35

Figure 3.40: New haplotype analysis. Microsatellite markers which were used in haplotype analysis are listed in first column. The white box represents the genotype of 129X1/SvJ strain and the black box represent a recombinant genotype, which means one of the sister chromatids had undergone recombination and represents a genotype of C57BL/6J strain. In summary one animal haplotype suggests a critical recombination from proximal position at marker M98 and another animal had a critical recombination at marker M140 from distal position. The distance between these two markers is 800 kb.

3.10.7.7 Identification of gene responsible for *nax* phenotype

The *nax* locus was identified between markers M98 and M140, the physical distance is 800 kb and the locus is mapped in the middle of chromosome 2 between 88.8 Mb and 89.6 Mb. In Celera database, within this region, 24 genes are reported. The next step was to analyse all the 24 genes and try to identify the mutation, which is responsible for *nax* phenotype in mice. We decided to screen for the *nax* mutation by three different approaches namely Northern blot analysis, which will detect any mutation which might result in lack of transcript or a truncated form of the transcript. RT-PCR will detect all those transcripts, which are weakly expressed and cannot be detected by Northern blot. Genomic sequencing will enable us to identify any mutation in the gene such as missense, nonsense, and frame-shift, which might lead to a non-functional protein. However till now no mutation was identified which can be attributed to *nax* phenotype. The results of mutational analysis are summarised in table 3.4.

GENE	Accession No.	Northern	RT-PCR	Sequencing
RabGEF	AAD12154.1	+	+	+
LIV-1 related	AAH19016	+	+	+
Dual phosphatase	BC014048	+	+	+
CUG binding protein	AAF78957.1	+	+	+
Nr1h3	AF085745	+	+	+
Pacsin3	AAH03884	+	+	+
Zfp289	AAH05112		+	
Fnbp4	AY059772	+	+	
Kelch related	AAA53472.2	+	+	+
0710001e10Rik	AAK92202.1		+	+
Rapsn	AAF34339		+	
Psmc3	AAH05783		+	
Sfpi	AAf78905.1	+	+	
Mybpc3	AAC04620.1		+	
2610043A19Rik	AAK16810.1		+	
NADH Ubiquinone	AAD40386.1	+	+	
Mitch2	AAH 02152			
L30 ribosomal				
mCG1040630	AF37515.1		+	
mCG134155				
P1 ribosomal	AAG13292.1			
mCG13490	AAH07460			
mCG1040633				
mCG13507	AAH13988	+	+	

Table 3.4: Summary of analysis of genes reported in *nax* locus. In column one, names of the unknown genes are listed and in case the name of gene is unknown than Celera identification number is given. Column two refers to corresponding accession number of the gene as reported in public database. The last three columns summarise the result of our analysis, a plus (+) sign in a box corresponding to a gene means analysis was complete and no mutation was found in that gene and a empty box refers to either incomplete analysis or a failed analysis.

4. Discussion

Sertoli cells greatly influence mammalian germ cells development at all stages of spermatogenesis. Sertoli cells generate the unique chemical and physical microenvironment in the adluminal compartment, which is required for spermiogenesis to proceed: Sertoli cells play nutritive and regulatory roles for development of all stages of germ cells. Furthermore, in vitro studies have shown that Sertoli cells secrete several proteins that exert their effect on the regulation of specific genes in germ cells. For example, cultured Sertoli cells express insulin-like growth factor II, and this growth factor is capable of altering the expression of *c-fos* in spermatogenic cells isolated from murine testes (Tsurala and O'Brien, 1995). The co-culture experiments have allowed identification of wide array of genes in germ cells, which are regulated by secreted factors from Sertoli cells (Skinner, 1991; Syed and Hecht, 1997; Lilienbaum et al., 2000; Vidal et al., 2001).

In first part of this thesis, functional characterisation of a novel gene named *Theg*, which is specifically expressed in spermatids, is discussed. *Theg* expression in spermatid cells is up regulated by some unknown factor/s from Sertoli cells (Nayernia et al., 1999). In latter section, human homologue of murine *Theg* is described and in the final section, a spontaneous mutation named *nax*, which arose during the generation of *Theg* knock-out mice is characterised.

4.1 Functional characterisation of *Theg* and its role in spermatogenesis

4.1.1 Expression analysis, cellular distribution and subcellular localisation of the *Theg* protein

Theg is a novel protein as no homologous protein and/or conserved motif was detected in Genbank database (<http://www.ncbi.nlm.nih.gov/>). When Western blot was performed with antibody generated against *Theg* peptides, specific immunostaining was observed only in testicular protein extract. *Theg* protein corresponded to a molecular weight of 42 kDa (figure 3.4), which was in agreement with the predicted molecular weight as deduced from its amino acid sequence. This suggests that during intracellular protein processing *Theg* does not undergo major post-translation modification, this is in contradiction to the deduced sequence analysis, where several putative sites for phosphorylation and glycosylation were detected (Nayernia et al., 1999). A notable point in Western blot analysis was that only a single protein band was observed, however from our previous studies and other studies, four different

spliced isoforms of *Theg* transcript were reported (Nayernia et al., 1999; Yanaka et al., 2000 and present study), in two out of these four isoforms only 24 amino acids are lacking and such difference in molecular weight cannot be detected in the condition applied in our Western blot analysis. In an earlier study, Nayernia et al. (1999) reported an isoform which could result into a truncated *Theg* protein of 20 kDa, perhaps this isoform is present in very low concentration in testicular protein extract, which is beyond the threshold level of Western blot detection, therefore we could not detect this truncated *Theg* protein in our Western blot analysis.

When the expression pattern of *Theg* protein was analysed in the testicular protein extract from different developing stages of mouse using Western blot, a weak immunostaining can be detected on testicular protein extract from stage P23 and latter stages of developing mice. The expression of *Theg* after stage P35 is comparable to that of adult mouse (figure 3.5). This result is in agreement with that of Northern blot analysis, where *Theg* expression was first detected in testicular RNA at stage P20 of mouse development, when haploid germ cells (round spermatids) start appearing in testis (Nayernia et al., 1999). The fact that no protein band was observed in testicular protein from stage P20 in the Western blot can be explained by presence of very low level of *Theg* protein, which is beyond the level of sensitivity of Western blot detection. From this finding, it cannot be concluded that there is some sort of translation delay in *Theg* protein synthesis, therefore any post-transcriptional regulation cannot be postulated. The notion that *Theg* is expressed in haploid germ cells was further consolidated by the result that *Theg* expression can only be detected in mouse mutants with haploid germ cells (*olt/olt* and *qk/qk*) but not in mouse mutants with either no germ cells (*W/Wⁿ*) or primary spermatocytes (*Tfm/y*) (figure 3.6).

The cellular distribution of *Theg* protein in testis was determined by immunofluorescence staining of mouse testis sections. Specific staining was only observed in round spermatids but neither in diploid germ cells (spermatogonias and spermatocytes) nor in maturing spermatozoas (figure 3.7 A-F). This highly restrictive expression of *Theg* in a narrow framework window of spermatogenesis particularly in the beginning of spermiogenesis, suggests that *Theg* might play a very important function in germ cell maturation. *Theg* might be one of the few genes, which are expressed only in round spermatid cells.

In the deduced amino acid sequence of *Theg* protein two putative nuclear localisation signals (NLS) were detected, which indicates that *Theg* encodes for a nuclear protein (figure 3.3). In order to determine the subcellular localisation of *Theg* protein two different experiments were performed. In first experiment, immunocytochemical staining with *Theg* antibody was carried

out on cellular suspensions isolated from adult mouse testes, specific immunostaining of *Theg* was predominately detected in the nucleus of round spermatids, whereas no specific staining was observed in other germ cell stages (figure 3.7 G-I). In a second approach, *Theg* protein expression was determined by generated Gfp-*Theg* fusion protein in NIH3T3 cell line, the fusion protein was mainly expressed in nucleus of transfected cells. This clearly demonstrates that *Theg* is localised predominately in the nucleus of round spermatids in mouse (figure 3.7 J-L). Other examples of nuclear proteins expressed during haploid stages of spermatogenesis are transition proteins (Tnp1 and Tnp2) (Kistler et al., 1996; Baskaran and Rao, 1991), protamines (Prm1, Prm2 and Prm3) (Balhorn et al., 1984; Schlüter and Engel, 1995; Schlüter et al., 1996) and ADAM family proteins (Wolfsberg et al., 1995a; Wolfsberg et al., 1995b). Different functions have been suggested for the nuclear proteins expressed in haploid germ cells, including nuclear shaping, histone removal, transcriptional repression, chromatin condensation, and, most recently, repair of the DNA strand breaks that normally transiently occur during the removal of the nucleosomes. (Zhao et al., 2001; Caron et al., 2001; Kierszenbaum, 2001)

4.1.2 Generation of *Theg* deficient mice to characterise its role in spermatogenesis

The temporal-spatial nuclear expression of *Theg* in round spermatid cells suggests that *Theg* might play an important role during a successful differentiation of the male germ cells. To elucidate the function of *Theg* in mice, a targeted mutation of the gene was generated deleting most of the C-terminal of the protein. We deleted a genomic locus of *Theg* that included exon 4 to exon 7 of the gene, encompassing important nuclear localisation signal sequences by homologous recombination in ES cells (figure 3.10 and 3.13). The heterozygous animals appeared phenotypically normal and were fertile, therefore haploinsufficiency of *Theg* allele have no effect in mice. The homozygous mice for *Theg* deleted allele were generated in two different genetic backgrounds namely, in inbred strain 129X1/SvJ and in hybrid strain C57BL/6J x 129X1/SvJ.

In our *Theg* deleted knock-out mice, we don't expect a *Theg* transcript to exist, as the truncated transcript which will result in the knock-out mice will be unstable under physiological condition due to lack of 3' untranslated region of the transcript. When Northern blot was performed using a *Theg* cDNA consisting of exons 3, 4, 5, 6, and 7, no signal was detected in testicular RNA from *Theg*^{-/-}, in heterozygous mice the intensity of signal was weaker than that of testicular RNA of wild type mice (figure 3.15A). However when full length *Theg* was used as a probe in Northern blot analysis, we observed a signal in testicular

RNA from *Theg*^{-/-} which was weaker compared to that of wild type testicular RNA (figure 3.15B). Further in Northern blot analysis with *Neomycin* as probe; we detected two bands in *Theg*^{+/-} and in *Theg*^{-/-} testicular RNA but no signal was detected in testicular RNA derived from wild type mice (figure 3.15D). This observation can be explained by the fact that during the generation of knock-out mice, a *Neomycin* cDNA under the control of *Phosphoglycerate kinase* promoter was introduced into mouse genome by homologous recombination, therefore one band corresponds to the endogenous transcript, the second transcript might be as a result of fusion of *Theg* truncated transcript with *Neomycin* cassette. In order to determine whether indeed a fusion transcript exists in *Theg*^{-/-} mice we performed RT-PCR with two different pairs of primers namely mTHEG1/NeoRI and mTHEG1/mTHEG2 on testicular RNA derived from *Theg*^{-/-} mice. A PCR product was amplified with mTHEG1/NeoRI, but no amplification product was observed with mTHEG1/mTHEG2 (figure 3.16). After sequence analysis of this amplified product we could confirm that indeed a fusion transcript exists in *Theg*^{-/-} mice. This fusion transcript is comprised of exon 1 and exon 2 of *Theg* fused with *Neomycin* transcript. The reason that this fusion transcript was generated in *Theg*^{-/-} mice could be because of inherent property of *Theg* to undergo alternative splicing. Under normal circumstances *Theg* give rise to four different splice variant (as discussed in 4.1.1), in one instance exon 3 is spliced out resulting in a *Theg* isoform, which only lacks exon 3. We believe that in *Theg*^{-/-} mice perhaps near *Neomycin* cassette there is a cryptic splice acceptor site, which generated a splice fusion *Theg* isoform (figure 4.1) and since *Neomycin* cassette included polyadenylation sequence, the fusion transcript generated is stable in vivo.

The next obvious question is what is the relevance of this fusion transcript in *Theg*^{-/-} mice. As wild type *Theg* protein is predominately localised in the nucleus of round spermatids, therefore we sought out the intracellular localisation of fusion *Theg* protein. Immunocytochemical staining of cellular suspension isolated from *Theg*^{-/-} testes showed a cytosolic localisation of *Theg* fusion protein (figure 3.20 C-F). This result was furthered supported by generated Gfp tagged *Theg* fusion protein, which is expressed only in cytoplasm of transfected NIH3T3 cells (figure 3.20 J-L). These result clearly shows that generated *Theg* fusion protein in *Theg*^{-/-} mice have lost its ability to localise predominately in the nucleus of round spermatid cells. However a rational question at this juncture, which still remains unanswered is what functional role (if any) this fusion protein play in *Theg*^{-/-} mice.

Phenotypically, homozygous mice from both genetic backgrounds (in inbred strain 129X1/SvJ and in hybrid strain 129X1/SvJ x C57BL/6J) appeared normal and were fertile.

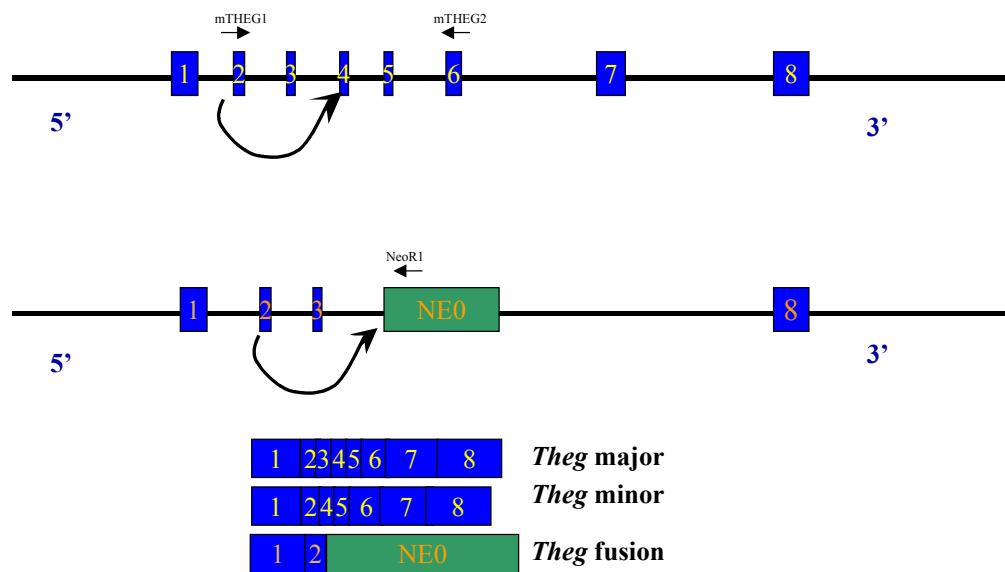


Figure 4.1: Schematic diagram showing alternative splicing in *Theg* gene, which in case of wild type allele give rise to *Theg major* and *Theg minor* transcript. However in case of mutant allele, alternative splicing in *Theg* gene give rise to *Theg fusion* transcript. mTHEG1, mTHEG2, and NeoR1 were the primers used for RT-PCR.

When a phenotype-genotype correlation was made, we discovered that in genetic background C57BL/6J x 129X1/SvJ the *Theg* deleted allele showed deviation from Mendelian mode of inheritance (figure 3.17). The ratio of heterozygous and homozygous mice are reduced as compared to that of wild type mice, suggesting that during segregation of alleles, *Theg* deleted allele is disadvantageous to wild type allele. However the mode of inheritance in *Theg* deleted allele in genetic background 129X1/SvJ did not show any deviation from Mendelian inheritance. An inferior *Theg* deleted allele in genetic background C57BL/6J x 129X1/SvJ could be due to genetic heterogeneity of the mice. There are numerous reports, which showed that genetic heterogeneity causes difference in phenotype in knock-out mice. For example Nayernia et al. (2002) showed that *Smcp*^{-/-} (sperm mitochondria-associated cysteine-rich protein) mice in mixed genetic background (C57BL/6J x 129X1/SvJ) are fully fertile, while they are infertile in the 129X1/SvJ background, in an another report Adham et al. (2001) showed that *Tnp2*^{-/-} (transition protein2) male mice were normal and fertile in mixed genetic background C57BL/6J x 129X1/SvJ, however they were totally infertile in the inbred 129X1/SvJ background. However a point worth mentioning is, in both examples discussed

above, the mixed genetic background of C57BL/6J x 129X1/SvJ serves as an advantage to the deleted allele but in contrary in our case, genetic heterogeneity resulted in minor inferiority to *Theg*-deleted allele. It is very difficult to speculate a reason when interactions of nature of such complex genetic heterogeneity is involved.

In another observation we noticed that testes of *Theg*^{-/-} mice appeared smaller than the testes of wild type (littermates) mice. When the growth curve of the testes development was determined, we observed a minor but significant retardation in case of *Theg*^{-/-} mice (figure 3.18). However histological examination of the testis revealed that in *Theg*^{-/-} mice the occurrence of spermatogenesis was normal and also the number of spermatozoa present in the epididymis was comparable to that of wild type mice. The spermatozoa from *Theg*^{-/-} mice were also able to fertilise oocytes in a normal fashion as determined by in vitro fertilisation (IVF) assay. The morphological appearance of *Theg* deficient sperm under the microscope also appeared to be normal.

We were also unable to detect any significant defect in motility of *Theg*^{-/-} sperm, as determined by computer assisted semen analysis (CASA) system (CEROS version 10, Hamilton Throne Research) (figure 3.19). So from the results of *Theg* deleted knock-out mice, we can conclude that deleted *Theg* is a non-essential protein in mice and spermatogenesis progress to its completion in *Theg* knock-out mice and produces normal motile spermatozoa. Although we observed some minor defects as discussed above however those abnormalities are subtle and does not effect *Theg*⁻ sperms ability to fertilise oocytes.

4.1.3 An insertional mutation in *Theg* locus caused a defect in spermatogenesis

During the course of time when *Theg* deletion knock-out mice were generated, Yanaka et al. (2000) reported a transgenic mouse, where a foreign DNA (human PDE5A gene under the control of cytomegalovirus promoter) was integrated into *Theg* locus. In their mice, several copies of the transgene were integrated into *Theg* locus thus producing a deletion of about 20 kb in the genomic locus. This insertion deleted the complete open reading frame (ORF) of *Theg* gene and also deleted an additional 7 kb fragment upstream of *Theg* gene (figure 4.2).

Yanaka et al. (2000) named this deleted transgenic locus as *Kisimo* (*ki*). The male mice homozygous for *ki* locus were sterile and sterility arose from a defect in spermatogenesis. They also reported that in *ki/ki* mice, there were virtually no spermatozoa in the lumina of seminiferous and epididymal tubules. On histological analysis of cross-sections of *ki/ki* testes, they discovered that elongated spermatids in the vicinity of the lumina of seminiferous tubules showed vacuolation and were occasionally phagocytosed by Sertoli cells (figure 4.3A&B).

Further characterisation of spermatids by electron microscopy revealed that the elongated spermatids have abnormal or completely non-existent flagella (figure 4.3C&D).

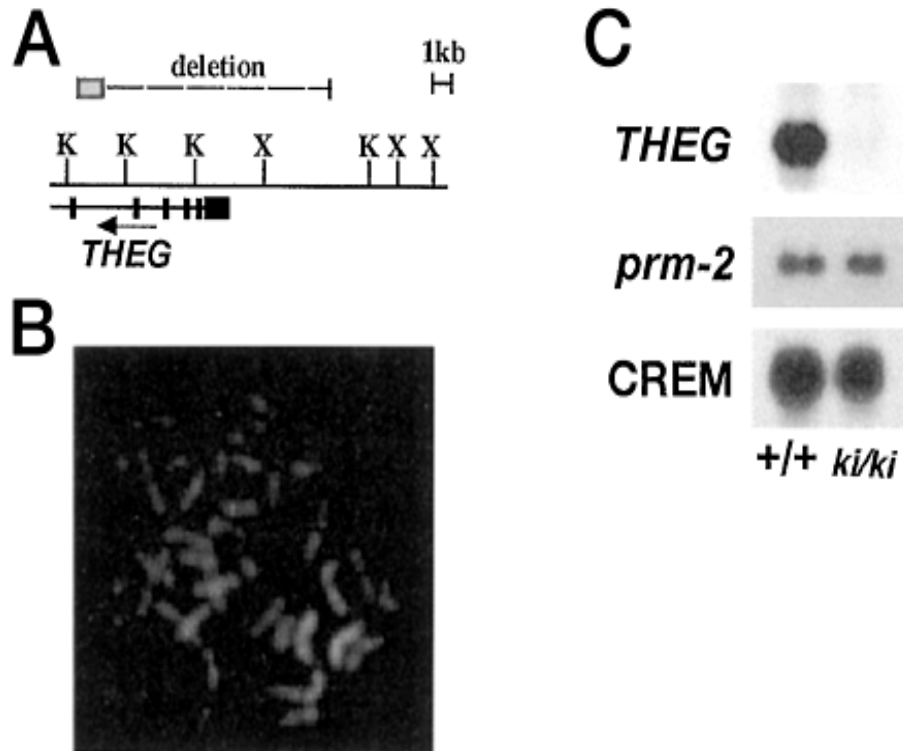


Figure 4.2: Identification of the mouse *ki* locus. A) Physical map of the *ki* locus. The direction of *Theg* transcription is indicated by an arrow. The genomic DNA flanking the transgene insertion site obtained by TAIL-PCR is shown by a bar. The hatched box represents a DNA fragment used as a probe for Northern blot analysis. Restriction enzymes are indicated as follows: K, *KpnI*; X, *XhoI*. B) Fluorescence in situ hybridization. The transgene insertion locus was determined to be localised in chromosome 10. C, Northern blot analysis probed with the *KpnI/XhoI* 5 kb genomic fragment from *ki* locus (*top panel*) or with testis-specific probes corresponding to protamine 2 (*prm2*) or CREM cDNA (adapted from Yanaka et al., 2000).

In addition, in the quantitative evaluation of spermatogenic cells in seminiferous tubules, elongated spermatids were significantly decreased at all stages of spermatogenesis (figure 4.3E). However, the number of spermatogonia cells and spermatocytes cells were comparable to that of wild type mice. Yanaka et al. (2000) postulate that in elongated spermatids *Theg* interacts with a cytoplasmic protein termed CCT ϵ . The CCT ϵ is one of the components of TCP1 (t-complex polypeptide), which is involved in microtubule assembly and required for the proper folding or assembly of tubulin proteins (Kubota et al., 1995; Silver et al., 1987). They suggest that abnormal or absent flagella in *ki/ki* mice is perhaps due to impairment of

the assembly of cytoskeletal proteins such as the tubulins that are major structural proteins of sperm tail flagella (Philips et al., 1993).

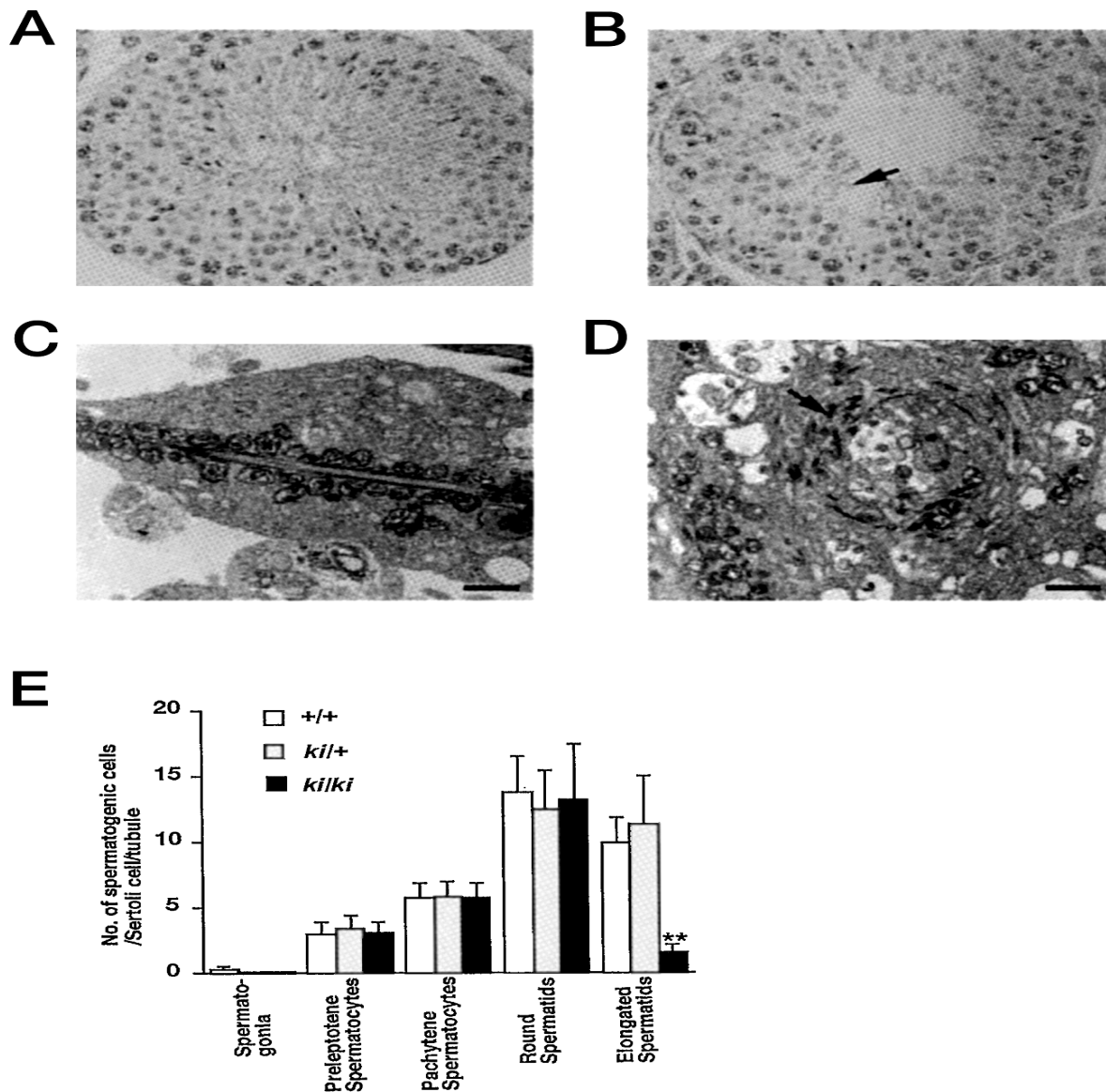


Figure 4.3: Spermatogenesis abnormality in *ki/ki* mice. Cross-sections of testes (stage VII) from wild-type *A* and *ki/ki* mice *B* are shown. Staining with hematoxylin and eosin revealed the abnormal elongate spermatids. The arrowhead indicates the increased intracytoplasmic vacuoles in elongated spermatids in *ki/ki* mice. Electron microscopy of spermatids from wild type *C* and *ki/ki* mice *D* are shown. In *ki/ki* mice, dense fibers are shown to form a circular array (arrowhead). *E*, quantitative evaluation of spermatogenic cells in seminiferous tubules (stage VII). The ratios of each type of spermatogenic cell to Sertoli cell per seminiferous tubule cross-section were calculated, and the obtained values were analyzed using a Dunnett-type multiple comparison test, with p , 0.01 (**) as minimum significance (adapted from Yanaka et al., 2000).

In contrast to the above report, our results clearly demonstrate that *Theg* is localised predominately in the nucleus of round spermatid, therefore further investigation such as generation of a new knock-out mice where N-terminal of *Theg* is deleted, was necessary to define the exact role of *Theg* in spermatogenesis.

One possible reason for difference in phenotype in our *Theg* knock-out mice and Yanaka et al. (2000) reported transgenic mice could be due to difference in genetic background of the mice. We generated *Theg* knock-out mice in mixed strain C57BL/6J x 129X1/SvJ and in inbred strain 129X1/SvJ, Yanaka and coworkers introduced the transgene in C3H/J inbred mice. In order to determine whether it is indeed a genetic factor which influences the phenotype of *Theg* deletion, we backcrossed our knock-out mice from strain 129X1/SvJ to strain C3H/J for seven generations to reach a 97 % of genomic homogeneity for C3H/J strain, thus creating a *Theg*^{-/-} congenic strain in C3H/J genetic background. The male *Theg*^{-/-} congenic mice were fertile and we did not observe any defect in spermatogenesis. The litter sizes of these mice were also comparable to that of wild type C3H/J mice (table 3.1). Therefore we can conclude that genetic background does not influence the phenotype of *Theg* deletion in mice.

4.1.4 Generation of a new *Theg* knock-out mice

From the results of our *Theg* knock-out mice and Yanaka and co worker's transgenic mice, one can speculate that N-terminal of *Theg* (exon 1 and exon 2), which constitutes a domain of 136 amino acids, is essential for in vivo function of *Theg*. In order to address this question we generated another *Theg* knock-out mice where we deleted exons 1, 2, 3 and exon 4 of *Theg* gene and replaced it with *Neomycin* cassette by homologous recombination in ES cells (figure 3.21 and 3.23). In our new knock-out mice we made one additional modification, we placed *Neomycin* transcription in a reverse manner to that of in vivo *Theg* transcription, thus ruling out any possibility of generating any sort of fusion transcript, which can be envisaged (figure 3.21). The heterozygous mice in both C57BL/6J x 129X1/SvJ and 129X1/SvJ background were phenotypically normal and fertile, which was as expected. The mice in C57BL/6J x 129X1/SvJ showed no deviation from Mendelian inheritance and the homozygous mice were fertile. The breeding of these mice is in progress and detailed phenotypic analysis can be performed in some latter point. We still have no data for homozygous mice from inbred 129X1/Sv, as mice are too young to test for fertility. There are many reports, which show the knock-out phenotype is more severe in inbred strain 129X1/SvJ as compared to mixed strain C57BL/6J x 129X1/SvJ (Nayernia et al., 2002; Adham et al., 2001). Perhaps we might be able

to see some severe spermatogenesis phenotype in 129X1/SvJ background. Nevertheless if we indeed observe any phenotype in 129X1/SvJ background, it will be due to some genetic modifier present in 129X1/SvJ genome.

4.1.5 Conclusion

We, with our first and second *Theg* knock-out mouse model systems attempted to genetically dissect the functional domain of *Theg* protein. From our results we can conclude that *Theg* is a non-essential protein for spermatogenesis and *Theg* protein function was perhaps compensated by some other redundant protein/s.

Gametogenesis is undoubtedly the most critical process in an organism therefore to ensure that a successful progression of spermatogenesis undergoes to its completion; there are numerous genes, which participate in redundant manner in testis. For example *transition protein2 (Tnp2)* (Zhao et al., 2001; Adham et al., 2001), testis-specific *histone1 protein (H1t)* (Fantz et al., 2001) and many others which are not well characterised (Baba et al., 1994; Adham et al., 1997), *Theg* might be one such gene. To determine, which gene is compensating for *Theg* function can be scope of future endeavour.

Now to address the ambiguity between our result and that of Yanaka et al. (2000), they generated a transgenic mice where more than 20 kb region near *Theg* locus was deleted, therefore it is very much possible the effect can be from any other gene which is near the proximity of this locus. We did a sequence analysis of the reported deleted region, no coding sequence of any other gene was reported in this region, however one cannot rule out a cis-acting effector sequence, which might influence any distance gene. The best way to address that question will be a detailed analysis of *ki/ki* mouse. For example, a comprehensive expression analysis of genes involved in spermatogenesis in *ki/ki* mouse to determine any other gene/s whose expression is altered in *ki/ki* mouse. Another possibility will be to determine any cis-acting effector/s sequence element present in *ki* locus, which is influencing expression of any gene/s which is/are involved in spermatogenesis. We wished to undertake this analysis and requested Yanaka and co workers to send us *ki/ki* mouse, unfortunately they turned down our request.

Therefore we urge to all the readers to take all the result together in perspective and decide themselves about possible function of *Theg*.

4.2 Characterisation of human *THEG* gene

Our interest to study human homologue of murine *Theg* was initiated with a vision that after functional characterisation of murine *Theg*, human *THEG* gene can be screened for possible mutation in a particular subgroup of male infertile patients. We would like to mention that in humans, male and female partners contribute more or less equally to the infertility problem. In ~15% of infertile couples, the concurrence of male and female factors is suggested to be responsible for infertility. Perhaps characterisation of *THEG* can serve as a useful tool for diagnostic purpose.

In this study, we describe the cloning, testis-specific expression, alternative splicing and chromosomal localisation of the human homologue of the mouse *Theg* gene. For human *THEG*, two transcripts with the nucleotide sequence of 1,290 bp (*THEG* major) and 1,218 bp (*THEG* minor) in length were identified (figure 3.25). These two transcripts are due to alternative splicing and differ 72 bp in the length of exon 3 sequences (figure 3.28B). This result is in agreement with previous reports for mouse *Theg* where it was shown that at least four alternatively spliced transcripts of mouse *Theg* exist (Nayernia et al., 1999; Yanaka et al., 2000). In these studies, it was demonstrated that alternative transcripts of *Theg* without exon 3 (72 bp in length), exon 4 (62 bp in length) sequences occur and moreover, an alternative transcript with an independent typical polyadenylation signal in the 3' untranslated region was detected. Interestingly, in the 3' untranslated region of the *THEG* cDNA sequence, 16 nucleotides upstream from the poly A⁺ tail, a modified polyadenylation signal (AATAAT) was found (figure 3.25). Modified polyadenylation signals were also described for transition protein2 (Kleene and Flynn, 1987; Luerssen et al., 1989), for subunits of cAMP-dependent protein kinases in male germ cells (Oyen et al., 1990) and for mouse and human outer dense fiber genes (Burfeind et al., 1991; Gastman et al., 1993). As was suggested (Oyen et al., 1990), those alternative signals could be important for germ cell-specific transcription and the translation of mRNAs after cessation of transcription. However, the cDNAs for several other testis-specific proteins, e.g. protamines and transition protein1, which are also transcribed in haploid spermatids, contain no alternative polyadenylation signal but the consensus polyadenylation signal sequence AATAAAA upstream of the poly A⁺ tail (Lee et al., 1987; Domenjoud et al., 1988; Luerssen et al., 1988). At nucleotide position 10 of *THEG* major and *THEG* minor transcripts, the start codon ATG of both open reading frames is found. The start codon is spanned by a sequence, which resembles the consensus sequence for initiation of translation by eucaryotic ribosomes (Kozak, 1987). This data is supported by the fact that we

have sequenced about 1 kb upstream of the ATG codon of genomic sequences from the PAC clone and no further in frame initiation codon was observed. The comparison of the *THEG* major sequence with the mouse *Theg* major sequence revealed an overall identity of 60% at the amino acid level (figure 4.4).

```

Theg: 78  SNLDGPLQODLEVEVEMSHLSITERTPSVSTAKGRKRSRR-LLELAKPKNWQCLDR 136
          S L+  L +DLE ++ E+S LSI+++ PS +  K RK+R RR L+ELA+PK NWQ L+DR
THEG: 85  SELERVLDKDLEEDIPEISRLSISQKLPSTTMTKARKRRRRRRRLMELAEPKINWQVLKDR 144

Theg: 137 TGRCKGYAWISPRKTNLQFCLYWPSVYWTERFIEDTTLTITVPVVSQRMEELSRPKRFY 196
          GRC KGYAWISP K +L FCL WPSVYWTERF+EDTTLTITVVP VS+R+EELSRPKRFY
THEG: 145 KGRCKGYAWISPKMSLHFCLCWPSVYWTERFLEDTTLTITVPAVSRRVEELSRPKRFY 204

Theg: 197 QEYYNNNRTPPIWSIPRSTLEYQASNRLKQLATPKVRNNIWSINMSEVSQVSRAAQMAMP 256
          EYYNNNRTPP+W IPRS+LEY+AS+RLK+LA PK+R+N WS+ MSEVSQVSRAAQMAMP
THEG: 205 LEYYNNNRTPPVWPIPRSSLEYRASSRLKELAAPKIRDNFWSMPMSEVSQVSRAAQMAMP 264

Theg: 257 TPRTLRLAKPRPPATLLEEWDPMPKPKPYVSDYNRLLQLATPKALSEKCVDRSPQWEVL 316
          + R L+L+KP+ PATLLEEWDP+PKPKP+VSD+NRLL LA PKA S+KCVDR P+WEVL
THEG: 265 SSRILQLSKPKAPATLLEEWDPVPPKPKPHVSDHNRLHLARPKAQSDKCVDRDRPRWEVL 324

Theg: 317 DVTKNAVASSRIISLAQPKIRKDLNEGYNPYYISPASL-VAQASP 360
          DVTK VAS RIISLA+PK+RK LNEGY+  ++ SL  +ASP
THEG: 325 DVTKKVVASPRIISLAKPKVRKGLNEGYDRRPLASMSLPPPKASP 369
    
```

Figure 4.4: Homology between mouse *Theg* and human *THEG*. The comparison of the human *THEG* major sequence with the mouse *Theg* major sequence revealed an overall identity of 60% and similarity of 81% at the amino acid level. Here plus (+) sign implies similar amino acid sequence.

The *THEG* gene was located in chromosome 19ptel-p13 by using fluorescence in situ hybridisation (figure 3.26). Previously, the homologous gene in the mouse was localised on chromosome 10 to region B5–C1, which is syntenic to human chromosome 19ptel.p13 (Nayernia et al., 1999). In addition, due to the release of the draft working sequences of the human genome, we identified two genomic clones located on chromosome 19 (clone CTD-3113P16, accession No. AC016588 and clone LLNL-FOS–22E10, accession No. AC010641), which contain the entire human *THEG* gene. The predicted exon/intron structure of the *THEG* gene revealed a similar exon/intron organization as found in *Theg*. Human *THEG* consists of eight exons, divided by seven intron sequences. In both mouse *Theg* and human *THEG* the

first exon contains 5' untranslated sequences, the initiation codon ATG and exon 8 contains the translational stop codon followed by the 3' untranslated region. By using these sequences by a computational procedure called electronic PCR (Schuler, 1998), we detected two sequence tagged sites (STS) in the reference interval ptel-D19S413 on human chromosome 19. In this region the human phosphatidic acid phosphatase type 2c gene was mapped (19p13) and its homologous gene in the mouse was mapped in the homologous region on mouse chromosome 10 (Zhang et al., 2000).

In the deduced amino acid sequences of both human THEG proteins, a four-residue pattern composed of either basic amino acids K and R or composed of three basic amino acids K, R and P was found at position 120 in the amino acid sequence (RKRR). An additional four-residue pattern RPKR was detected in the predicted THEG proteins at position 199 (THEG major) and 175 (THEG minor), respectively (figure 3.25). These patterns resemble putative nuclear targeting signals (NLS) of the SV40 large T antigen type (Hicks and Raikhel, 1995). Two additional seven-residue patterns found in the human THEG sequences also representing NLS are located at positions 197 and 342 (THEG major) and at positions 173 and 318 (THEG minor). In the deduced mouse Theg amino acid sequences two of the four-residue pattern NLS in the N-terminal part and one of the seven-residue pattern NLS in the C-terminal part of the protein are conserved. The presence of these nuclear targeting signals in both the human THEG and mouse Theg proteins indicates that THEG might also encode a nuclear protein.

A high concentration of genes and a number of candidate genes for recognisable syndromes (Saccone et al., 1992; Lamb et al., 1989; Altherr et al., 1991; Overhauser et al., 1989; Kuwano et al., 1991; Koehler et al., 1996) are known to be present in telomeric regions. The human telomeric regions represent a major diagnostic challenge in clinical cytogenetics. Since *THEG* is localised in teleomeric region of chromosome 19, further analysis of *THEG* might assist in answering many questions, specifically for chromosome 19.

4.3 Identification and characterisation of a novel mutation named *nax*

nax is a spontaneous autosomal recessive mutation, which arose in ES cells during the generation of first *Theg* knock-out mice. The phenotype of *nax* was very intriguing and at the same time very interesting. *nax* mice exhibit complete lack of hair after birth and displays growth retardation during development. However there is a delayed appearance of hair but the posterior part of adult mice remains naked. Strikingly *nax* mice starts to show an ataxic gait around stage P10, they walk around incessantly in their cage. Macroscopic inspection showed a pronounced reduction in the brain size. However the most severely affected region of brain is cerebellum. To comprehend and appreciate the beauty of *nax* phenotype, it is necessary to describe cerebellum biogenesis and few mutant mice models known in literature. In forthcoming discussion an abridged description about cerebellum development in mouse will be presented.

4.3.1 Overview of cerebellum development

4.3.1.1 Genesis of cerebellum

(Goldowitz and Hamre, 1998)

The cerebellum is derived from the alar (dorsal, sensory) plate of the neural tube and resides in the metencephalon. All cells in the cerebellum arise from a germinal matrix region. The cerebellum is one of the unusual structures in the brain where neuronal populations arise not from a single germinal zone but from at least two different germinal zones. The initial germinal matrix in the cerebellum is bounded by the isthmus anteriorly and the choroid plexus posteriorly.

The germinal matrix initially consists of a typical neuroepithelial ventricular zone and a more caudal germinal trigone region, also known as the rhombic lip (figure 4.5).

The first neurons to leave the ventricular zone are the nuclear neurons (embryonic days E10-12 in the mouse) that will settle deep to the cerebellar cortex. The birth and exit of Purkinje cells (E11-13 in the mouse) from the ventricular zone to form a temporary plate-like structure occur soon after the birth of the first nuclear neurons. The formation of a secondary germinal matrix from the rhombic lip-the external granular layer (EGL) occurs at about the time when nuclear and Purkinje cells have stopped dividing. As these EGL cells migrate over the cerebellar surface in a subpial position, the Golgi neurons are born from the diminishing ventricular zone. Postnatally, the EGL will seed the internal granular layer (IGL) with an

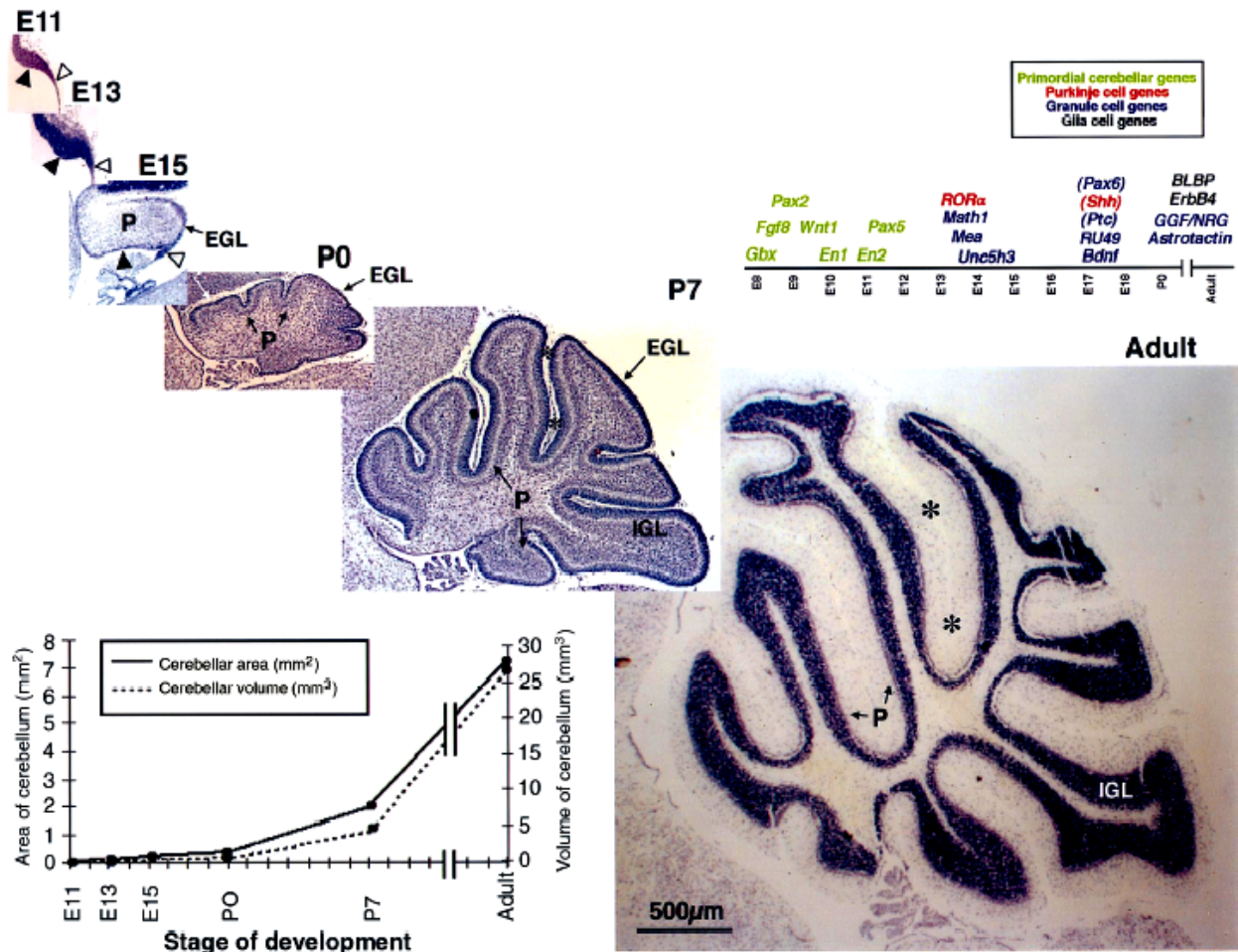


Figure 4.5: The cerebellum undergoes dramatic increases in size and changes in patterns of gene expression from its initial specification as the cerebellar anlage to its adult form. Sagittal sections of a mouse cerebellum from embryonic days E11, E13 and E15, as well as from postnatal days P0 (day of birth) and P7 and adulthood, are shown from left to right in the center of the figure to highlight the changes in size and complexity during development. In the photomicrographs, asterisks denote the primary fissure, P (with accompanying small arrows) points to the Purkinje cell layer, IGL labels the internal granule cell layer and EGL demarcates the external granular layer. In the three embryonic pictures, the filled arrowheads point to the ventricular neuroepithelium, and the unfilled arrowheads point to the germinal trigone. The scale bar for all photomicrographs is the same. The graph in the lower inset further emphasizes the change in cerebellar size. The axis on the left shows the area of the sections that are pictured, while the axis on the right provides estimates of the total cerebellar volume at these ages. The timeline in the upper right inset illustrates genes/loci that are known to have relatively specific effects on cerebellar development when mutated. The position of genes along the timeline denotes the most likely point at which the effects of the mutant gene are manifested, based upon descriptions of the phenotypes in the literature. Genes critical to the earliest stages of cerebellar development are shown in green. Those genes that are specific to granule cell, Purkinje cell, or radial glial cell development are shown in blue, red, or black, respectively, with genes in parentheses of presumed cellular localisation. (adapted from Goldowitz and Hamre, 1998).

abundance of granule cells. Also generated at this time are the stellate and basket cells that colonise the molecular layer (figure 4.5 and 4.6). Granule cell neuroblasts migrate in an inward, radial manner with the presumed aid of radial glial fibers (more properly known as the Bergmann fibers of the Golgi epithelial cells), through the molecular layer and past the developing Purkinje cells (figure 4.5 and 4.6). It is during this time that the cerebellum is transformed from a curved, sausage-shaped structure into one with deep fissures and large finger-like appendages called folia. The EGL ceases to exist by the end of the third postnatal week in the mouse, and this period also sees the maturation of the two major extrinsic inputs to the cerebellar cortex: the inferior olive's climbing fibers to the Purkinje cells and the spinal and reticular mossy fibers to the granule cells. The cerebellum undergoes over a 1000-fold increase in volume during this time (figure 4.5), and with all such growth there is a beginning of actual differentiation of cerebellum.

4.3.1.2 Maturation of the cerebellum

During early postnatal life, at least in rats and mice, the Purkinje cell undergoes numerous changes, most notably the spreading of the cell somas from a multi-layered structure to a monolayer and the orientation of the Purkinje cell dendrites towards the pial surface with the outgrowth of its dendritic branches (figure 4.7). In these changes, the impact of the granule cells or their precursors in the EGL are unequivocal. In cases where the granule cell population is disrupted, the Purkinje cells are poorly aligned in a rather ragged layer (Rakic and Sidman, 1973; Ross et al., 1990), and the Purkinje cell dendrite exhibits altered morphologies ranging from stunted to misoriented dendritic trees. The settling of Purkinje cells into a monolayer has been hypothesised to be a mechanical process (due to the migration of granule cells below and growth of parallel fibers above the Purkinje cells) (Altman and Bayer, 1997). The issue of whether there are other forces i.e. chemorepulsive, or molecules that might contribute to this process is still unexplored. Similarly, it is unclear whether the granule cells influence on Purkinje cell dendritic outgrowth (figure 4.7) is due to the expression of specific molecules or via electrical activity or a combination of both.

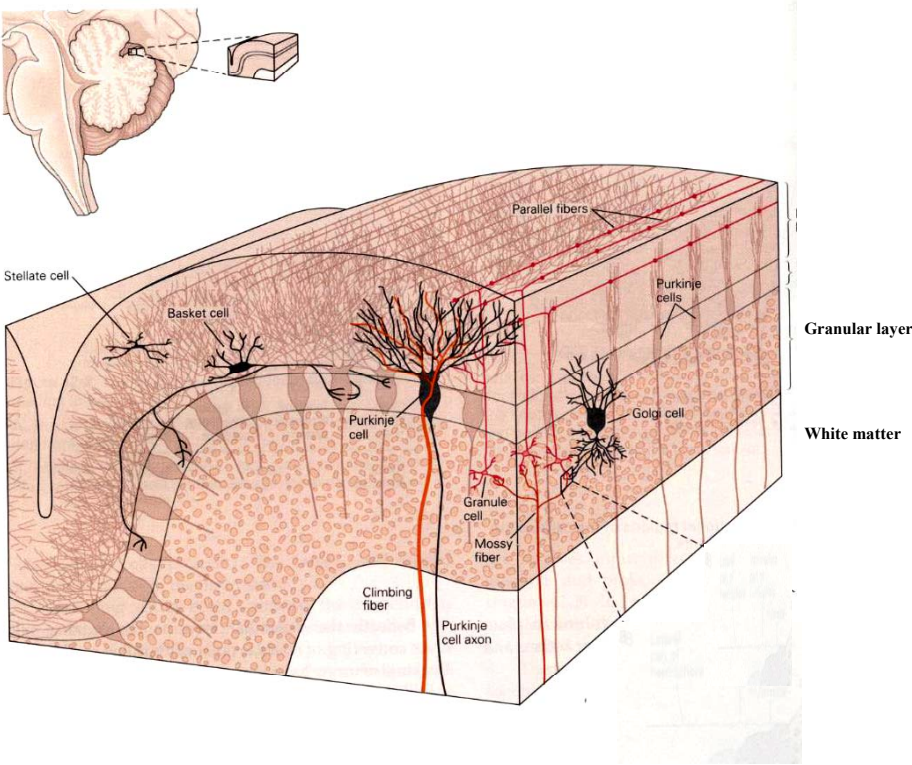


Figure 4.6: The cerebellar cortex is organised into three layers (Molecular layer, Purkinje cell layer, and Granular layer) and contains five different types of neurons (Stellate cells, Basket cells, Purkinje cells, Granule cells, and Golgi cells). A vertical section of a single cerebellar folium, in both longitudinal and transverse planes, illustrates the general organization of the cerebellar cortex (adapted from Ghez and Thach, 2000).

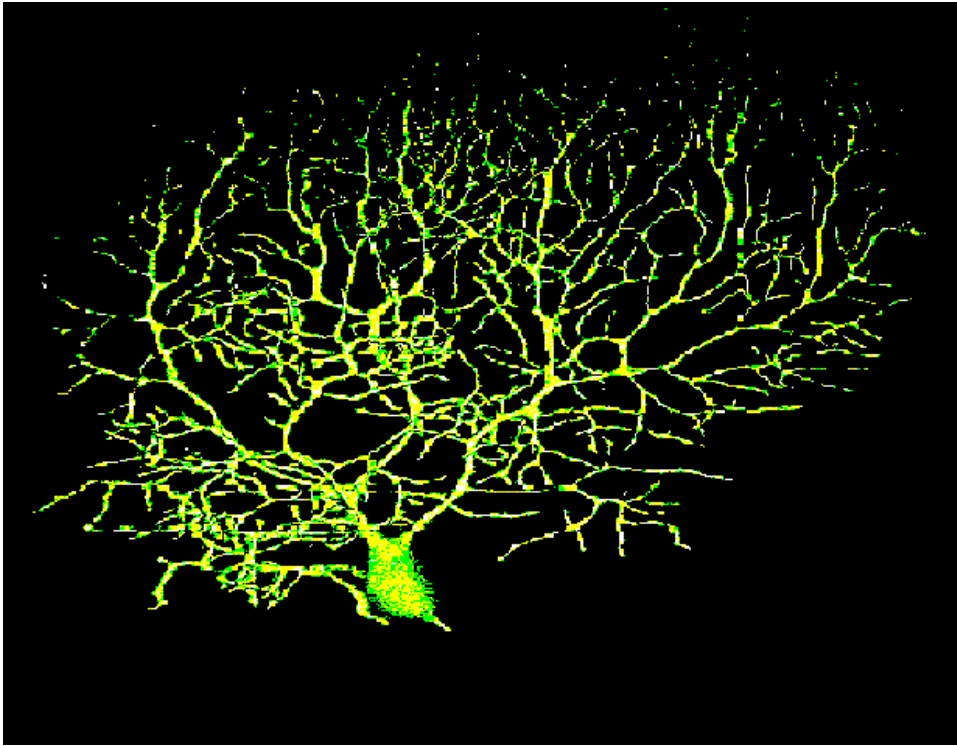


Figure 4.7: Confocal image of a Purkinje cell. This image shows a Purkinje cell that was patch-clamped in a sagittal cerebellar slice from a rat with a Lucifer Yellow-filled pipette and imaged with a laser scanning confocal microscope. Note the bush-like arrangement of dendrites, which arises from the soma, and the branching axon, which gives rise to three presynaptic terminals within the cerebellar granular layer. (adapted from Rossi et al., 1995).

4.3.1.3 A genetic dissection of the origin of the cerebellum

With the analysis of naturally occurring and induced mutations that affect cerebellar development, the mapping of patterns of early gene expression and the analysis of cell lineage using chimeric, mosaic and retroviral techniques, a veritable revolution is taking place in our understanding of the origin of the cerebellum and its cellular components.

The regional origin of the cerebellum

Traditionally, the cerebellum was thought to arise exclusively from the metencephalic territory of the developing neural tube (a region that extends posteriorly from the choroid plexus to the constriction between the mesencephalon and metencephalon anteriorly, termed the isthmus). However, experimental studies using the chick-quail chimera system provided compelling evidence for a dual, mesencephalic-metencephalic (mes-met) origin of the cerebellum (Martinez and Alvarado-Mallart, 1989; Hallonet et al., 1990). The junction between the mes- and metencephalon, the isthmus, has been shown, again using the chick-quail chimera paradigm, to have an organising potential on surrounding neural tissue (Martinez et al., 1991; Marin and Puelles, 1994; Martinez et al., 1995). For example, mes-met grafts to the diencephalon can organise the donor tissue to form mesencephalon, while mes-met grafts to the posterior rhombencephalon can organise the formation of a cerebellum. The special nature of the isthmus region is inherent in the number of important secreted and regulatory proteins that are expressed in this region: *Fgf8*, *Wnt1*, *En1*, *En2*, *Pax2*, *Pax5*, *Pax6*, *Pax8*, *Otx1*, *Otx2*, *Gbx2*, *Wnt5a* and *Nkx2.2*. The isthmus is a transition region for the expression of many of these molecules (Joyner, 1996). For example, *Fgf8* is expressed at the caudal portion of the isthmus, whereas *Wnt1* is expressed rostrally, in a non-overlapping, complementary fashion (Crossley and Martin, 1995). As testimony to the importance of these molecules, mutations in, or ectopic expression of, several of these genes have marked impact upon cerebellar development. The best-documented examples are *Wnt1* and *En1*. Knock-outs of these genes largely or totally eliminate the cerebellum (McMahon and Bradley, 1990; Thomas and Capecchi, 1990; Wurst et al., 1994). In mouse, as in fruit fly, it is believed that *En* expression is regulated by the secreted glycoprotein Wnt (McMahon et al., 1992)). The expression of these molecules might, in turn, be regulated by paired box-containing genes of the *Pax* family (Song et al., 1996). Indeed *Pax2* is found to presage *En1* and *Wnt1* expression in the presomitic embryo (around E8 in the mouse). Inactivation of *Pax2* leads to a frequent loss of the cerebellum and posterior midbrain (Favor et al., 1996). The next *Pax* gene to be expressed in the midbrain-hindbrain region is *Pax5*. Although deletions of *Pax5* have little effect on cerebellar development, when a *Pax5* knock-out is complemented with a hemizygous deletion of *Pax2* there is complete agenesis of the cerebellum in addition to the inferior colliculus (Urbanek et al., 1997). Other molecules that are expressed in the neighborhood of the mes-met junction, namely *Fgf8* and the homeodomain proteins *Gbx2* and *Islet3*, are known to be important players in cerebellar development, as their elimination results in large deletions of

the cerebellum (Kikuchi et al., 1997; Wasserman et al., 1997; Meyers et al., 1998). These three molecules are believed to act upstream from *Wnt1* and *En*. These results are starting to form a molecular framework of the initial development of the cerebellar anlagen (figure 4.5, inset).

The contribution of the ventricular zone to neuronal populations in the cerebellum

The cells that colonise the cerebellum arise from either the ventricular neuroepithelium or the rhombic lip. The ventricular neuroepithelium is responsible for the generation of most of the neuronal populations within the cerebellum, including the nuclear, Purkinje and Golgi neurons. Fundamental questions are just being answered about the lineage relationships of these various neurons, although little is known about the cellular interactions and molecules that are crucial to commitment to each lineage. The intriguing work of Mathis et al. (1997), using a novel cell-marking construct, has found that all cerebellar neurons, with the exception of the granule cells, are clonally related. They found evidence that Purkinje cells and nuclear cells arise from a common progenitor.

A further issue related to the ventricular epithelium concerns the generation of molecular layer (ML) inter-neurons. Until recently, the prevailing dogma was that ML inter-neurons were generated in the secondary germinal zone (the EGL). This belief was based on the fact that the generation of both of these neuronal populations occurs postnatally and that the only substantive germinal zone at this time is the EGL. Recently, several lines of evidence made it clear that the formation of the ML lineage occurs independently of the formation of the EGL. The analysis of chick-quail chimeras demonstrated that the origin of the molecular layer inter-neurons and granule cells was not equivalent and that, unlike granule cells, the molecular layer inter-neurons were generated from both the mesencephalon and the metencephalon (Martinez and Alverado-Mallart, 1989; Alvarez et al., 1993; Hallonet and Alverado-Mallart, 1997). Evidence from several independent analyses (Mathis et al., 1997; Zhang and Goldman, 1996) has shown a similar lineage separation in mammals. Zhang and Goldman, using retroviral infections either superficially into the EGL or deep into the white matter of cerebellar cortex, found that injections into the EGL generated only granule cells (Zhang and Goldman, 1996). Alternatively, injections into the white matter generated both astrocytes and molecular layer inter-neurons. Because only cells undergoing mitosis take up the retrovirus, the molecular layer inter-neurons must be undergoing a round of mitosis within the white matter. This led Zhang and Goldman to hypothesize that the generation of ML inter-neurons

is a two-step process, with the initial generation occurring in the ventricular layer and subsequent migration into the white matter of the cerebellar cortex, where they undergo further cell division before translocation to the molecular layer.

The contribution of the secondary germinal zone to the cerebellum

Granule cells have been recognised to arise from the secondary germinal zone since the time of Ramón y Cajal (Ramón, 1911). This has been supported by both retroviral labelling studies and cell culture studies in which only granule cells were generated from cultured EGL cells. Granule cell lineage is, on one hand, very clear: the cells arise from the EGL, and the EGL appears to come exclusively from the metencephalon. On the other hand, the early events of colonisation of the rhombic lip, the number of precursors that establish the granule cell lineage and any evidence for heterogeneity of this massive population of cells are all either lacking or rudimentary. A major finding that illuminated the early development of the granule cell was the discovery that the helix–loop–helix transcription factor *Math1* is essential for the establishment of the granule cell lineage (Ben-Arie et al., 1997).

4.3.1.4 Migration of cerebellar neurons and neuron-glia interaction

The developing cerebellum can well serve as a model for most of the recognised means by which neurons migrate in the CNS. There is an outward radial (Purkinje and nuclear neurons) and inward radial (granule cells) migration (Altman and Bayer, 1997), as well as a circumferential (formation of the EGL) and tangential migration of cells (EGL) (Ryder and Cepko, 1994). The initial migratory opportunity in the cerebellum is the departure of the first cohorts of cells from the ventricular zone and germinal trigone/rhombic lip, which occurs at about E13 in the mouse. These migrating cells define the future boundaries of the cerebellum, but this process was not appreciated until the description of the mutant phenotype in the rostral cerebellar malformation mutant mouse and the identification of the mutated gene responsible for the phenotype (gene symbol *rcm*, now called *Unc5h3*) (Ackermann et al., 1997). In this mutation, there is an exuberant migration of granule and Purkinje cells beyond the normal confines of the cerebellum into the inferior colliculus and hindbrain.

The spate of newly identified genes that are known to be critical to neuronal migration (*reelin*, *disabled*, *cdk5*, *p35*) offers keys to the genetic underpinnings of migratory behavior in the CNS (figure 4.8).

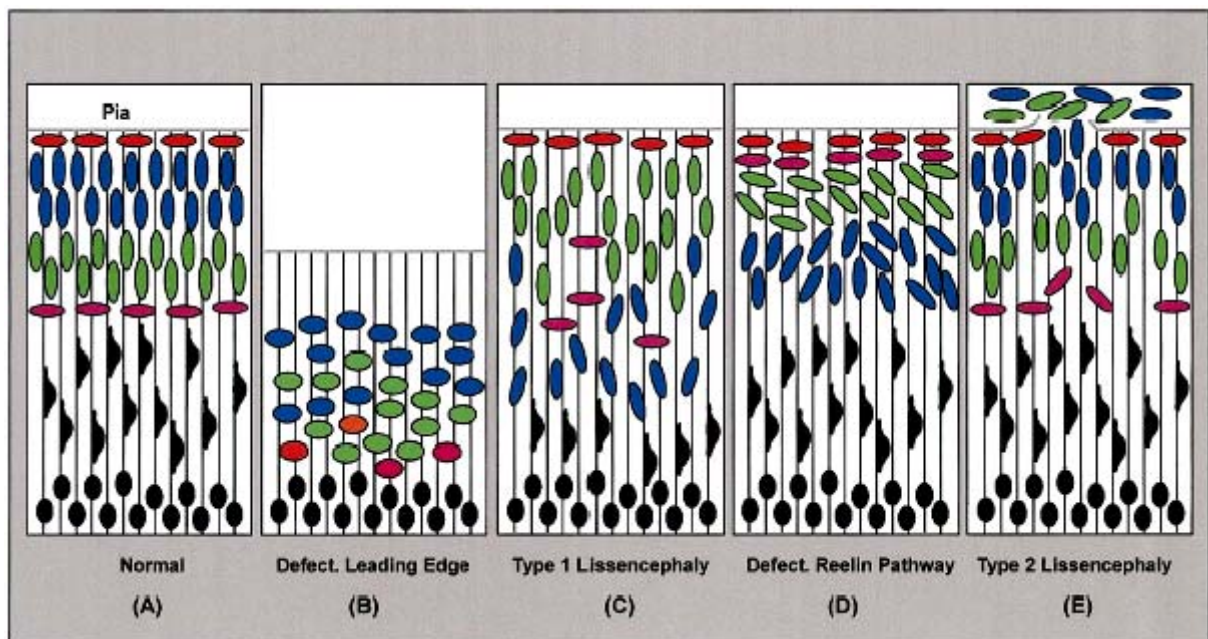


Figure 4.8: Schematic illustration of normal (A) and different abnormalities of cerebellum cortical migration (C-E). In the normal cortex, neurons proliferate in the ventricular zone and migrate along radial fibres. Early post migratory neurons settle horizontally in the marginal zone (red) and sub plate (pink). Reelin produced in the marginal zone (red neurons) helps organize post migratory neurons into layers. Early born cortical neurons (green) are found in the depth of the cortex and younger cells (blue) more superficially, according to the so-called 'inside-outside' gradient. When leading edge extension is defective (B) as in familial heterotopias, neurons cannot migrate from the ventricular zones and settle locally. In case of defective nucleokinesis (C), as occurs probably in type 1 lissencephaly, initial migration is satisfactory but many cells settle at sub cortical level. In the absence of Reelin signalling (D), cortical neurons (green) settle obliquely in the cortex and displace the early contingent (red and pink) 'en bloc' in the marginal zone; the next generation (blue) cannot cross the first one and the gradient is directed from outside to inside. Finally, when the limiting membrane is defective (E), as in type 2 lissencephaly, streams of neurons over migrate in the marginal zone and in the meninges (adapted from Lambert de Rouvroit and Goffinet, 2001).

In mutant *reeler* and *scrambler* mice [caused by mutations in an extracellular matrix molecule, Reelin (D'Arcangelo et al., 1997), and a tyrosine kinase signaling adaptor molecule Disabled (encoded by *Dab1*) (Sheldon et al., 1997; Howell et al., 1997; Ware et al., 1997), respectively], Purkinje cells fail to migrate and largely remain in several ectopic clusters deep in the cerebellar cortex, while the granule cells appear to migrate normally but are greatly reduced in number (Goffinet et al., 1984; Goldwitz et al., 1997) (figure 4.9). In neocortex it is clear that there is a single population of Reelin-positive cells (the Cajal–Retzius cells) that are spatially juxtaposed to Disabled-positive cortical and pyramidal cells. In the cerebellum, however, two populations of Reelin positive cells have been identified: a subpopulation of

cerebellar nuclear neurons and cells of the external granular layer (D’Arcangelo et al., 1995; Miyata et al., 1996). The responding cells in the cerebellum appear to be the Purkinje cells in that they both bind Reelin and express the Disabled protein (Sheldon et al., 1997; Miyata et al., 1996). While these two molecules are believed to be in the same molecular pathway (Sheldon et al., 1997; Goldwitz et al., 1997), it is not clear how *cdk5* and *p35* fit into this scheme, as the phenotypes of these mutants are different from *reeler* and *scrambler* [*cdk5*^{-/-} mice die as newborns (Ohshima et al., 1996) and *p35*^{-/-} mice are reported to have a normal cerebellum (Chae et al., 1997)].

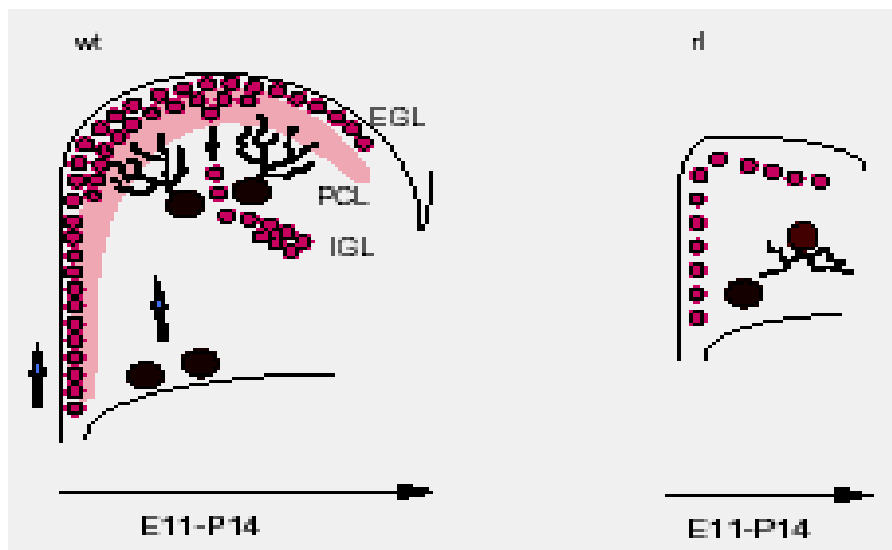


Figure 4.9: Reelin and the development of the cerebellum. Cerebellum development begins around embryonic day 11. Granule cells (red) originate from the rhombic lip and migrate tangentially over the surface of the developing cerebellum to form the EGL. These cells produce Reelin (shaded red) in normal animals (wt). Purkinje cells (brown) migrate radially from the ventricular zone to a location near the Reelin producing granule cells. In *reeler* mice, granule cells do not produce Reelin, Purkinje cells remain in a deep sub cortical location, and granule cells are reduced in number. As a result, the *reeler* cerebellum is very small. During postnatal days (P) 0–14, granule cells in the normal cerebellum migrate inwardly across the PCL to form the IGL (adapted from D’Arcangelo and Curran, 1999).

The cerebellum may be the ideal place to decipher the interplay of cells and molecules in the migratory process because of the limited numbers of cells and interactions. It is just this simplicity that has been exploited by Rakic and co-workers to study the molecular underpinnings of cell migration from the EGL. They have identified antibodies that recognise radial glia and the junctional complexes between glia and neurons, and have shown that some of these antibodies can perturb neuron-glia recognition and neuronal migration (Cameron and Rakic, 1994; Cameron et al., 1997; Anton et al., 1996). In addition, Komuro and Rakic, in an elegant set of studies, have demonstrated an important role for NMDA receptor activation in the migration of granule cells from the EGL (Komuro and Rakic 1993; Komuro and Rakic

1996). The weaver mutation, a single base pair substitution in the G-protein-coupled inward-rectifying K⁺ channel protein *Girk2* (Patil et al., 1995), appears to specifically affect the preparation of cells to migrate (Sneyne and Goldowitz, 1989). At the cellular level, mutant mice have taught us that granule cell survival is compromised and migration does not occur when radial glia are disordered and 'atrophic' (Rakic and Sidman, 1973; Ross et al., 1990). Furthermore, from chimera analyses of these mutants, it has been determined that the glial abnormality is secondary to a defect in granule cells (Hamre and Goldowitz, 1997; Goldowitz, 1989). Additionally, it has been determined that, if glia are experimentally eliminated from the developmental equation, granule cell precursors suffer large-scale cell death (Delaney et al., 1996). The molecular underpinnings of this two-way interaction appear to be, on the one hand, a neuron-glial trophic signal and, on the other, a glial-neuron buffering of the surrounding milieu. Astrotactin and glial growth factor (or Neuregulin) are two molecules that are expressed in migrating cerebellar granule cells and provide a trophic support for glial development (Zheng et al., 1996; Rio et al., 1997). While a specific receptor has not yet been found for Astrotactin, the EGF receptor tyrosine kinase family ErbB is known to be activated following Neuregulin stimulation. A recent analysis has shown that Neuregulin and ErbB4 are localised to granule cells and glia, respectively, and that this signalling pathway is important to granule cell migration as well as the induction of radial glia morphology (Rio et al., 1997). Another molecule that has been localised to the glia cell and appears important for neuron and glia cell differentiation is brain lipid-binding protein (BLBP) (Feng et al., 1994; Feng and Heintz, 1995).

4.3.1.5 Corollary remarks

The great advances that have been made concerning the molecular and cellular bases of cerebellar development within the last decade presage even more exciting times to come. Currently, our state of knowledge of the molecular underpinnings of development is only rudimentary. Furthermore, it will be an endeavour to understand how these molecules subserve the cellular interactions that are critical for cerebellar development. Finally, the most challenging goal should be relating the emerging picture of cerebellar development to the normal and abnormal function of the cerebellum.

4.3.2 Characterisation of *nax* phenotype

nax is an autosomal recessive mutation, heterozygous mice are indistinguishable from normal mice, whereas homozygous mice can be distinguished from its littermates by phenotypic observation. There is a complete lack of hair in their body, the skin appears to be wrinkled and they have long protruding ears. At stage P10 they can be clearly distinguished from their siblings because by this stage all wild type mice develop hair, in contrary *nax* mice are completely devoid of any body hair and they remain naked. Therefore they were named *nax*. A striking feature, which was observed at stage P10, is that *nax* mice show symptoms of ataxia, tremors, and imbalance. They walk around incessantly in their cage and frequently they fall down. It seemed that they were unable to keep their body balance and locomotion is uncoordinated, which could be because of motor imbalance resulting from improper or disrupted function of brain. Neurological abnormalities affecting behaviour or locomotion have been identified frequently in inbred mouse strains. Lineages carrying these mutations are maintained at the Jackson Laboratory in Bar Harbor, U.S.A. (<http://www.informatics.jax.org>). Several mutants exhibit abnormal locomotion associated with aberrant cerebellar development or function. A catalogue of neurological mouse mutants established the first database of spontaneous neurological mutations (Sidman et al., 1965). In this catalogue, four mutations were reported to have a dramatic effect on cerebellar development: *reeler*, *weaver*, *straggerer*, and *leaner* (Sidman et al., 1965).

Macroscopic analysis of *nax* brain showed that it is significantly smaller than that of wild type brain. Although all subdivisions of brain are present in *nax* mice, however subtle distortion in some subdivisions can be observed. Analysis of sagittal sections of brain showed that most severely affected subdivision of brain is cerebellum (figure 3.33), which will be discussed in detail in forthcoming sub-section.

Other apparent abnormalities observed in *nax* mice are: firstly, they are smaller in size than their wild type littermates. A growth curve analysis of their body weight showed significant growth retardation, during *nax* mice development the growth is reduced by (almost) 50% compared to that of wild type mice (figure 3.32). Adult *nax* mice are almost half the size to that of their normal littermates. Secondly preliminary analysis of skin section of *nax* mice showed some major abnormalities, however further studies are necessary to comment on the aetiology of skin.

4.3.2.1 *nax* cerebellum and neurodevelopment

The cerebellum is composed of a cortex and deep nuclei. But unlike the more complex and heterogeneous neocortex, the adult mouse cerebellar cortex contains two principal cell types granule and Purkinje cells, each forming in a distinct layer (figure 4.5 and 4.6).

Our phenotypic analysis of the *nax* cerebellum raises interesting issues concerning mechanisms of cerebellar development. The anatomical structure of the *nax* cerebellum is very different from those mouse mutants previously reported to have a dramatic effect on cerebellar development namely: *reeler*, *weaver*, *straggerer*, and *leaner* (Sidman et al., 1965), table 4.1 summarise description of cerebellar defects in major identified spontaneous mutant mouse.

Mutant name	Symbol	Primary defect	Gene
<i>swaying</i>	<i>sw</i>	Absence of anterior lobes	Wnt-1, extracellular protein
<i>reeler</i>	<i>rl</i>	Purkinje cell ectopia, granule cell reduction	Reelin, extracellular protein
<i>weaver</i>	<i>wv</i>	Granule cell degeneration	GIRK2, K-channel subunit
<i>tottering/leaner/rolling mouse Nagoya</i>	<i>tg/tg^{la}/tg^{td}</i>	Purkinje cell and granule cell loss in anterior lobes	α 1A-subunit of Ca channel
<i>staggerer</i>	<i>sg</i>	Purkinje cell reduction and ectopia	ROR- α , orphan receptor
<i>rosto-cerebellar-malformation</i>	<i>rcm</i>	Granule cell ectopia in the anterior lobes	RCM, similar to netrin receptors
<i>lurcher</i>	<i>lc</i>	Purkinje cell degeneration	δ ₂ -glutamate receptor

Table 4.1: List of identified spontaneous cerebellar mutations (adapted from D’Arcangelo and Curran, 1998).

The cytoarchitectural analysis of *nax* cerebellum (with crystal violet staining of sagittal sections of cerebellum) revealed a normal foliation of cerebellar cortex and they also form normal fissures in each folium, although fissures are thinner and shallower than that of wild type cerebellar cortex (figure 3.34A). However the most striking abnormality observed in adult mouse cerebellar cortex is that there is a complete lack of inner granular layer (IGL) (figure 3.34A). It is very difficult to determine the cause of lack of IGL, however we postulate two possible reasons for lack of IGL: first, granule cells from external granule layer (EGL) fails to migrate inwards in a radial manner, for example, as in cyclin-deficient kinase 5 (*cdk5*) mice (Ohshima et al., 1997). The second reason could be during cerebellum differentiation, granule cells were degenerated, like in *weaver* mice (D’Arcangelo and Curran, 1998). It is

very difficult to comment about granule cell fate in *nax* cerebellum due to absence of any granule cell specific marker.

The Purkinje cell bodies also show an ectopic positioning, instead of forming a single thin layer of Purkinje cell layer (PCL), the Purkinje cell bodies are placed in dispersed manner in *nax* cerebellum, as determined by immunostaining for Calbindin, which is a specific marker for Purkinje cell bodies. However the ectopic placement of PCL in *nax* cerebellum is very different from that of known mouse mutants like *reeler* and *scrambler* (Miyata et al., 1997; Goldowitz et al., 1997). In *reeler* and *scrambler* mouse the number of Purkinje cell count is remarkably reduced and the PCL is almost non-existent. However in *nax* cerebellum, the number of Purkinje cell number is comparable to that of wild type cerebellum but instead of forming a thin PCL, it seems that PCL is much more broader and dispersed.

The order of axon and dendrites also appears to be distorted in *nax* cerebellum. A Bergmann fiber, which originates from molecular layer (ML) show a completely distorted arrangement instead of a radial positioning as can be seen in wild type cerebellum with immunostaining with glial filament acidic protein (GFAP) (figure 3.34C).

Analysis of sagittal sections of cerebellum from different developmental stage of mice indicates that the onset of *nax* phenotype is prenatal. The cerebellum from stage P20, showed a dispersed PCL in *nax* cerebellum as observed in adult *nax*. However the extend of ectopic placement of Purkinje cell bodies is diminished (figure 3.35A). It could be due to disruption of signalling pathway involved in Purkinje cell migration or Purkinje cell fails to respond to signalling ligand. Analysis on cerebellum from stage P0 also show an ectopic placement of PCL (figure 3.35C), therefore suggesting that the onset of *nax* phenotype is at stage of prenatal development of cerebellum. It is very difficult to obtain *nax* embryo for prenatal analysis, as breeding of *nax* homozygous mice is very difficult due to behavioural phenotype, although both male and female *nax* mice are fertile. However after lots of failed attempt, we finally managed to obtain *nax* embryos and analysis of cerebellum from stage E16 is in progress. Analysis on the developing cerebellum from different prenatal stages of mouse will enable us to determine the exact time of onset of *nax* phenotype.

4.3.3 Identification of gene/s responsible for *nax* phenotype

As mentioned before *nax* mutation arose during the generation of *Theg* knock-out mice. From our previous experience in generation of knock-out, we concluded that there are two different possible reason for generation of *nax* mutation. It could be due to insertion of partial fragment

of *Theg* targeting vector randomly into the genome of ES cell or *nax* is result of a spontaneous mutation in ES cell, which is quite frequent during passaging of ES cell line.

The first possibility is that an insertion of *Theg* targeting vector is causative reason for *nax* phenotype was ruled out by Southern blot analysis. The ideology behind this experiment was if there is any random integration of any fragment of *Theg* construct in mouse genome then in Southern blot analysis on genomic DNA digested with various restriction enzymes, we would be able to detect additional band/s other than wild type fragments. We used five different fragments of *Theg* targeting vector as probe for this Southern blot screening, however we could not detect any additional band in our blots (figure 3.29 and 3.30). Therefore *nax* could have arisen due to a novel spontaneous mutation.

4.3.3.1 Mapping of *nax* locus by linkage analysis

In order to perform linkage analysis, we undertook an intercross breeding strategy. The original mutation arose in ES cells, which are derived from 129X1/SvJ strain. Our F1 animals were from mixed genetic background (C57BL/6J x 12X1/SvJ) strain. We derived our F2 *nax* animals (which we used for linkage analysis) from intercross F1 animals, in a similar fashion we obtained F3 and F4 *nax* animals also. The breeding strategy is described in detail in methods section 2.2.23.1. Our linkage analysis was based upon classical two-step strategy. The goal of the first stage was to identify a definitive linkage to a defined sub-chromosomal interval. This can be accomplished by typing a relatively small set of markers on a subset of phenotypically typed animals within the larger panel. Once this first stage was completed, it became possible to proceed to the second stage, which should focus on the construction of a high-resolution map just in the vicinity of the locus of interest with a selected set of markers and a selected set of animal samples. The ultimate goal of this entire protocol is the identification of a handful of markers and recombinant animals that bracket a very small interval containing an interesting gene that can then be subjected to positional cloning.

We carried out our initial genome-wide linkage analysis with 47 microsatellite markers on 48 animals, which included 42 affected *nax* animals and 6 parental controls. The coordinates of the microsatellite markers based upon a comprehensive framework map were taken from Jackson Laboratory's database. The association of linkage to the locus of interest was determined by simply comparing the pattern of allele segregation for each marker. Essentially, the frequency of recombination for each anchor locus (marker) was calculated, one at a time, to identify one or more anchors that show a significant departure from the independent assortment frequency of 50%. This task is performed most easily by entering the

accumulated allele segregation data into an electronic file that is analysed by a special computer program developed for this type of analysis. By this method we were able to identify a strong linkage to marker D2mit206, whose coordinates are, chromosome 2; 51.6 cM. The allelic frequency for genotype of 129X1/SvJ strain at locus, D2mit 206 was 95%, therefore the odd against maximum likelihood to this locus was infinitesimally small (table 3.3).

After mapping the *nax* locus to a sub-chromosomal interval, we proceeded to the second stage of our linkage, haplotype analysis, which focuses on the construction of a high-resolution map just in the vicinity of the locus of interest i.e. D2mit206, with a selected set of markers and a selected set of animal sample. The ultimate goal of this entire protocol was to narrow down the locus of interest to a minimal region possible based upon information provided by animals, which underwent critical recombination breakpoint event with the help of a handful of markers. With this approach we were able to place *nax* locus between markers D2mit158 (45.5cM) and D2mit100 (47.5cM) (figure 3.37 and 3.38). One *nax* animal had undergone critical recombination breakpoint in one of the sister chromatids of chromosome 2 near marker D2mit158 locus from the proximal position and three animals showed critical recombination breakpoint event from the distal position at marker D2mit100 in one sister chromatids of chromosome 2 (figure 3.37).

In summary from our linkage analysis we were able to map *nax* locus within a genetic distance of 2 cM between microsatellite markers D2mit158 and D2mit100 in chromosome 2.

4.3.3.2 Determining gene order: generating a physical map

Once linkage has been demonstrated for a new locus, it is usually straightforward to determine its relative physical position of anchor markers on the chromosome framework map. In order to determine the physical position of markers D2mit158 and D2mit100 on chromosome 2, we searched for above markers sequences in Celera database. To our surprise, the result from Celera database indicated that the physical distance between markers D2mit 158 and D2mit100 is approximately 14 Megabase pair (Mb), spanning five different scaffold of chromosome 2 in Celera database (figure 3.39). In genome wide level (in mouse), 1 cM approximately corresponds to a physical distance of 2 Mb, however a possible deviation from this relation is not unlikely also. For example in the vicinity of recombination hot spot the

actual physical distance will be smaller and in case of rare recombination event spot the distance will be longer.

However we again checked the coordinates of markers D2mit 158 and D2mit100, but this time we looked them in the comprehensive map provided by Whitehead Institute of Biomedical Research (WIBR) database, as originally these markers were designed by WIBR. We discovered that WIBR database indicates different coordinates than that provided by Jackson Laboratory (JL) database, the coordinate from WIBR database were 44.8 cM for D2mit158 and 50.3 cM for marker D2mit100. Therefore the genetic distance between these two markers will be 5.5 cM, which will be in accordance with physical distance determined from Celera database. An intensive comparison of framework map between JL and WIBR database revealed many more errors in JL database, therefore we would suggest our readers to use WIBR coordinates for their linkage analysis.

In this region of 14 Mb, we identified two candidate genes, *Neurod1* and *γ-catenin*. *Neurod1* is a helix loop helix transcription factor, which is highly expressed in brain and in pancreas, a knock-out mouse model showed a growth retardation, diabetes and cerebellum defects (Liu et al., 2000). *γ-catenin* is a postulated member of Wnt signalling pathway and in mouse mutant *swaying* (which shows cerebellum defect), *Wnt-1* was mutated (D’Arcangelo and Curran, 1998). However we could not find any mutation in both genes in *nax* mouse. The *nax* locus spanned a region of about 14 Mb, with more than 200 genes reported in this region, a region of this size is still quite large for undertaking gene identification studies, and thus it makes sense to try to narrow it down further.

4.3.3.3 A new haplotype analysis

The goal of this step is to identify the closest limiting markers both sides of the locus of interest (*nax* locus), in order to establish a minimal interval within which the *nax* locus must lie. In public database, this locus was coarsely mapped, therefore we generated a high-density map of this locus by screening for polymorphisms in about 400 putative microsatellite markers, which were identified by searching for tandem repeat elements in the sequence of *nax* locus (*nax* locus sequence was downloaded from Celera database). Our new map comprised of 9 novel microsatellite markers, which were evenly distributed within this region of 14 Mb (figure 4.10). Based upon our newly characterised map, we genotyped 45 affected *nax* animals for additional new markers and constructed a new haplotype for each affected animal. With this approach we were able to refine further the putative susceptibility *nax* locus to a physical distance of 800 kb, which was between marker M98 (88.8 Mb) and marker

M140 (89.6 Mb) (figure 3.40). At this point, we reached a stage called limiting marker stage and we exhausted the available source of polymorphic markers within a region of 800 kb, therefore it was impossible to refine the locus of interest further, under current breeding scheme i.e. between C57BL/6J x 12X1/SvJ.

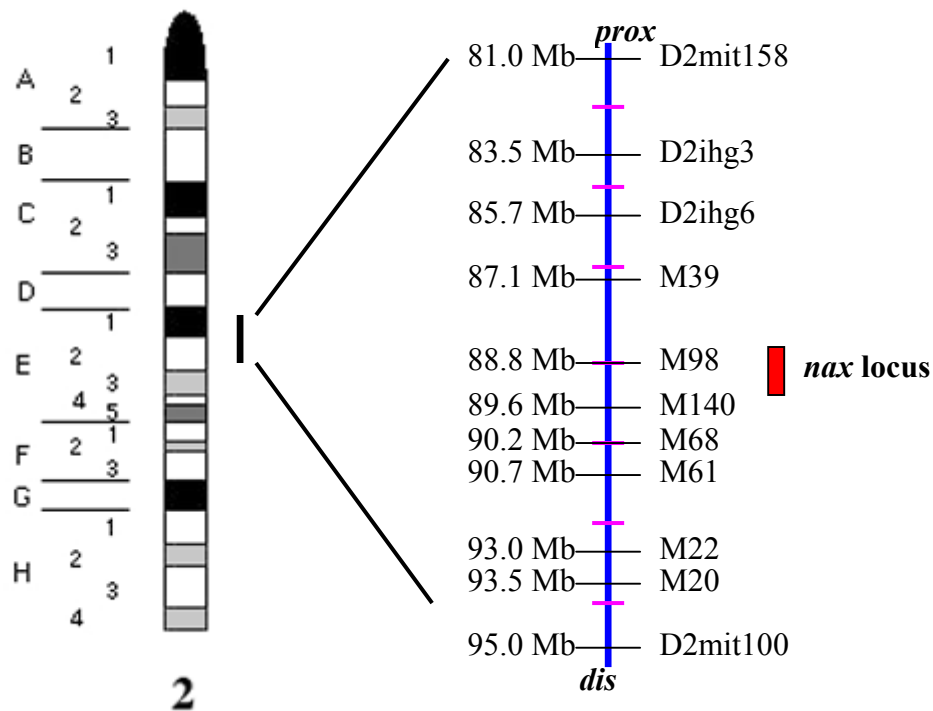


Figure 4.10: Ideogram of mouse chromosome 2. Physical map showing approximate physical location and order of the 11 microsatellite markers in the fine mapping panel that spans the regional assignment for *nax* locus. The physical distances between the markers are in megabase pair (Mb). *prox*, proximal tip of chromosome tip; *dis*, distal tip of chromosome.

4.3.3.4 In pursuit for candidate gene?

At this stage of analysis, *nax* locus resides within a region of 800 kb. Within this region 24 genes were reported in Celera database (figure 4.11). A number of these genes have been shown to be expressed in embryonic brain and adult cerebellum of mouse. However, there is no known transcript that is expressed exclusively in cerebellum and therefore none of them (genes) are obvious candidate/s for *nax* phenotype based upon their expression pattern.

We initiated a systematic analysis, which included Northern blot, RT-PCR and genomic sequence analysis for all the 24 genes. However no mutation was found that could be

attributed to *nax* phenotype (table 3.4). We also didn't observe any change in level of expression of genes (present in *nax* locus) between *nax* and wild type mouse.

The reason for not finding the mutation responsible for *nax* phenotype in our analysis could be due to possibility that there are some unknown genes present in *nax* locus, which are not report in Celera database. This is likely as Celera database mainly relies on computational program to detect genes and this method has its own inherent drawbacks.

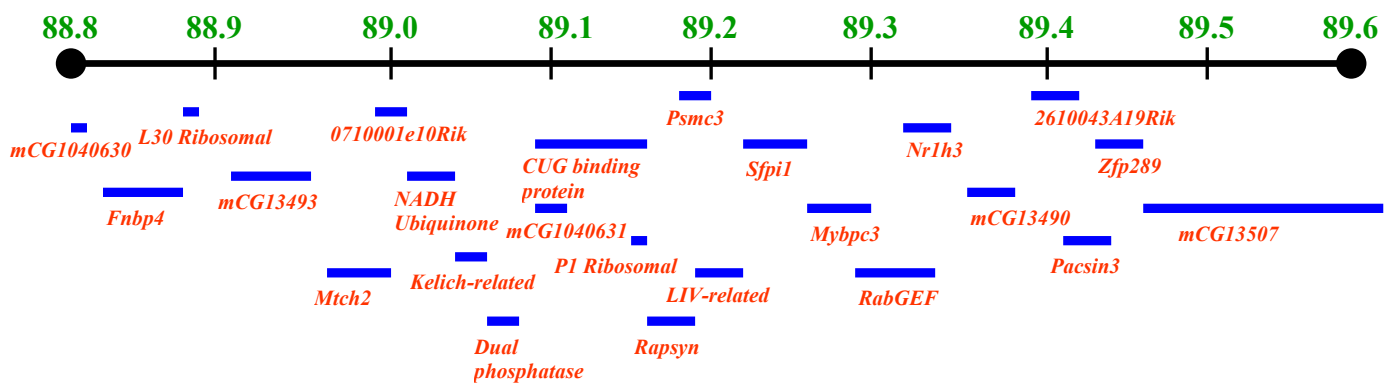


Figure 4.11: Schematic representation of *nax* locus. Illustrating the position of all the 24 genes reported in Celera database.

4.3.3.5 Future perspective

1. We identified more than 80 ESTs, which are present in *nax* locus. Further characterisation of ESTs in the interval of *nax* locus and identification of other transcripts are therefore necessary to identify candidate gene/s.
2. In order to narrow down the locus further, *nax* mutation was introduced into DBA/2J mice strain background, as *nax* locus is identical between 129X1/SvJ and C57BL/6J background, it was not possible to score for polymorphic microsatellite markers. However *nax* locus is highly polymorphic between C57BL/6J and DBA/2J strain. By introducing *nax* mutation from C57BL/6J background to DBA/2J strain, we will be

able to narrow down the *nax* locus further based upon our scoring for polymorphic Single Nucleotide Polymorphism (SNP) markers, which were identified from Celera database.

3. In another approach, complementation for *nax* mutation will be performed; we have identified BAC clones spanning the complete *nax* locus from Celera database. These BAC clones will be injected into fertilized oocytes, which are homozygous for *nax* mutation and then transferred into pseudopregnant mice. By this method we hope to rescue *nax* phenotype by one or more BAC clones.
4. We are also undertaking a differential display mRNA technique. This will enable us to identify all those genes, which are down regulated or up regulated in *nax* cerebellum compared to that of wild type cerebellum. In parallel, we will perform Atlas array hybridisation on two different cDNA libraries: Adult mouse cerebellum library and whole embryo (stage E17) with labelled cDNA derived from cerebellum of *nax* and wild type mouse (stage P10). Taken together these two different but redundant approaches will increase the sensitivity of data. Our ultimate goal will be to characterise any pathway involved in cerebellum biogenesis, which is defective in *nax* mouse. At the same time, it will also assist us to identify the gene/s which is/are mutated in *nax* mouse.
5. A complete phenotypic characterisation of *nax* mutant mouse, which will include sagittal section analysis of other subdivisions of brain to determine if there are any subtle defects in these structures. A detailed analysis of skin sections and other tissues are also necessary to identify other abnormalities present in *nax* mouse.
6. Establishment of primary cell culture system for cerebellar cells derived from *nax* cerebellum. This will enable us to study the cellular defects manifested by *nax* cerebellar cells and even will let us to identify genes whose expression is altered in *nax* mouse cerebellum.

5. Synopsis

Stage-specific interactions between Sertoli cells and germ cells are central in the regulation of spermatogenesis, but owing to its complex cytoarchitecture the mediators of Sertoli cell-germ cell interactions remains still poorly understood. In vitro analysis conducted on simplified culture systems may offer useful alternatives. In the first part of present study, a novel murine germ cell specific gene, named *Theg*, which was identified as differentially expressed gene in co-culture of primary germ cells (spermatocytes/spermatids) and Sertoli cells is characterised. *Theg* is specifically expressed in spermatids. Its expression is up regulated by some unknown factor/s from Sertoli cells. The gene is approximately 12 kb in size and contains 8 exons and it was localised in chromosome 10 to region B5-C1. The ORF of 376 amino acids encodes for a 42 kDa putative nuclear protein. The cellular type and the subcellular localisation of the THEG protein in the testicular sections of adult mice were determined by immunostaining with a polyclonal antibody against THEG protein. A predominant signal was detected in the nucleus of round spermatids, whereas no specific staining was observed in other germ cell stages and in somatic cells.

To elucidate the function of *Theg* protein and its role in spermatogenesis, we decided to disrupt the *Theg* gene by homologous recombination in mouse. *Theg* was identified as a novel protein therefore for functional analysis of *Theg* protein's domain structure, two different knock-out approach were undertaken. In first knock-out mice the C-terminal of *Theg* protein, which includes exons 3-8, was deleted. Both male and female mice heterozygous for *Theg* deletion appeared normal and fertile. Homozygous male and female mice also exhibited normal phenotype and *Theg*^{-/-} male mice were fertile. Thus suggesting that C-terminal of *Theg* does not play any important role for a successful progression of spermatogenesis. In the second knock-out mice, we deleted the N-terminal domain of *Theg* protein by homologous recombination, which involved deletion of exons 1-4 of *Theg*. Both heterozygous and homozygous male mice were fertile. Further phenotypic and physiological analyses to identify any subtle abnormalities are awaited. However from our results from both knock-out mouse model systems, we can conclude that *Theg* is a non-essential protein for spermatogenesis and perhaps its function is compensated by some other redundant protein/s.

We also describe the molecular cloning and characterisation of the human homologue of murine *Theg*. Expression studies by using dot blot technique revealed that *THEG* is expressed specifically in the testis. The complete genomic structure of *THEG* was identified from

sequence analysis of a genomic *THEG* PAC clone and from the release of the working draft sequence of the human genome in the GeneBank. Furthermore, we assigned the *THEG* gene to human chromosome 19ptel-p13 region by fluorescence in situ hybridisation, which is syntenic to mouse chromosome 10 B5-C1 region. The comparison of the *THEG* and *Theg* sequences showed an overall identity of 60% at the amino acid level. Taken together these results confirm that *THEG* is indeed a true homologue of murine *Theg*.

During the generation of *Theg* knock-out mice, an intriguing but nevertheless very interesting mutation was discovered, which we named as *nax*, which is a novel (autosomal recessive) spontaneous mutation. *nax* mice exhibit complete lack of hair after birth and displays growth retardation during development. However there is a delayed appearance of hair but the posterior part of adult mice remains naked. Strikingly *nax* mice starts to show an ataxic gait around stage P10, they walk around incessantly. Macroscopic inspection showed a pronounced reduction in the brain size. Further histological examination revealed a distorted deep cerebellar nuclei (DCN) and disruption of spatial arrangement of the cerebellar cortex. The granule cells from external granule layer (EGL) fail to migrate inward in a directional manner to form internal granule layer (IGL). To understand the aetiology of cerebellar defects in *nax* mice, we examined postnatal development of *nax/nax* cerebellum. Our results demonstrate that cerebellum defects are apparent at birth; we are currently investigating various prenatal stages to identify the onset of phenotype. Further, we are also investigating subtle histological defects in other parts of brain, abnormalities in skin and other tissues.

For mapping the mutant gene locus, we undertook a genome wide linkage search using microsatellite markers. Analysis of first 42 (affected) intercross (C57BL/6J x 129X1/SvJ) progeny demonstrated a strong linkage to marker D2mit206 in the middle of chromosome 2 with recombination fraction of .05. Further haplotype analysis of 32 recombinant progeny placed the *nax* locus between the marker D2mit158 and D2mit100. We retrieved the sequence between these two markers from Celera database. The locus was more than 14 Mb long and contains about 200 genes. To facilitate fine mapping, we undertook an intensive search within this sequence to identify additional polymorphic microsatellite markers and initiated a linkage dis-equilibrium analysis with 45 affected animals. A haplotype analysis with new markers enabled us to narrow down the locus to 800 kb. At present we have 24 candidate genes, however no mutation was identified, which can be attributed to *nax* phenotype.

References

Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B., Knowles B.B. (1997). The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein. *Nature*. **386**: 838-42.

Adham I.M., Nayernia K., Burkhardt-Gottges E., Topaloglu O., Dixkens C., Holstein A.F., Engel W. (2001). Teratozoospermia in mice lacking the transition protein 2 (*Tnp2*). *Mol Hum Reprod*. **7**: 513-20.

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

Altherr M.R., Bengtsson U., Elder F.F., Ledbetter D.H., Wasmuth J.J., McDonald M.E., Gusella J.F., Greenberg F. (1991). Molecular confirmation of Wolf-Hirschhornsyndrome with a subtle translocation of chromosome 4. *Am J Hum Genet*. **49**: 1235-42.

Altman J., Bayer S.A. (1997). The development of the cerebellar system: In relation to its evolution, structure and functions. pp: 109-52 (CacheRiver Press, Clearwater).

Alvarez-Otero R., Sotelo C., Alvarado-Mallart R.M. (1993). Chick/quail chimeras with partial cerebellar grafts: an analysis of the origin and migration of cerebellar cells. *J Comp Neurol*. **333**: 597-615.

Anton E.S., Cameron R.S., Rakic P. (1996). Role of neuron-glia junctional domain proteins in the maintenance and termination of neuronal migration across the embryonic cerebral wall. *J Neurosci*. **16**: 2283-93.

Aoki A., Fawcett D.W. (1975). Impermeability of Sertoli cell junctions to prolonged exposure to peroxidase. *Andrologia*. **7** : 63-76.

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

References

Baba T., Azuma S., Kashiwabara S., Toyoda Y. (1994). Sperm from mice carrying a targeted mutation of the *acrosin* gene can penetrate the oocyte zona pellucida and effect fertilization. *J Biol Chem.* **269**: 31845-9.

Balhorn R., Weston S., Thomas C. (1984). DNA packaging in mouse spermatids. Synthesis of protamine variant and four transition proteins. *Exp Cell Res.* **150**: 298-308.

Baskaran R., Rao M.R.S. (1991). Mammalian spermatid specific protein, TP2, is a zinc metalloprotein with two finger motifs. *Biochem Biophys Res Commun.* **179**: 1491-99.

Bellve A.R., Zheng W. (1989). Growth factors as autocrine and paracrine modulators of male gonadal functions. *J Reprod Fertil.* **85**: 771-93.

Ben-Arie N., Bellen H.J., Armstrong D.L., McCall A.E., Gordadze P.R., Guo Q., Matzuk M.M., Zoghbi H.Y. (1997). *Math1* is essential for genesis of cerebellar granule neurons. *Nature.* **390**: 169-72.

Birnboim H.C., Doly J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA, *Nucleic Acids Res.* **7**: 1513-23.

Bradford M.M. (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.

Burfeind P., Hoyer-Fender S. (1991). Sequence and developmental expression of an mRNA encoding a putative protein of rat sperm outer dense fibers. *Dev Biol.* **148**: 195–204.

Cameron R.S., Rakic P. (1994). Identification of membrane proteins that comprise the plasmalemmal junction between migrating neurons and radial glial cells. *J Neurosci.* **14**: 3139-55.

References

- Cameron R.S., Ruffin J.W., Cho N.K., Cameron P.L., Rakic P. (1997). Developmental expression, pattern of distribution, and effect on cell aggregation implicate a neuron-glia junctional domain protein in neuronal migration. *J Comp Neurol.* **387**: 467-88.
- Caron N., Veilleux S., Boissonneault G. (2001). Stimulation of DNA repair by the spermatid TP1 protein. *Mol Reprod Dev.* **58**: 437-43.
- Chae T., Kwon Y.T., Bronson R., Dikkes P., Li E., Tsai L.H. (1997). Mice lacking p35, a neuronal specific activator of *Cdk5*, display cortical lamination defects, seizures, and adult lethality. *Neuron.* **18**: 29-42.
- Cheng C.Y., Mather J.P., Byer A.L., Bardin C.W. (1986). Identification of hormonally responsive proteins in primary Sertoli cell culture medium by anion-exchange high performance liquid chromatography. *Endocrinology.* **118**: 480-8.
- Chien A., Edgar D.B., Trela J.M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol.* **127**: 1550-57.
- Clark J.M. (1988). Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16**: 9677-86.
- Costa G.L., Weiner M.P. (1994). Polishing with T4 or *Pfu* polymerase increases the efficiency of cloning of PCR fragments. *Nucleic Acids Res.* **22**: 2423.
- Cooke H.J., Saunders P.T. (2002). Mouse models of male infertility. *Nat Rev Genet.* **3**: 790-801.
- Crossley P.H., Martin G.R. (1995). The mouse *Fgf8* gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development.* **121**: 439-51.
- Dagert M., Ehrlich S.D. (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene.* **6**: 23-8.

References

D'Agostino A., Monaco L., Stefanini M., Geremia R. (1984). Study of the interaction between germ cells and Sertoli cells in vitro. *Exp Cell Res.* **150**: 430-5.

D'Arcangelo G., Miao G.G., Chen S.C., Soares H.D., Morgan J.I., Curran T. (1995). A protein related to extracellular matrix proteins deleted in the mouse mutant reeler. *Nature.* **374**: 719-23.

D'Arcangelo G., Curran T. (1998). Reeler: new tales on an old mutant mouse. *Bioessays.* **20**: 235-44.

de Kretser D.M., Kerr J.B. (1988). The cytology of the testis. In Knobil E., Neil J. (eds): The physiology of reproduction. pp: 837 (Raven Press, New York).

Delaney C.L., Brenner M., Messing A. (1996). Conditional ablation of cerebellar astrocytes in postnatal transgenic mice. *J Neurosci.* **16**: 6908-18.

Deryckere F., Gannon F.A. (1994). One-hour miniprep technique for extraction of DNA-binding proteins from animal tissues. *Biotechniques.* **16**: 405.

Denhardt D.T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochem Biophys Res Commun.* **23**: 641-46.

DePhilip R.M., Danahey D.G. (1987). Germ cell binding to rat Sertoli cells in vitro. *Biol Reprod.* **37**: 1271-82.

Domenjoud L., Fronia C., Uhde F., Engel W. (1988). Sequence of human protamine 2 cDNA. *Nucl Acids Res.* **16**: 7733.

Dym M., Fawcett D.W. (1975). The blood-testis barrier in the rat and the physiological compartmentation of the seminiferous epithelium. *Biol Reprod.* **3**: 308-26.

Enders G.C., Millette C.F. (1988). Pachytene spermatocyte and round spermatid binding to Sertoli cells in vitro. *J Cell Sci.* **90**: 105-14.

References

Evans M.J., Kaufman M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*. **292**: 154-6.

Fantz D.A., Hatfield W.R., Horvath G., Kistler M.K., Kistler W.S. (2001). Mice with a targeted disruption of the *Hlt* gene are fertile and undergo normal changes in structural chromosomal proteins during spermiogenesis. *Biol Reprod*. **64**: 425-31.

Favor J., Sandulache R., Neuhauser-Klaus A., Pretsch W., Chatterjee B., Senft E., Wurst W., Blanquet V., Grimes P., Sporle R., Schughart K. (1996). The mouse *Pax2* (1Neu) mutation is identical to a human *PAX2* mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney. *Proc Natl Acad Sci U S A*. **93**: 13870-5.

Feinberg A.P., Vogelstein B. (1989). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem*. **123**: 6-13.

Feng L., Hatten M.E., Heintz N. (1994). Brain lipid-binding protein (BLBP): a novel signaling system in the developing mammalian CNS. *Neuron*. **12**: 895-908.

Feng L., Heintz N. (1995). Differentiating neurons activate transcription of the brain lipid-binding protein gene in radial glia through a novel regulatory element. *Development*. **121**: 1719-30.

Fritz I.B., Tung P.S. (1986). Role of interactions between peritubular cells and Sertoli cells in mammalian testicular function. In Gall J. (ed): Gametogenesis and early embryo. pp: 151 (Alan R Liss, New York).

Fritz I.B., Tung P.S., Ailenberg M. (1993). Proteases and antiproteases in the seminiferous tubules. In Russell L.D., Griswold M.D. (eds): The Sertoli Cell. pp: 217 (CacheRiver Press, Clearwater).

References

- Gastmann O., Burfeind P., Gunther E., Hameister H., Szpirer C., Hoyer-Fender S. (1993). Sequence, expression, and chromosomal assignment of a human sperm outer dense fiber gene. *Mol Reprod Dev.* **36**: 407–418.
- Ghez C., Thach N.T. (2000). The cerebellum. In Kandel E.R., Schwartz J.H., Jessell T. M. (eds): Principles of neural Science. pp: 832-67 (McGraw-Hill Companies, NewYork).
- Goffinet A.M., So K.F., Yamamoto M., Edwards M., Caviness V.S. Jr. (1984). Architectonic and hodological organization of the cerebellum in *reeler* mutant mice. *Brain Res.* **318**: 263-76.
- Goldowitz D. (1989). Cell allocation in mammalian CNS formation: evidence from murine interspecies aggregation chimeras. *Neuron.* **3**: 705-13.
- Goldowitz D., Cushing R.C., Laywell E., D'Arcangelo G., Sheldon M., Sweet H.O., Davisson M., Steindler D., Curran T. (1997). Cerebellar disorganization characteristic of *reeler* in scrambler mutant mice despite presence of reelin. *J Neurosci.* **17**: 8767-77.
- Goldowitz D., Hamre K. (1998). The cells and molecules that make a cerebellum. *Trends Neurosci.* **21**: 375-82.
- Gershoni J.M., Palade G.E. (1982). Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to a positively charged membrane filter. *Anal Biochem.* **124**: 396-405.
- Griswold M.D. (1993). Unique aspects of the biochemistry and metabolism of Sertoli cells. In Russell L.D., Griswold M.D. (eds): The Sertoli Cell. pp: 485 (CacheRiver Press, Clearwater).
- Grootegeod J.A., Jansen R., Van der Molen H.J. (1984). The role of glucose, pyruvate and lactate in ATP production by rat spermatocytes and spermatids. *Biochim Biophys Acta.* **767**: 248-56.
- Grove B.D., Pfeiffer D.C., Allen S., Vogl A.W. (1990). Immunofluorescence localization of *vinculin* in ectoplasmic ("junctional") specializations of rat Sertoli cells. *Am J Anat.* **188**: 44-56.

References

Hagenas L., Ploen L., Ritzen E.M., Ekwall H. (1977). Blood-testis barrier: maintained function of inter-Sertoli cell junctions in experimental cryptorchidism in the rat, as judged by a simple lanthanum-immersion technique. *Andrologia*. **9**: 250-4.

Hallonet M., Alvarado-Mallart R.M. (1997). The chick/quail chimeric system: a model for early cerebellar development. *Perspect Dev Neurobiol*. **5**: 17-31.

Hallonet M.E., Teillet M.A., Le Douarin N.M. (1990). A new approach to the development of the cerebellum provided by the quail-chick marker system. *Development*. **108**: 19-31.

Hamre K.M., Goldowitz D. (1997). *meander tail* acts intrinsic to granule cell precursors to disrupt cerebellar development: analysis of meander tail chimeric mice. *Development*. **124**: 4201-12.

Hicks G.R., Raikhel N.V. (1995). Protein import into the nucleus: an integrated view. *A Rev Cell Dev Biol*. **11**: 155-58.

Hodge R. (1994). Preparation of RNA gel blots. *Methods Mol Biol*. **28**: 49–54.

Howell B.W., Hawkes R., Soriano P., Cooper J.A. (1997). Neuronal position in the developing brain is regulated by mouse *disabled-1*. *Nature*. **389**: 733-7.

Hu G. (1993). DNA polymerase-catalyzed addition of non-templated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol*. **12**: 763–770.

Ioannou P.A., Amemiya C.T., Garnes J., Kroisel P.M., Shizuya H., Chen C., Batzer M.A., de Jong A.J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet*. **6**: 84–89

Joyner A.L. (2000). Gene Targeting, 2nd Edn., A Practical Approach. pp: 138 (Oxford University Press, New York).

References

Joyner A.L. (1996). *Engrailed*, *Wnt* and *Pax* genes regulate midbrain-hindbrain development. *Trends Genet.* **12**: 15-20.

Jutte N.H., Grootegoed J.A., Rommerts F.F., van der Molen H.J. (1981). Exogenous lactate is essential for metabolic activities in isolated rat spermatocytes and spermatids. *J Reprod Fertil.* **62**: 399-405.

Jutte N.H., Jansen R., Grootegoed J.A., Rommerts F.F., van der Molen H.J. (1983). FSH stimulation of the production of pyruvate and lactate by rat Sertoli cells may be involved in hormonal regulation of spermatogenesis. *J Reprod Fertil.* **68**: 219-26.

Kierszenbaum A.L. (2001). Transition nuclear proteins during spermiogenesis: unrepaired DNA breaks not allowed. *Mol Reprod Dev.* **58**: 357-8

Kikuchi Y., Segawa H., Tokumoto M., Tsubokawa T., Hotta Y., Uyemura K., Okamoto H. (1997). Ocular and cerebellar defects in zebrafish induced by overexpression of the LIM domains of the islet-3 LIM/homeodomain protein. *Neuron.* **18**: 369-82.

Kissinger C., Skinner M.K., Griswold M.D. (1982). Analysis of Sertoli cell-secreted proteins by two-dimensional gel electrophoresis. *Biol Reprod.* **27**: 233-40.

Kistler W.S., Henriksen K., Mali P. (1996). Sequential expression of nucleoproteins during rat spermatogenesis. *Exp Cell Res.* **225**: 374-81.

Kleene K.C., Flynn J.F. (1987). Characterization of a cDNA clone encoding a basic protein, TP2, involved in chromatin condensation during spermiogenesis in the mouse. *J Biol Chem.* **262**: 17272-77.

Koehler M.R., Bosserhoff A., von Beust G., Bauer A., Blesch A., Buettner R., Schlegel J., Bogdahn U., Schmid M. (1996). Assignment of the human melanoma inhibitory activity gene (MIA) to 19q13.32-q13.33 by fluorescence in situ hybridization (FISH). *Genomics.* **35**: 265-67.

References

Komuro H., Rakic P. (1993). Modulation of neuronal migration by NMDA receptors. *Science*. **260**: 95-7.

Komuro H., Rakic P. (1996). Intracellular Ca^{2+} fluctuations modulate the rate of neuronal migration. *Neuron*. **17**: 275-85.

Kozak M. (1987). An analysis of 5' noncoding sequences from 699 vertebrate messenger RNAs. *Nucl Acids Res*. **15**: 8125-48.

Kubota H., Hynes G., Willison K. (1995). The chaperonin containing t-complex polypeptide 1 (TCP-1). Multisubunit machinery assisting in protein folding and assembly in the eukaryotic cytosol. *Eur J Biochem*. **230**: 3-16.

Kuwano A., Ledbetter S.A., Dobyns W.B., Emanuel B.S., Ledbetter D.H. (1991). Detection of deletions and cryptic translocations in Miller-Dieker syndrome by in situ hybridization. *Am J Hum Genet*. **49**: 707-14.

Lacroix M., Parvinen M., Fritz I.B. (1981). Localization of testicular plasminogen activator in discrete portions (stage VII and VIII) of the seminiferous tubule. *Biol Reprod*. **25**: 143-6.

Laemmli U.K. (1970). Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature*. **227**: 680-85.

Laird P.W., Zijderveld A., Linders K., Rudnicki M.A., Jaenisch R., Berns A. (1991). Simplified mammalian DNA isolation procedure. *Nucleic Acids Res*. **19**: 4293.

Lamb J., Wilkie A.O., Harris P.C., Buckle V.J., Lindenbaum R.H., Barton N.J., Reeders S.T., Weatherall D.J., Higgs D.R. (1989). Detection of breakpoints in submicroscopic chromosomal translocation, illustrating an important mechanism for genetic disease. *Lancet*. **2**: 819-24.

Lambert de Rouvroit C., Goffinet A.M. (2001). Neuronal migration. *Mech Dev*. **105**: 47-56.

References

Leblond C.P., Clermont Y. (1952). Definition of the stages of the cycle of the seminiferous epithelium in the rat. *Ann NY Acad Sci.* **55**: 548.

Lee C.H., Hoyer-Fender S., Engel W. (1987). The nucleotide sequence of human protamine 1 cDNA. *Nucl Acids Res.* **15**: 7639.

Liang P., Pardee A.B. (1992). Differential display of eucaryotic messenger RNA by means of the polymerase chain reaction. *Science.* **257**: 967-71.

Lichter P., Cremer T., Borden J., Manuelidis L., Ward D.C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum Genet.* **80**: 224-34

Lilienbaum A., Sage J., Memet S., Rassoulzadegan M., Cuzin F., Israel A. (2000). NF-kappa B is developmentally regulated during spermatogenesis in mice. *Dev Dyn.* **219**: 333-40.

Liu M., Pleasure S.J., Collins A.E., Noebels J.L., Naya F.J., Tsai M.J., Lowenstein D.H. (2000). Loss of BETA2/NeuroD leads to malformation of the dentate gyrus and epilepsy. *Proc Natl Acad Sci U S A.* **97**: 865-70.

Luerssen H., Hoyer-Fender S., Engel W. (1988). The nucleotide sequence of human transition protein 1 cDNA. *Nucl Acids Res.* **16**: 7723.

Luerssen H., Maier W.M., Hoyer-Fender S., Engel W. (1989). The nucleotide sequence of rat transition protein 2 (TP2) cDNA. *Nucl Acids Res.* **17**: 3585.

Marin F., Puelles L. (1994). Patterning of the embryonic avian midbrain after experimental inversions: a polarizing activity from the isthmus. *Dev Biol.* **163**: 19-37.

Martin G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA.* **78**: 763 4-8.

References

Martinez S., Alvarado-Mallart R.M. (1989). Transplanted mesencephalic quail cells colonize selectively all primary visual nuclei of chick diencephalon: a study using heterotopic transplants. *Brain Res Dev.* **47**: 263-74.

Martinez S., Alvarado-Mallart R.M., Martinez-de-la-Torre M., Puelles L. (1991). Retinal and tectal connections of embryonic nucleus superficialis magnocellularis and its mature derivatives in the chick. *Anat Embryol (Berl)*. **183**: 235-43.

Martinez S., Marin F., Nieto M.A., Puelles L. (1995). Induction of ectopic *engrailed* expression and fate change in avian rhombomeres: intersegmental boundaries as barriers. *Mech Dev.* **51**: 289-303.

Mathis L., Bonnerot C., Puelles L., Nicolas J.F. (1997). Retrospective clonal analysis of the cerebellum using genetic *lacZ/lacZ* mouse mosaics. *Development.* **124**: 4089-104.

McGinley D.M., Posalaky Z., Porvaznik M., Russell L. (1979). Gap junctions between Sertoli and germ cells of rat seminiferous tubules. *Tissue Cell.* **11**: 741-54.

McMahon A.P., Bradley A. (1990). The *Wnt-1 (int-1)* proto-oncogene is required for development of a large region of the mouse brain. *Cell.* **62**: 1073-85.

McMahon A.P., Joyner A.L., Bradley A., McMahon J.A. (1992). The midbrain-hindbrain phenotype of *Wnt-1/Wnt-1* mice results from stepwise deletion of engrailed-expressing cells by 9.5 days postcoitum. *Cell.* **69**: 581-95.

Meyers E.N., Lewandoski M., Martin G.R. (1998). An *Fgf8* mutant allelic series generated by Cre- and Flp-mediated recombination. *Nat Genet.* **18**: 136-41.

Mita M., Hall P.F. (1982). Metabolism of round spermatids from rats: lactate as the preferred substrate. *Biol Reprod.* **26**: 445-55.

References

Miyata T., Nakajima K., Mikoshiba K., Ogawa M. (1997). Regulation of Purkinje cell alignment by reelin as revealed with CR-50 antibody. *J Neurosci.* **17**: 3599-609.

Nayernia K., Adham I.M., Burkhardt-Gottges E., Neesen J., Rieche M., Wolf S., Sancken U., Kleene K., Engel W. (2002). Asthenozoospermia in mice with targeted deletion of the sperm mitochondrion-associated cysteine-rich protein (*Smcp*) gene. *Mol Cell Biol.* **22**: 3046-52.

Nayernia K., von Mering M.H., Kraszucka K., Burfeind P., Wehrend A., Kohler M., Schmid M., Engel W. (1999). A novel testicular haploid expressed gene (THEG) involved in mouse spermatid-sertoli cell interaction. *Biol Reprod.* **60**: 1488-95.

Neaves W.B. (1973). Permeability of Sertoli cell tight junctions to lanthanum after ligation of ductus deferens and ductuliefferentes. *J Cell Biol.* **593**: 559-72.

Newton S.C., Millette C.F. (1992). Sertoli cell plasma membrane polypeptides involved in spermatogenic cell-Sertoli cell adhesion. *J Androl.* **13**: 160-71.

Ning Y., Flint J. (1996). A complete set of human telomeric probes and their clinical application. *Nature Genet.* **14**: 86–89.

Ohshima T., Ward J.M., Huh C.G., Longenecker G., Veeranna I., Pant H.C., Brady R.O., Martin L.J., Kulkarni A.B. (1996). Targeted disruption of the *cyclin-dependent kinase 5* gene results in abnormal corticogenesis, neuronal pathology and perinatal death. *Proc Natl Acad Sci U S A.* **93**: 11173-8.

Overhauser J., Bengtsson U., McMahon J., Ulm J., Butler M.G., Santiago L., Wasmuth J.J. (1989). Prenatal diagnosis and carrier detection of a cryptic translocation by using DNA markers from the short arm of chromosome 5. *Am J Hum Genet.* **45**: 296–303.

Oyen O., Myklebust F., Scott J.D., Cadd G.G., McKnight G.S., Hansson V., Jahnsen T. (1990). Subunits of cyclic adenosine 3',5'-monophosphate-dependent protein kinase show differential

References

and distinct expression patterns during germ cell differentiation: alternative polyadenylation in germ cells gives rise to unique smaller-sized mRNA species. *Biol Reprod.* **43**: 46–54.

Palombi F., Ziparo E., Rommerts F.F., Grootegoed J.A., Antonini M., Stefanini M. (1979). Morphological characteristics of male germ cells of rats in contact with Sertoli cells in vitro. *J Reprod Fertil.* **57**: 325-30.

Parvinen M. (1982). Regulation of the seminiferous epithelium. *Endocr Rev.* **3**: 404-17.

Patil N., Cox D.R., Bhat D., Faham M., Myers R.M., Peterson A.S. (1995). A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nat Genet.* **11**: 126-9.

Pelletier R.M. (1988). Cyclic modulation of Sertoli cell junctional complexes in a seasonal breeder: the mink (*Mustela vison*). *Am J Anat.* **183**: 68-102.

Phillips D.M., Pilder S.H., Olds-Clarke P.J., Silver L.M. (1993). Factors that may regulate assembly of the mammalian sperm tail deduced from a mouse t complex mutation. *Biol Reprod.* **49**: 1347-52.

Rakic P., Sidman R.L. (1973). Organization of cerebellar cortex secondary to deficit of granule cells in weaver mutant mice. *J Comp Neurol.* **152**: 133-61.

Ramo'n y Cajal S. (1911). *Histologie du syteme nerveux de l'homme et des vertebres*. Oxford University Press.

Rassoulzadegan M., Paquis-Flucklinger V., Bertino B., Sage J., Jasin M., Miyagawa K., van Heyningen V., Besmer P., Cuzin F. (1993). Transmeiotic differentiation of male germ cells in culture. *Cell.* **75**: 997-1006.

Rio C., Rieff H.I., Qi P., Khurana T.S., Corfas G. (1997). Neuregulin and erbB receptors play a critical role in neuronal migration. *Neuron.* **19**: 39-50.

References

Romrell L.J., Belve A.R., Fawcett D.W. (1976). Separation of mouse spermatogenic cells by sedimentation velocity. *Dev Biol.* **49**: 119-31.

Ross M.E., Fletcher C., Mason C.A., Hatten M.E., Heintz N. (1990). *Meander tail* reveals a discrete developmental unit in the mouse cerebellum. *Proc Natl Acad Sci U S A.* **87**: 4189-92.

Ross M.I. (1975). The Sertoli cell and the blood testis barrier: an electron microscopical study. In Holstein A.F., Horstmann E. (eds): *Morphological aspects of andrology.* pp: 83 (Grosse, Berlin).

Rossi D.J., Alford S., Mugnaini E., Slater N.T. (1995). Properties of transmission at a giant glutamatergic synapse in cerebellum: The mossy fiber-unipolar brush cell synapse. *J.Neurophysiol.* **74**: 24-42.

Russel L.D. (1993). Morphological and functional evidence for Sertoli-germ cell relationship. In Russell L.D., Griswold M.D. (eds): *The Sertoli Cell.* pp: 365–90 (CacheRiver Press, Clearwater).

Russell L.D. (1977). Observations on rat Sertoli ectoplasmic ('junctional') specializations in their association with germ cells of the rat testis. *Tissue Cell.* **9**: 475.

Russell L.D. (1979a). Further observations on tubulobulbar complexes formed by late spermatids and Sertoli cells in the rat testis. *Anat Rec.* **194**: 213-32.

Russell L.D. (1979b). Spermatid-Sertoli tubulobulbar complexes as devices for elimination of cytoplasm from the head region late spermatids of the rat. *Anat Rec.* **194**: 233-46.

Russell L.D., Peterson R.N. (1985). Sertoli cell junctions: morphological and functional correlates. *Int Rev Cytol.* **94**: 177-211.

Russell L.D., Tallon-Doran M., Weber J.E., Wong V., Peterson R.N. (1983). Three-dimensional reconstruction of a rat stage V Sertoli cell: III. A study of specific cellular relationships. *Am J Anat.* **167**: 181-92.

References

Ryder E.F., Cepko C.L. (1994). Migration patterns of clonally related granule cells and their progenitors in the developing chick cerebellum. *Neuron*. **12**: 1011-28.

Saccone S., DeSario A., Della-Valle G., Bernardi G. (1992). The highest gene concentrations in the human genome are in telomeric bands of metaphase chromosomes. *Proc natl Acad Sci USA*. **89**: 4913-17.

Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., Erlich H.A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science*. **239**: 487-91.

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

Sanger F., Nicklen S., Coulson A.R. (1977). DNA sequencing with the chain terminating inhibitors, *Proc Natl Acad Sci U S A*. **74**: 5463-67.

Schlüter G., and Engel W. (1995). The rat *Prm3* gene is an intronless member of the protamine gene cluster and is expressed in haploid male germ cells. *Cytogenet Cell Genet*. **71**: 352-55.

Schlüter G., Celik A., Obata R. (1996). Sequence analysis of conserved protamine gene cluster shows that it contains a fourth expressed gene. *Mol Reprod Dev*. **43**: 1-6.

Schuler G.D. (1998). Electronic PCR: bridging the gap between genome mapping and genome sequencing. *Trends Biotech*. **16**: 456-59.

Sertoli E. (1865). E lesistenza di particolari cellule ramificate nei canalicoli seminiferi dell'testicolo umno. *Morgagni*. **7**: 31.

Setchell B.P. (1986). The movement of fluids and substances in the testis. *Aust J Biol Sci*. **39**: 193-207.

References

Setchell B.P., Waites G.B.H. (1986). The blood-testis barrier. In Hamilton D.W., Greep R.O. (eds): Handbook of physiology. pp: 143 (Williams and Wilkins Press, Baltimore).

Shabanowitz R.B., DePhilip R.M., Crowell J.A., Tres L.L., Kierszenbaum A.L. (1986). Temporal appearance and cyclic behavior of Sertoli cell-specific secretory proteins during the development of the rat seminiferous tubule. *Biol Reprod.* **35**: 745-60.

Sharpe R.M. (1993). Experimental evidence for Sertoli cell-gem cell and Sertoli-Leydig cell interactions. In Russell L.D., Griswold M.D. (eds): The Sertoli Cell. pp: 391 (CacheRiver Press, Clearwater).

Sheldon M., Rice D.S., D'Arcangelo G., Yoneshima H., Nakajima K., Mikoshiba K., Howell B.W., Cooper J.A., Goldowitz D., Curran T. (1997). Scrambler and yotari disrupt the disabled gene and produce a reeler-like phenotype in mice. *Nature.* **389**: 730-3.

Sidman R.L., Green M.C., Appel S.H. (1965). Catalogue of the neurological mutants of the mouse. pp 22-206 (Harvard University Press, Boston).

Silver L.M. (1985). Mouse haplotypes. *Annu Rev Genet.* **19**: 179-208.

Silver L.M. (1995). Mouse Genetics: concepts and application. pp: 101 (Oxford University Press, NewYork).

Silver L.M., Kleene K.C., Distel R.J., Hecht N.B. (1987). Synthesis of mouse t complex proteins during haploid stages of spermatogenesis. *Dev Biol.* **119**: 605-8.

Skinner M.K. (1991). Cell-cell interactions in the testis. *Endocr Rev.* **12**: 45-77.

Skinner M.K. (1993). Sertoli cell-peritubular myoid cell interactions. In Russell L.D., Griswold M.D. (eds): The Sertoli Cell. pp: 477 (CacheRiver Press, Clearwater).

References

Skinner M.K., McLachlan R.I., Bremner W.J. (1989). Stimulation of Sertoli cell inhibin secretion by the testicular paracrine factor PmodS. *Mol Cell Endocrinol.* **66**: 239-49.

Smeyne R.J., Goldowitz D. (1989). Development and death of external granular layer cells in the weaver mouse cerebellum: a quantitative study. *J Neurosci.* **9**: 1608-20.

Song D.L., Chalepakis G., Gruss P., Joyner A.L. (1996). Two Pax-binding sites are required for early embryonic brain expression of an *Engrailed-2* transgene. *Development.* **122**: 627-35.

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

Sternlicht H., Farr G.W., Sternlicht M.L., Driscoll J.K., Willison K., Yaffe M.B. (1993). The t-complex polypeptide 1 complex is a chaperonin for tubulin and actin in vivo. *Proc Natl Acad Sci U S A.* **90**: 9422-6.

Syed V., Hecht N. (1997). Up-regulation and down-regulation of genes expressed in co-culture of rat Sertoli cells and germ cells. *Mol Reprod Dev.* **47**: 380-89.

Sylvester S.R. (1993). Secretion of transport and binding proteins. In Russell L.D., Griswold M.D. (eds): *The Sertoli Cell.* pp: 201 (CacheRiver Press, Clearwater).

Thomas K.R., Capecchi M.R. (1990). Targeted disruption of the murine *int-1* proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. *Nature.* **346**: 847-50.

Tres L.L., Kierszenbaum A.L. (1983). Viability of rat spermatogenic cells in vitro is facilitated by their coculture with Sertoli cells in serum-free hormone-supplemented medium. *Proc Natl Acad Sci U S A.* **80**: 3377-81.

Tsurula J.K., O'Brien D.A. (1995). Sertoli cell-spermatogenic cell interaction: the insulin-like growth factor II/cationindependent mannose 6-phosphate receptor mediates changes in spermatogenic cell gene expression in mice. *Biol Reprod.* **53**: 1454-64.

References

Urbanek P., Fetka I., Meisler M.H., Busslinger M. (1997). Cooperation of *Pax2* and *Pax5* in midbrain and cerebellum development. *Proc Natl Acad Sci U S A.* **94**: 5703-8.

Vidal F., Lopez P., Lopez-Fernandez L.A., Ranc F., Scimeca J.C., Cuzin F., Rassoulzadegan M. (2001). Gene trap analysis of germ cell signaling to Sertoli cells: NGF-TrkA mediated induction of *Fra1* and *Fos* by post-meiotic germ cells. *J Cell Sci.* **114**: 435-43.

Vincent S., Segretain D., Nishikawa S., Nishikawa S.I., Sage J., Cuzin F., Rassoulzadegan M. (1998). Stage-specific expression of the Kit receptor and its ligand (KL) during male gametogenesis in the mouse: a Kit-KL interaction critical for meiosis. *Development.* **125**: 4585-93.

Vogelstein B., Gillespie D. (1979). Preparative and analytical purification of DNA from agarose, *Proc Natl Acad Sci U S A.* **76**: 615-19.

Ware M.L., Fox J.W., Gonzalez J.L., Davis N.M., Lambert de Rouvroit C., Russo C.J., Chua S.C. Jr., Goffinet A.M., Walsh C.A. (1997). Aberrant splicing of a mouse disabled homolog, *mdab1*, in the scrambler mouse. *Neuron.* **19**: 239-49.

Wassarman K.M., Lewandoski M., Campbell K., Joyner A.L., Rubenstein J.L., Martinez S., Martin G.R. (1997). Specification of the anterior hindbrain and establishment of a normal mid/hindbrain organizer is dependent on *Gbx2* gene function. *Development.* **124**: 2923-34.

Wilson R.M., Griswold M.D. (1979). Secreted proteins from rat Sertoli cells. *Exp Cell Res.* **123**: 127-35.

Wolfsberg T. G., Primakoff P., Myles D. G. and White J. M. (1995a). ADAM, a novel family of membrane proteins containing a disintegrin and metalloprotease domain: multipotential functions in cell-cell and cell-matrix interactions. *J Cell Biol.* **131**: 275-78.

Wolfsberg T. G., Straight P. D., Gerena R. L., Huovila, A. P. Primakoff P., Myles D. G. and White J. M. (1995b). ADAM, a widely distributed and developmentally regulated gene family

References

encoding membrane proteins with a disintegrin and metalloprotease domain. *Dev Biol.* **169**: 378-83.

Wright W.W., Parvinen M., Musto N.A., Gunsalus G.L., Phillips D.M., Mather J.P., Bardin C.W. (1983). Identification of stage-specific proteins synthesized by rat seminiferous tubules. *Biol Reprod.* **29**: 257-70.

Wurst W., Auerbach A.B., Joyner A.L. (1994). Multiple developmental defects in Engrailed-1 mutant mice: an early mid-hindbrain deletion and patterning defects in forelimbs and sternum. *Development.* **120**: 2065-75.

Yaffe M.B., Farr G.W., Miklos D., Horwich A.L., Sternlicht M.L., Sternlicht H. (1992). TCP1 complex is a molecular chaperone in tubulin biogenesis. *Nature.* **358**: 245-8.

Yanaka N., Kobayashi K., Wakimoto K., Yamada E., Imahie H., Imai Y., Mori C. (2000). Insertional mutation of the *kisimo* locus caused a defect in spermatogenesis. *J Biol Chem.* **275**: 14791-97.

Zhang L., Goldman J.E. (1996). Developmental fates and migratory pathways of dividing progenitors in the postnatal rat cerebellum. *J Comp Neurol.* **370**: 536-50.

Zhang L., Goldman J.E. (1996). Generation of cerebellar interneurons from dividing progenitors in white matter. *Neuron.* **16**: 47-54.

Zhang N., Copeland N.G., Gilbert D.J., Jenkins N.A., Gridley T. (2000). Cloning, expression, and chromosomal localization of a mouse gene homologous to the germ cell migration regulator *wunen* and two type 2 phosphatidic acid phosphatases. *Genomics.* **63**: 142-44.

Zhao M., Shirley C.R., Yu Y.E., Mohapatra B., Zhang Y., Unni E., Deng J.M., Arango N.A., Terry N.H., Weil M.M., Russell L.D., Behringer R.R., Meistrich M.L. (2001). Targeted disruption of the transition protein 2 gene affects sperm chromatin structure and reduces fertility in mice. *Mol Cell Biol.* **21**: 7243-55.

References

Zheng C., Heintz N., Hatten M.E. (1996). CNS gene encoding astrotactin, which supports neuronal migration along glial fibers. *Science*. **272**: 417-9.

Publication

[Mannan A.](#), Lucke K., Dixkens C., Neesen J., Kamper M., Engel W., Burfeind P. (2000). Alternative splicing, chromosome assignment and subcellular localisation of the testicular haploid expressed gene (*THEG*). *Cytogenet Cell Genet.* **91**: 171-9.

Acknowledgements

In the course of the past three years, many people have contributed to my work in numerous ways and I would like to mention them in chronological order.

Firstly I would like to express my deep sense of gratitude to Prof. Dr. W. Engel. He has created an extraordinary and stimulating research environment at the Institute of Human Genetics. I cannot imagine another Doktorvater who would have permitted me to work as independently as he did and would have encouraged outside contacts and collaborations so generously, not to speak of numerous national and international meetings I was able to attend. I am also indebted to him for editing my thesis, despite having a very busy schedule. It has been a pleasure to work with Dr. Karim Nayernia. During the course of time in the Institute, we had some inspiring discussions. I would also like to thank him for being an exacting proofreader, who kept embarrassing mistakes in my thesis away from seeing the light of day.

I would like to thank all my institute colleagues for being so friendly and helpful during my stay. My special thanks to Maneula, Michal, Jin, Ilona, Tarvo and Dr. Peter Burfiend.

I cherished the warm atmosphere of the Lab 107, where everybody is always in cheerful mood and working environment is fantastic, you can't ask for a better place to work. It has been a pleasure to have Ewelina, Lukas, Manyu, Ania, Christian, Iris, Marita and Ksenija and former members Murat, Pawel, Rashimi, Zoryana, Michaela, and Tomek as my colleagues.

I am indebted to Graduate program, "Molekulare Genetik der Entwicklung" for a generous stipend that made life in Göttingen affordable. From all the academic activities organised by graduate committee, I learned a lot in the course of my study. My special thanks to Prof. Dr. T. Pieler for his continuous encouragement.

I would like to thank Dr. S. Hoyer-Fender for being my second referee and to Prof. Dr. D. Gradmann and to Prof. Dr. O. Götze for being my dissertation examiner.

I would also like to thank Dr. Riedesel and Max Planck-Institut für Experimentelle Medizin, Göttingen for their technical assistance in generation of knock-out mice.

Acknowledgements

Many thanks to Prof. Dr. K. Krieglstein from Zentrum Anatomie, Abt. Anatomie mit Schwerpunkt Neuroanatomie, Göttingen for performing histological studies on the cerebellum of *nax* mice

Beyond Göttingen, the collaboration with Prof. Dr. A. Reis from Institute of Human Genetic, Erlangen, has been very stimulating. I am particularly indebted to him for allowing me to do my linkage analysis in his Institute and for the interest he showed in my project. Thanks are due to Conny, Martin, Kathrin, Arif, Alxendaro, and Andrea for being helpful and inspiring colleagues in Erlangen.

My parents, my brothers and my chuchi have contributed to the genesis of this work through their loving support in more wonderful ways than a page can hold. I dedicate this thesis to them.

Resume

Personals

Name : Ashraf-ul Mannan
Date of Birth : Oct 31, 1973
Sex : Male
Marital Status : Single
Nationality : Indian
Correspondence address : Theoder-Heuss-Straße 11, 111
37075, Göttingen, Germany.

Academics

1980-1990 : All India Secondary Examination,
Central Board of Secondary Education. (C.B.S.E)
New Delhi, India

1990-1992 : All India Senior Secondary Examination,
Central Board of Secondary Education. (C.B.S.E)
New Delhi, India

1992-1995 : Bachelor of Science (B.Sc), majored in Biochemistry
University of Delhi, New Delhi, India

1995-1998 : Master of Science (M.S) majored in Molecular Biology
Division of Biological Science, Indian Institute of Science
Bangalore, India.
Dissertation for M.Sc. Thesis: “ Molecular Genetic
Characterisation of *Saccharomyces cerevisiae* Prp17 protein and
It’s interacting partners”.

Since 1999 : Doctorate of Philosophy (Ph.D)
Institute of Human Genetics, Georg-August-Universität,
Göttingen, Germany.
Dissertation for Ph.D Thesis: “Elucidation of *Theg* Gene Role in
Spermatogenesis and Characterisation of a Novel Spontaneous
Mutation Named *nax* in Mouse”.