The *Tff* gene cluster encoding gastroprotective trefoil peptides and targeted disruption of the *Tff2* gene in mouse

D i s s e r t a t i o n

der Fakultät für Biologie
der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

vorgelegt von

Tuncay Kayademir
aus Herrenberg
2001

Tag der mündlichen Prüfung:	11. Mai 2001
Dekan:	Prof. Dr. H.-U. Schnitzler
1. Berichterstatter:	Prof. Dr. N. Blin / PD Dr. P. Gött
2. Berichterstatter:	Prof. Dr. G. Stamp

Babama

ACKNOWLEDGEMENTS

Prof. Dr. Nikolaus Blin für die Überlassung des Dissertationsthemas, sein immer reges Interesse an meiner Arbeit, seine unermüdliche Diskussionsbereitschaft und seine Unterstützung, sowie der Ermöglichung von Kongreßreisen und dem Aufenthalt in Porto.

PD Dr. Peter Gött im besonderen, für die intensive Zusammenarbeit, sein stetes Interesse an meinem Projekt, die vielen anregenden Diskussionen und seine große Unterstützung (in guten wie in schlechten Tagen) halfen mir diese Arbeit erfolgreich zu beenden.

I would especially like to thank Prof. Dr. Gordon Stamp, ICSM (Hammersmith Hospital, London, UK) for his warm hospitality during my visit to London. His unhesitating decision to become second referee of my thesis is an honour.

Ian Rosewell, Ph.D., and his colleagues (S. Wilson,...) at the ICRF (Transgenic Laboratory, Clare Hall, UK) deserve many many thanks for their work which was crucial to the completion of this thesis. Their knowledge and labwork were extremely valuable.

Our partner at the ICRF (Prof. Dr. Clive Dickson and Prof. Dr. Gordon Stamp), the co-operation program enabled this project; as well as for providing the pL2-neo vector.

Dem IZKF-Tübingen (Projekt IIIB 1) für die finanzielle Unterstützung meiner Projekte und dem DAAD-CRUP Exchange-Programm für meinen Porto Aufenthalt.

Prof. Dr. Horst Hameister, Universität Ulm, für die FISH-Analyse.

Allen Mitgliedern der TFF-Arbeitsgruppe für die gute Zusammenarbeit und ihre Geduld bei den ‘kurzen’ Vorträgen am Freitagmorgen.

Allen anderen Mitgliedern der Abt. Molekulare Genetik für die angenehme Atmosphäre und ihre freundschaftliche Unterstützung; im besonderen Hakan Esmer für die Sequenzdaten, sağol; und meinen Sekretärinnen für die vielen Telefonvermittlungen.

Für das Korrekturlesen der Arbeit konnte ich drei Freunde gewinnen; Kristina Schröder, Wolfgang Maier und Kirsten Gay, sağol.

Agradeço de coração a todos os colegas e amigos do IPATIMUP pelos dias “ensolarados” em Porto; em especial ao José Carlos Machado, pelo apoio e amizade.

Allen Freunden, nah und fern, dank Ich hiermit gern;
zur Hilfe jederzeit, waren Sie stets bereit.

und nicht zuletzt meiner Familie; her zaman yanımda olduğunuz için sonsuz teşekkürler ... Sayenizde... .

TABLE OF CONTENTS I

ABBREVIATIONS IV

1. INTRODUCTION 2
1.1 TREFOIL PEPTIDES 3
1.2 MEMBERS OF THE TREFOIL PEPTIDE FAMILY 5
1.2.1 Trefoil peptides in mammals 5
1.2.2 Amphibian trefoil proteins 8
1.3 GENOMIC STRUCTURE OF THE HUMAN TFF GENES 9
1.4 EXPRESSION OF TREFOIL PEPTIDES 10
1.4.1 Expression in normal human tissues 10
1.4.2 Expression in other mammals 13
1.4.3 Expression of trefoil peptides under pathological conditions 14
1.5 FUNCTION OF TREFOIL PEPTIDES 16
1.5.1 Transgenic and knockout studies 20
1.6 GENE TARGETING 22
1.6.1 Engineering a gene targeting construct 23
1.6.2 Transfection of embryonic stem (ES) cells 25
1.6.3 Assessing ES cells for homologous recombination and screening 26
1.6.4 Generating chimeric mice by blastocyst injection 26
1.6.5 Breeding chimeric mice to homozygosity 27
1.7 AIMS OF THE PROJECT 29

2. MATERIAL 31
2.1 ENZYMES AND REACTION KITS 31
2.2 CHEMICALS 31
2.3 VECTORS 31
2.4 FILTER, MEMBRANES, COLUMNS, FILM MATERIAL, TUBES 31
2.5 INSTRUMENTS 32
2.6 COMPUTER SOFTWARE, INTERNET 32
2.7 OLIGONUCLEOTIDES 33
2.8 BACTERIAL CLONING SYSTEMS AND BACTERIA STOCKS 34
2.8.1 Bacterial Artificial Chromosomes (BACs) 34
2.8.2 Bacteria strains 34
2.9 GENE TARGETING ("MOUSE") FACILITY 35
2.10 MEDIA FOR BACTERIA 35
2.11 BUFFERS, SOLUTIONS, AND DNA MARKERS 36

3. METHODS 39
3.1 ELECTROPHORESIS AND DNA RECOVERY FROM AGAROSE GEL 39
3.2 RESTRICTION ENDONUCLEASES 39
3.3 PCR 39
3.4 LIGATION 41
3.5 SCREENING BAC LIBRARY 42
3.6 TRANSFORMATION 44
3.7 RANDOM OLIGONUCLEOTIDE-PRIMED SYNTHESIS 45
3.8 SOUTHERN BLOT 45
3.9 SPOT HYBRIDISATION 46
3.10 HYBRIDISATION 47
3.11 COLONY HYBRIDISATION 47
3.12 DNA EXTRACTION 48
3.13 SEQUENCING ANALYSIS 50
3.14 GENE TARGETING 51
3.14.1 Construction of a replacement vector 51
3.14.3 Detection of homologous recombinants 53
3.14.4 ES cell injection 54
3.14.5 Chimeras 54
3.14.6 Mouse tail biopsies 55
3.15 PULSED-FIELD GEL ELECTROPHORESIS (PFGE) 55
3.16 AUTORADIOGRAPHY 56
3.16.1 Exposure to X-ray film 56
3.16.2 Using storage phosphor and phosphor imaging 56

4. RESULTS 58
4.1 CONSTRUCTION OF THE TARGETING VECTOR FOR DISRUPTION OF TFF2 58
4.2 GENOMIC STRUCTURE OF THE TFF GENE CLUSTER 68
4.2.1 Characterisation of the YAC/ BAC contig 68
4.2.2 Mapping of the Tff cluster to mouse chromosome 17 70
4.2.3 Distances and orientation of Tff genes by long distance PCR display the murine Tff gene cluster within 40 kb 70
4.3 SCREENING OF ES CELLS FOR HOMOLOGOUS RECOMBINANT CLONES BY PCR AND SOUTHERN BLOT ANALYSIS 73
4.4 BREEDING OF CHIMERAS AND HETEROZYGOUS MICE 76
4.4.1 Germline transmission confirmed by genotyping of F_1 offspring 77

5. **DISCUSSION** **80**

5.1 Replacement constructs substitute the endogenous target gene with the disrupted sequences 81

5.2 Use of increased size of homology and isogenic DNA improve the probability of correct targeting 82

5.3 Enrichment for homologous recombinant clones by positive-negative selection ... 83

5.4 Tff2 gene inactivation is achieved by deleting the signal peptide sequence and one trefoil domain 84

5.5 A successful strategy: a combination of two screening procedures 85

5.6 Genomic structure of the mouse Tff gene cluster 87

5.6.1 Only one BAC recombinant contains all three Tff genes 88

5.6.2 Mapping of the Tff cluster to mouse chromosome 17 using FISH 88

5.6.3 Murine gene distances but not the overall TFF cluster architecture are different from the human counterpart 89

5.6.4 Highest homology levels of amino acid sequences are found between corresponding trefoil domains from mouse and man 89

5.7 Outlook 91

6. **SUMMARY** **93**

ZUSAMMENFASSUNG **94**

7. **BIBLIOGRAPHY** **95**

Abbreviations

^{o}C	Degree Celsius
A, C, G, T	Nucleotides: adenine, cytosine, guanine, thymine
Amp	Ampicillin
APS	Ammonium persulfate
BAC	Bacterial Artificial Chromosome
BCEI	Breast cancer estrogen-inducible
bp	Base pair(s)
BSA	Bovine serum albumin
CHEF	Contour clamped homogenous electric field
cM	Centi Morgan
Cre	Causes recombination
ddNTP	Dideoxynucleotide
DNA	Deoxynucleic acid
dNTP	Deoxynucleotide
ds DNA	Double stranded DNA
E. coli	Escherichia coli
EDTA	Ethylenediamine Tetraacetic acid
e.g.	Exempli gratia (for example)
EGF	Epidermal growth factor
ES	Embryonic stem
EtOH	Ethanol
Fig.	Figure
g	Acceleration due to gravity
H_2O_{dd}	Double deionised water
h	Hour(s)
HCl	Hydrochloric acid
hr(s)	Hour(s)
i.e.	Id est (that is)
IPTG	Isopropyl-ß-d-thiogalactopyranoside
ITF	Intestinal trefoil factor
Kan	Kanamycin
kb	Kilo basepairs
kDa	Kilo Dalton
loxP	Locus of cross-over (x) in P1
µg	Micrograms

µl	Microliter
M	Molar
MCS	Multiple cloning site
min	Minute(s)
ml	Milliliters
mRNA	Messenger ribonucleic acid
ng	Nanograms
nm	Nanometers
O/N	Overnight
PCR	Polymerase Chain Reaction
RNase	Ribonuclease
rpm	Rotations per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
sec	Seconds
SP	Spasmolytic polypeptide
ss DNA	Single stranded DNA
SSC	Standard saline citrate
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
TFF	Trefoil FActor
T_m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
U	Unit
UACL	Ulcer associated cell lineage
V	Volt
Vol.	Volume
wt	Wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YAC	Yeast Artificial Chromosome

INTRODUCTION

1. Introduction

During the last ten years a considerable amount of knowledge has been accumulated on the function of trefoil peptides. Indeed, they not only play an important role in the protecting and healing processes of the stomach and intestine, but also demonstrate how crucial these small proteins are for the normal functioning of the gastrointestinal (GI) tract. Trefoil peptides are a small family of polypeptides, secreted predominantly by mucous cells in the gut. This family is composed of 3 mammalian members: TFF1 (or pS2), a gastric peptide; TFF2 or (spasmolytic polypeptide, SP), a distal gastric peptide; and TFF3 (or intestinal trefoil factor, ITF), which is found mainly in the intestine and colon. Two gene knockout studies in particular (Lefebvre *et al.*, 1996; Mashimo *et al.*, 1996) have demonstrated that normal trefoil peptide expression is vital to the dynamic processes of epithelial restitution and surface protection.

The mouse is one of the most widely used laboratory animals for studying mammalian biochemistry and developmental biology. Its popularity is largely due to an impressive range of transgenic and genetic techniques which allow researchers to manipulate or even "knockout" mouse gene expression.
In connection with the exponential increase in the number of genes identified by various genome projects, e.g. HUGO (Lander *et al.*, 2001; Venter *et al.*, 2001), it has become imperative to use efficient methods respectively for determining gene function *in vivo* in a mammalian organism.
Transgenic mice are produced by the injection of one or more transgenes into the pronucleus of a fertilised mouse oocyte, which results in the birth of a transgenic mouse carrying one to several copies of the transgene.
The use of gene knockout (selective removal of a gene by homologous recombination) has become a extremely effective approach that allows scientists to study and understand what individual genes may do. Two major elements are needed to produce a mutant mouse strain by homologous recombination: (1) an embryonic stem (ES) cell line (Mansour *et al.*, 1988; Ramirez-Solis *et al.*, 1993; Zimmer, 1992) capable of contributing to the germline and (2) a targeting construct containing target-gene sequences with the desired mutation. Once identified, mutant ES cell clones can be injected into a blastocyst in order to produce a chimeric mouse. The breeding of germ-line chimeras yields animals that are heterozygous for the mutation. These can be interbred to produce homozygous mutant “knockout” mice.

1.1 Trefoil peptides

Trefoil factors (TFFs) constitute a family of mucin-associated peptides containing one or more structurally characteristic trefoil domains (Thim, 1989).

The trefoil domain

Trefoil peptides share similarity in cysteine-rich regions of 38 or 39 amino acids known as trefoil domains/motifs (also termed P-domains) in which cysteine residue-disulphide bonds are formed according to an invariable 1-5, 2-4, 3-6 design. This amino acid sequence, together with the disulphide bonds, form a characteristic three-leaved (clover leaf) structure which has given the peptide family its name (Fig. 1). The minimum requirement for a trefoil domain would be the presence of the consensus sequence ($CX_{9\text{-}10}CGX_8CX_4CCX_9WCF$)(Hoffmann and Hauser, 1993; Thim, 1989; Thim, 1997).

The suggested disulphide arrangement was later confirmed by several three-dimensional structure determinations of trefoil domains (Carr, 1992; Carr *et al.*, 1994; Gajhede *et al.*, 1993).

Since this structural motif is extremely stable, these peptides have an extraordinary resistance to acid hydrolysis and proteolysis, and thus, making them well suited for activity in the harsh, luminal environment of the gut.

Members of this family have been discovered in various species. In mammals (e.g., man, mouse, and rat) three paralogous members designated as TFF1, TFF2, and TFF3 have been identified. In addition, in *Xenopus*, an African clawed frog, TFF-related proteins containing either one, two, four or six copies of the TFF domain have been documented. Table 1 summarises the collected data on the number of trefoil domains and main sites of expression. Interestingly enough, all TFFs and TFF-related proteins are normally expressed in mucin secreting tissues, namely the GI tract mucosa and the skin of *Xenopus laevis.*

When taken together, these data demonstrate that the trefoil domain, which has been well-preserved throughout evolution from amphibia to mammals, is likely to display important conserved functions (Sommer *et al.*, 1999).

Nomenclature

As a result of a nomenclature discussion on trefoil (P-domain) peptides at a Philippe Laudat Conference (autumn 1996, Aix-les-Bains, France), the scientific organising committee of the meeting recommended the use of a TFF (TreFoil Factor) nomenclature system for mammalian peptides, i.e. TFF1 for pS2/BCEI, TFF2 for SP and TFF3 for ITF/P1.B (P-domain 1.B)(Wright *et al.*, 1997).

It was further recommended that different species be marked by a lowercase letter prefixed to the peptide name, e.g. hTFF3 for human, mTFF2 for mouse rTFF1 for rat etc.. Moreover genes encoding the trefoil peptides were to be written in italics (*TFF1,...*) and the corresponding peptides in non-italic letters (TFF1,...). In addition, the standard "Genetic Nomenclature for Mice" was to be used. Mouse locus symbols, then, begin with a capital letter followed by lower-case letters (*Tff1*, *Tff2*, *Tff3*).

Table 1: Number of trefoil domains and major sites of expression of trefoil peptides grouped according to their homology (FIM: Frog integumentary mucin, X: Xenopus)

Name	Trefoil domains	Major sites	References
human TFF1	1	Stomach	Masiakowski *et al.* 1982, Rio *et al.* 1988
mouse TFF1	1	Stomach	Lefebvre *et al.* 1993, Otto *et al.* 1996,
rat TFF1	1	Stomach	Itoh *et al.* 1996
X. laevis xP1	1	Stomach	Hauser and Hoffmann, 1991
human TFF2	2	Stomach	Tomasetto *et al.* 1990
mouse TFF2	2	Stomach	Tomasetto *et al.* 1990, Lefebvre *et al.* 1993
rat TFF2	2	Stomach	Jeffrey *et al.* 1994
porcine TFF2	2	Pancreas	Jørgensen *et al.* 1982, Tomasetto *et al.* 1990
human TFF3	1	Intestine	Hauser *et al.* 1993, Podolsky *et al.* 1993
mouse TFF3	1	Intestine	Mashimo *et al.* 1995
rat TFF3	1	Intestine	Suemori *et al.* 1991, Chinery *et al.* 1992
X. laevis xP2	2	Skin	Hauser *et al.* 1992
X. laevis xP4	4	Stomach	Hauser and Hoffmann, 1991
X. laevis FIM-A.1[a]	4	Skin	Hoffmann 1988, Hauser *et al.* 1990
X. laevis FIM-C.1[a]	6	Skin	Hauser and Hoffmann, 1992

[a] larger proteins containing also mucin–like sequences

1.2 Members of the trefoil peptide family

1.2.1 Trefoil peptides in mammals

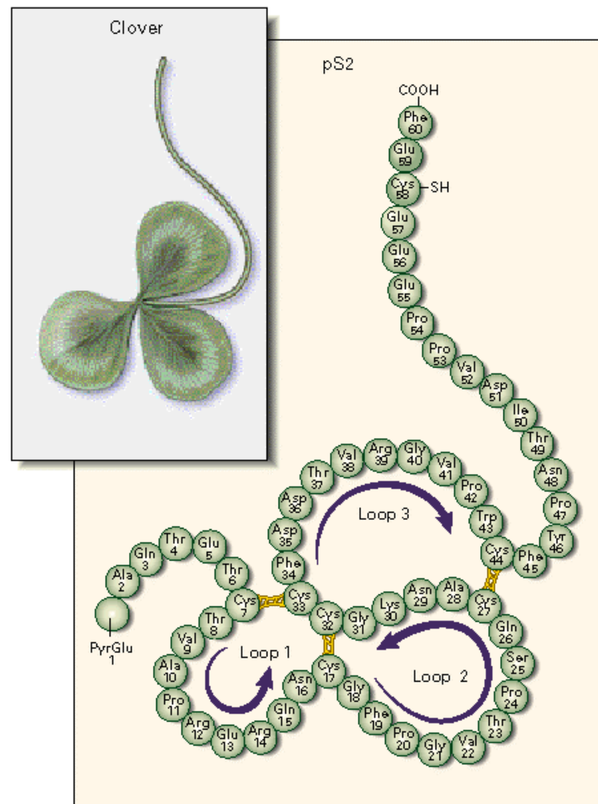
TFF1

Human TFF1 was discovered in 1982 while screening for oestrogen-inducible mRNAs in the library of the breast carcinoma cell line MCF-7. Masiakowski *et al.* (Masiakowski *et al.*, 1982) cloned this oestrogen-inducible cDNA designated pS2. The human pS2/TFF1 gene codes for a small pre-protein of 84 amino acids contains an aminoterminal signal peptide characteristic of secreted proteins (Jakowlew *et al.*, 1984; May and Westley, 1988). The secreted mature form of TFF1 is a 6,6 kDa protein of 60 amino acids (Fig. 1A)(Mori *et al.*, 1990) and displays an N-terminal pyro-glutamic acid residue (Rio *et al.*, 1988), a trefoil domain, and a carboxy-terminal acidic domain. The nucleotide sequence of the oestrogen-induced element, formerly called pS2, was published two years later (Jakowlew *et al.*, 1984; Prud'homme *et al.*, 1985).
Although TFF1 is derived from a breast cancer cell line and expressed in about 60% of human breast carcinomas (which can be sometimes of prognostic value)(Rio and Chambon, 1990), it is not a major constituent of normal breast tissue. In contrast to the insignificant amounts of mRNA and protein detectable in the normal breast, TFF1 is usually represented with high levels in the human stomach (Rio *et al.*, 1988).

The discovery of TFF1 homologues in mice, with proteins of either 62 or 66 amino acids (Lefebvre *et al.*, 1993; Otto *et al.*, 1996) and in rats (Itoh *et al.*, 1996; Table 1) has confirmed that these mammalian peptides are conserved during evolution (Sands and Podolsky, 1996).

Certain evidence suggests that TFF1 can exist as a dimer, bridging over the seventh cysteine residue (Cys58) to another TFF1 molecule. Though this ability might be functionally important, it further illustrates the possible role of specific residues in binding (Chadwick *et al.*, 1997; Polshakov *et al.*, 1997). Western transfer analysis has shown, that TFF1 is present in different molecular forms intimately associated with mucus in normal stomach (Newton *et al.*, 2000). Tomasetto *et al.* also demonstrated that mouse TFF1 interacts directly with the VWFC (von Willebrand factor C) cysteine-rich domains of mucins (Tomasetto *et al.*, 2000).

A



B

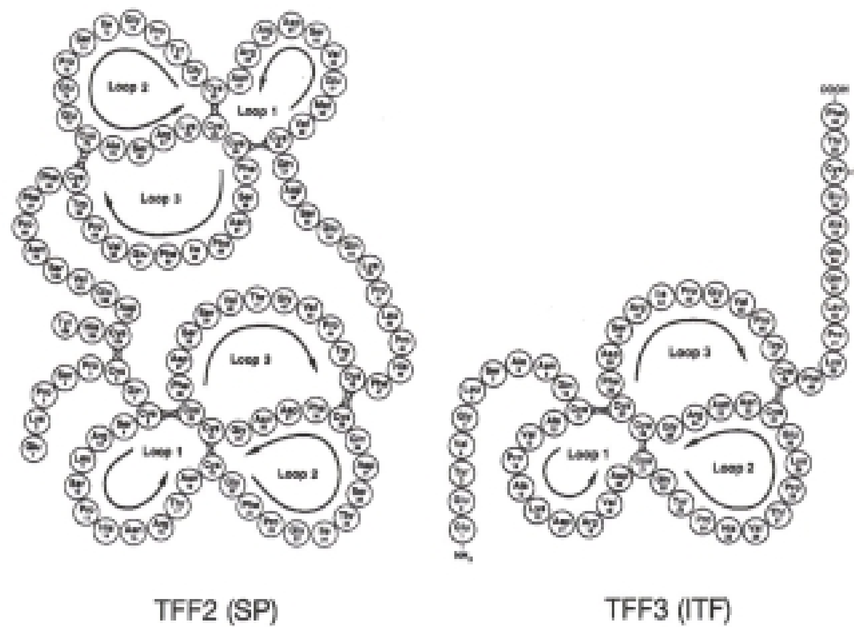


Figure 1: Amino acid sequence and proposed secondary structure of human trefoil peptides pS2/TFF1 (in comparison to a clover leaf) (**A**), SP/TFF2, and ITF/TFF3 (**B**) with the disulphide bonding of cysteine residues 1-5, 2-4, 3-6. Adapted after Plaut (1997) and Thim (1997).

TFF2

TFF2 was identified in 1990 (Tomasetto *et al.*, 1990) as a human homologue of a peptide isolated from the porcine pancreas, and originally called pancreatic spasmolytic polypeptide (pSP). The pSP was first isolated by Jørgensen *et al.* (Jørgensen *et al.*, 1982a) in 1982 within a side-fraction obtained by purifying insulin from a porcine pancreas. It was found to be abundant in porcine pancreas (100 mg/kg)(Thim *et al.*, 1982), in contrast to pS2/TFF1, where enough peptide was available to perform various physiological and pharmacological studies.

The human variant was named spasmolytic polypeptide (SP), since the pancreas did not seem to be a major site of expression. Several years later it was discovered that these peptides (pS2 and pSP) share a common unique sequence motif (Baker, 1988; Thim, 1988) later named the trefoil domain (Thim *et al.*, 1989).

The mature hTFF2 is a 12 kDa protein composed of 106 amino acids and contains two trefoil domains separated by a short sequence of amino acid residues (Fig. 1B left). The structure of pSP has been studied in detail by 3-D structure analysis (Carr *et al.*, 1994; De *et al.*, 1994), and regions have been identified on the surface which are candidates for interacting with oligosaccharide chains or putative receptors. The TFF2 protein was recently found to be glycosylated via an N-linkage (May *et al.*, 2000) but little is known concerning the possible function of this glycosylation.

Homologous peptides to pSP have been identified in other species such as rat and mouse (Table 1). A considerable conservation of amino acid residues is evident among TFF2 from different mammalian species (Poulsom, 1996). The first trefoil motif (the most N-terminal) is more strongly related to the single-trefoil motifs present in either TFF1 or TFF3. The (second) trefoil motif close to the C-terminus perhaps evolved later by an incomplete duplication event (Kayademir *et al.*, 2000).

TFF3

The third mammalian protein in the family, TFF3 (Fig. 1B right; originally known as intestinal trefoil factor-ITF- or hP1.B) was later discovered as a rat cDNA sequence from a rat jejunal library in 1991 (Suemori *et al.*, 1991), while the human cDNA sequence was reported in 1993 (Hauser *et al.*, 1993; Podolsky *et al.*, 1993). Clones for the rat were picked using oligonucleotide probes based on the peptide sequence of human TGIF (transformed growth-inhibitory factor; Podolsky *et al.*, 1988). Consequently, a relationship between TFF3 and TGIF has not been confirmed.

Sequence analysis has revealed that TFF3 from human, rat, and mouse sources (Table 1; Mashimo *et al.*, 1995) are single trefoil peptides (like TFF1) of 59 amino acids residues. Gel filtration analysis of extracts taken from human and rat tissues revealed only a monomeric ~6.6 kDa form (Taupin *et al.*, 1995). A dimerization of TFF3,

however, seems to occur naturally (Chinery and Playford, 1995), when expressed in bacteria (Chinery *et al.*, 1995) and yeast (Thim *et al.*, 1995). Characterisation of purified rTFF3 verified that both monomers and dimers can be observed under reducing and non-reducing conditions, respectively.
Site-directed mutagenesis studies illustrate that Cys57 is necessary for rTFF3 dimer formation. Samples of human gastrointestinal tissue following biopsy also demonstrate the presence of reducible human TFF1 and TFF3 covalent dimers (Chinery *et al.*, 1995). Thus, these results raise the intriguing possibility that a spectrum of biological activities may be generated by combining different single-trefoil peptides.

1.2.2 Amphibian trefoil proteins

A trefoil protein, originally called spasmolysin (Hoffmann, 1988), was first cloned from the skin of the clawed frog *Xenopus laevis*. Initially, this relatively large (~130 kDa) protein was presumed to undergo post-translational processing to yield two single trefoil-motif containing peptides, as well as another peptide bearing two trefoil motifs, and a region of extensive glycosylation (Hoffmann, 1988). Subsequently, this protein, called **f**rog **i**ntegumentary **m**ucin-A.1 (FIM-A.1), was found not to be cleaved. Thus FIM-A.1 exists in *X. laevis* skin as a peptide containing four trefoil domains, in addition to extensively glycosylated domains containing mucin-type, O-glycosidally-linked oligo-saccharide side chains (Hauser *et al.*, 1990).
Other trefoil-motif containing proteins secreted onto the frog's skin include a second FIM, which contains six trefoil domains (FIM-C.1; Hauser and Hoffmann, 1992) and a double trefoil domain protein, called APEG protein (Gmach *et al.*, 1990). In contrast, xP2, a protein with two trefoil domains, though present in skin, does not appear to be secreted onto mucus coating of the skin (Hauser *et al.*, 1992).
With respect to the identification of family members within the amphibian stomach, two proteins were found and designated xP1 (homologue of the TFF1 gene product) and xP4, containing one and four trefoil motifs. Both of these proteins are expressed by surface mucous cells of *X. laevis'* gastric mucosa (Hauser and Hoffmann, 1991). Because of the quasi-tetraploid genome (Thiebaud and Fischberg, 1977) of *X. laevis,* two xP4 genes were found, xP4.1 and xP4.2, differentially expressed (Botzler *et al.*, 1999).
The evolutionary conservation of the trefoil motif from amphibian to human proves the functional importance of this structure.

1.3 Genomic structure of the human *TFF* genes

The three human *TFF* genes are clustered within 50 kb in the chromosomal region 21q22.3 (Fig. 2)(Chinery *et al.*, 1996a; Gött *et al.*, 1996; Hattori *et al.*, 2000; Schmitt *et al.*, 1996; Theisinger *et al.*, 1992; Tomasetto *et al.*, 1992). This gene cluster is positioned within a CpG-rich region and is tandemly oriented (tel-*TFF1*-*TFF2*-*TFF3*-cen) with a transcription of all three genes directed towards the centromere (Gött *et al.*, 1996; Seib *et al.*, 1997).
Analysis of the exon structure of the three trefoil peptide genes revealed an extremely similar organisation. The first exon encodes the secretion signal sequence; the second (and the third exon of *TFF2*) encodes the trefoil domain(s); and the third exon (the fourth of *TFF2*) encodes four to five residues of the carboxyl terminus. This conserved structural organisation may have evolved from gene duplication and exon shuffling (Gött *et al.,* 1996).

Figure 2: The region on chromosome 21q22.3 and the genomic organisation of the gene cluster that encodes the human trefoil peptides. Coloured boxes represent the trefoil genes (*TFF1*,...).

The *TFF1* gene, formerly known as *BCEI* (breast cancer oestrogen inducible) gene, is formed by three exons (Rio and Chambon, 1990), with the single trefoil domain encoded by exon 2. The promoter region of *TFF1* contains several control elements, all located within 1 kb: an oestrogen responsive element (ERE) as well as elements responsive to phorbol esters, epidermal growth factors, c-Ha-*ras*, and c-*jun* (Berry *et al.*, 1989; Cavailles *et al.*, 1989; Nunez *et al.*, 1989).
The structure of the *TFF2* gene, whose locus was designated *SML1* (spasmolysin 1) contains four exons and is the only gene encoding two trefoil domains by exons 2 and 3 (Tomasetto *et al.*, 1990). These exons are separated by an intron containing an unique 25 bp tandem repeat cluster, repeated approximately 48 to 53 times (Kayademir *et al.*, 1998).

Analysis of the *TFF3* gene reveals the same basic structure of three exons as *TFF1* (Hauser *et al.*, 1993; Podolsky *et al.*, 1993; Seib *et al.*, 1995).

Besides gene clustering and transcriptional orientation, all three genes display a collection of regulatory signal sequences in their 5'-region sharing several motifs (e.g., the HNF-3 binding site; Beck *et al.*, 1999) with almost identical sequences or high sequence homology and spacing, suggesting a coordinated regulation and/or a common locus-controlling region (Gött *et al.*, 1996).
Two motifs with identical sequence and position are shared exclusively by the stomach-specific genes *TFF1* and *TFF2*, thus, presenting possible targets for stomach specific gene regulation. In addition, the genes react in a comparable fashion to stimuli such as osmolarity changes (Ludeking *et al.*, 1998) or show (only *TFF1*) a unique activation response to steroid hormones (Beck *et al.*, 1997).

1.4 Expression of trefoil peptides

1.4.1 Expression in normal human tissues

The predominant site of expression of all three mammalian (and human) TFF is in the gastrointestinal (GI) epithelial cells (Table 2).
Under normal circumstances TFF1 and TFF2 are mainly expressed in the stomach (Rio *et al.,* 1988; Tomasetto *et al.,* 1990), and TFF3 in the small and large intestine (Suemori *et al.,* 1991)(Fig. 3).

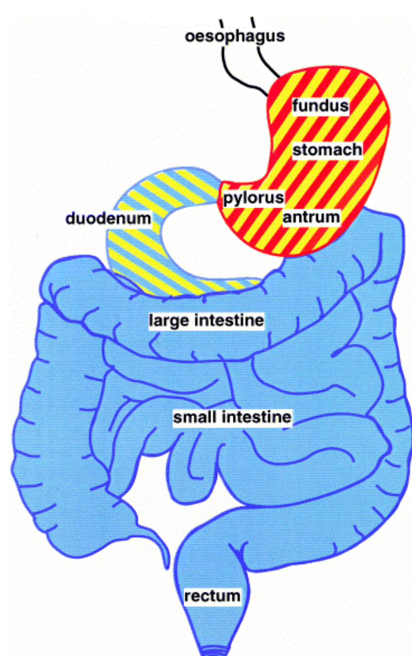


Figure 3: Restricted expression of TFFs along the GI tract. TFF1 is expressed in the stomach (red/yellow stripes), TFF2 in stomach and duodenum (yellow/blue stripes), and TFF3 in the small and large intestine (blue). Adapted after Ribieras *et al.* (1998).

Table 2 summarises data (incomplete) concerning sites of physiological expression of trefoil factors in normal human tissues.

Table 2: Sites of expression of TFF domain peptides in normal human tissues

Trefoil peptide	Organ / Site	Description	References
TFF1	Stomach	Mucus cells from neck upwards, all regions	Rio *et al.,* 1988; Luqmani *et al.,* 1989; Hanby *et al.,* 1993
	Small Intestine	Ductal luminal cells of Brunner's glands	Hanby *et al.,* 1993
	Large Intestine	Goblet cells near surface of crypts	Singh *et al.,* 1998
	Pancreas	Focally in duct epithelium	Wright *et al.,* 1990
	Salivary glands	Patchy epithelial expression	Rio *et al.,* 1988; Devine *et al.,* 2000
	Breast	Focally in duct luminal cells	Poulsom *et al.,* 1997
	Eye	Conjunctival goblet cells	Langer *et al.,* 1999
	Lung respiratory tract	Focally in epithelial (goblet and ciliated) cells of trachea and bronchi	dos Santos Silva *et al.,* 2000
TFF2	Stomach	Fundus; mucous neck cells of antrum; mucous cells in base of glands	Hanby *et al.,* 1993
	Small intestine	Brunner's glands (acini and distal ducts)	Piggott *et al.,* 1991
	Pancreas	Focally in duct epithelium	Wright *et al.,* 1990
	Gall bladder	Patchy epithelial expression	Seitz *et al.,* 1991
TFF3	Small intestine	Brunner's glands (acini and ducts, goblet cells)	Suemori *et al.,* 1991; Chinery *et al.,* 1992; Hauser *et al.,* 1993
	Large intestine	Superficial goblet cells	Suemori *et al.,* 1991; Chinery *et al.,* 1992; Hauser *et al.,* 1993
	Uterus	Epithelium	Hauser *et al.,* 1993
	Breast	Focally in duct luminal cells	Poulsom *et al.,* 1997
	Brain	Hypothalamus, Pituitary	Probst *et al.,* 1996
	Salivary glands	Abundantly in most mucous cells	Devine *et al.,* 2000
	Lung respiratory tract	Mucous cells in the acini of sub-mucosal glands, goblet cells of bronchi	Wiede *et al.,* 1999; dos Santos Silva *et al.,* 2000
	Eye	Conjunctival goblet cells	Langer *et al.,* 1999

TFF1 is abundantly expressed in the mucosa cells of the superficial and foveolar epithelium (Rio *et al.* 1988; Luqmani *et al.*, 1989; Hanby *et al.,* 1993). A high level of TFF1 has also been described in upper duct and surface cells of Brunner glands in the duodenum (Hanby *et al.*, 1993a). In the pancreas, only a few cells in large ducts appear positive (Wright *et al.*, 1990a), which might be due to a response to damage as seen in the gall bladder (Seitz *et al.*, 1991).

Salivary glands were reported as weakly immunopositive by *in situ* hybridisation and RT-PCR (Rio *et al.,* 1988; Devine *et al.*, 2000). The small intestine generally appears not to express TFF1 (though staining of the tips of villi in the ileum and jejunum is reported; Piggott *et al.*, 1991). In the large intestine, expression has been demonstrated in some goblet cells, particularly in the distal regions (Singh *et al.*, 1998). Outside of the GI tract, in other mucous epithelia, TFF1 has been observed focally in duct luminal cells of normal breast (Poulsom *et al.*, 1997) and focally in the epithelial (goblet and ciliated) cells of trachea and bronchi (dos Santos *et al.*, 2000). Another location of TFF1 is in the conjunctiva, where the goblet cells secrete this peptide together with the mucin MUC5AC as a constituent of the ocular mucus (Langer *et al.*, 1999).

The expression of TFF2 appears to be highly correlated with that of TFF1 in terms of organ specificity. However, although they are co-expressed in stomach cells (Tomasetto *et al.* 1990), TFF1 and TFF2 show striking differences in their cellular localisation, since TFF2 is observed in mucous neck cells of the fundus and basal cells of the antral and pyloric glands of human stomach (Hanby *et al.*, 1993b). In addition, TFF2 is expressed within Brunner's glands that are confined to the proximal duodenum (Piggott *et al.*, 1991). Focal expression is also observed in the duct epithelium of pancreas (Wright *et al.*, 1990a) and in gall bladder epithelium (Seitz *et al.*, 1991). The principal sites of TFF2 are indicated in Table 2 and *in situ* hybridisation data which usually are in agreement with the distribution of the peptide. Some exceptions are noted for TFF2; in the gastric mucosa. There is also TFF2 mRNA in the gland bases and in the foveolar surface. The peptide, however, is undetectable in the surface cells (Hanby *et al.*, 1993b).

While TFF1 and TFF2 are mainly expressed in the stomach, TFF3 is predominantly expressed in goblet cells throughout the intestine, in the acini of glands, and distal ducts of Brunner's glands (Suemori *et al.*, 1991; Podolsky *et al.*, 1993; Hauser *et al.*, 1993). TFF3 expression has been also observed in human uterus (Hauser *et al.*, 1993), normal breast (Poulsom *et al.*, 1997), in salivary glands (Devine *et al.*, 2000), in human hypothalamus and pituitary (Probst *et al.*, 1996), and the mucous cells in the acini of submucosal glands and goblet cells of bronchi of the lung (dos Santos *et al.*, 2000; Wiede *et al.*, 1999). It is also a secretory product of conjunctival goblet cells beside TFF1. Together with MUC5AC, they may contribute to the rheological properties of the tear film (Langer *et al.*, 1999).

1.4.2 Expression in other mammals

Despite the general conservation of expression patterns of different trefoil peptides, some variation in the sites of expression among species have been observed.

Pigs

When pSP (pTFF2) was first isolated from pancreas, this organ was thought to be major site for trefoil peptide expression (Thim *et al.*, 1992). Later studies have shown that the pig is unique with respect to TFF2. Indeed, the expression of pTFF2 has been localised in acinar cells of the pancreas, mucous cells of the stomach and duodenum, and epithelial cells in the jejunum and the ileum (Rasmussen *et al.*, 1992).

Rats

TFF2 expression in rats was found mainly in the antrum (Jeffrey *et al.*, 1994). By Northern blot analysis, TFF1 was found expressed abundantly in the stomach and only faintly in the duodenum (Itoh *et al.*, 1996). TFF3 was detected in the duodenum and in goblet cells of the small intestine and colon (Chinery *et al.*, 1992; Taupin *et al.*, 1995). In contrast to this apparent gastrointestinal specificity, TFF3 has been observed in some regions of the hypothalamus and in the pituitary gland (Probst *et al.*, 1995; Schwarzberg *et al.*, 1999).

Mice

In mice, TFF1 as well as TFF2 expression have been found in the stomach. A comparison of *TFF1* and *TFF2* gene pattern expression using Northern blots and *in situ* analysis revealed that the peptides were specifically expressed in different tissues and/or at different cellular levels in the mouse GI tract, thus, resulting in a complementary pattern of expression.

TFF1 mRNA was shown to be expressed in gastric mucous cells lining the surface epithelium and the pits of fundus, antrum, and antrum-pyloric regions. In contrast, TFF2 was traced in cells, presumably mucous neck cells, confined to the junction between the pits and the base of glands within the fundic mucosa, in the base of the antrum, and antrum-pyloric glands (Tomasetto *et al.*, 1990; Lefebvre *et al.*, 1993). In the case of TFF2, gene expression was also found within the Brunner gland's confined to the proximal duodenum (Lefebvre *et al.*, 1993) as well as in acinar cells of the pancreas (Tomasetto *et al.*, 1990; Lefebvre *et al.*, 1996; Terada *et al.*, 2001). TFF3 was found in goblet cells of the intestine and colon and in extremely small amounts in the stomach (Mashimo *et al.*, 1995).

Furthermore, embryonic development of the mouse demonstrated that while the expression of the TFF genes is highly dynamic, their distribution generally follows morphological boundaries and reflects the development of the tissue in which they are expressed. TFF expression is developmentally controlled in the GI tract and precedes mucous cell differentiation in several tissues (Otto and Patel, 1999).

1.4.3 Expression of trefoil peptides under pathological conditions

TFF1 was originally isolated from human breast carcinoma cell lines (Masiakowski *et al.*, 1982). Since this discovery, several studies have documented an association between trefoil peptides and human epithelial malignancies.
Table 3 summarises selective data regarding expression of human trefoil peptides in carcinomas of different organs.

TFF1 was found to be expressed in a variety of other carcinomas including those of the pancreas (Collier *et al.*, 1995; Welter *et al.*, 1992), lung (Higashiyama *et al.*, 1994), endometrium (Henry *et al.*, 1991), ovary (particularly mucinous carcinoma; Dante *et al.*, 1994), prostate (Bonkhoff *et al.*, 1995), gall bladder (Seitz *et al.*, 1991), colorectum (Labouvie *et al.*, 1997; Welter *et al.*, 1994), skin (Hanby *et al.*, 1998), and oesophagus (Labouvie *et al.*, 1999). In breast carcinoma, TFF1 expression can be detected in more than 50% of the tumours and is significantly associated with oestrogen receptor expression (Pallud *et al.*, 1993; Rio *et al.*, 1987; Skilton *et al.*, 1989), responsiveness to hormone therapy (Henry *et al.*, 1990), and favourable prognosis (Cappelletti *et al.*, 1992; Foekens *et al.*, 1990).
Interestingly, TFF2 is not expressed in breast carcinomas (Tomasetto *et al.*, 1990), but TFF3 expression is induced in breast carcinomas in a hormone (by oestrogen; May and Westley, 1997) dependent manner (Poulsom *et al.*, 1997).

The expression of TFF1 is lower in gastric adenomas and carcinomas than in adjacent normal gastric mucosa and hyperplastic polyps (Machado *et al.*, 1996a), and its expression was observed in a frequency ranging from 48-66% (57%-Luqmani *et al.*, 1989; 56%-Henry *et al.*, 1991; 48%-Muller and Borchard, 1993; 66%-Machado *et al.*, 1996b). Expression has been found to correlate with diffuse morphology of the stomach or with tumour development (Machado *et al.*, 1996), and one series showed a significant relationship between TFF1 expression and tumour stage (Muller *et al.*, 1993).

TFF3 expression in carcinomas is reflected in a small number of studies. Immunoreactive expression has been reported to occur in enterocytes in colonic adenomas and carcinoma, with significant associations among TFF3 expression and the degree of differentiation and mucin presence (Taupin *et al.*, 1996). It has also been detected in skin mucinous carcinoma (Hanby *et al.*, 1998), as well as in breast carcinoma (Theisinger *et al.*, 1996; Poulsom *et al.*, 1997) and is co-expressed with TFF1 (Poulsom *et al.*, 1997).

TFF2 expression in human carcinomas has been reported in squamous cell carcinoma of the oesophagus (Labouvie *et al.*, 1999), of the gastric (Theisinger *et al.*, 1991; Machado *et al.*, 2000), and biliary tract (Seitz *et al.*, 1991), and pancreatic carcinoma (Welter *et al.*, 1992).

Table 3: Expression sites of human trefoil peptides in carcinomas of different organs, as well as in different types of putative preliminary phases.

Trefoil peptide	Organ / Site	Description	References
TFF1	Oesophagus	Barrett's metaplasia, Carcinoma	Hanby *et al.,* 1994; Labouvie *et al.,* 1999
	Stomach	Carcinoma: diffuse > intestinal	Luqmani *et al.,* 1989; Henry *et al.* 1991; Theisinger *et al.,* 1991
	Small intestine	Gastric metaplasia	Hanby *et al.,* 1993; Khulusi *et al.,* 1995
	Large intestine	Hyperplastic polyps, Adenoma, Adenocarcinoma	Hanby *et al.,* 1993; Welter *et al.,* 1994; Labouvie *et al.* 1997
	Pancreas	Carcinoma	Welter *et al.,* 1992; Collier *et al.,* 1995
	Breast	Carcinoma	Foekens *et al,* 1990; Cappelletti *et al.,* 1996; Poulsom *et al.,* 1997
TFF2	Oesophagus	Barrett's metaplasia, Carcinoma	Hanby *et al.,* 1994; Labouvie *et al.,* 1999
	Stomach	Carcinoma	Theisinger *et al.,* 1991; Machado *et al.,* 2000
	Gall bladder	Carcinoma	Seitz *et al.,* 1991
	Pancreas	Carcinoma	Welter *et al.,* 1992
TFF3	Large intestine	Hyperplastic polyps, Adenoma, Carcinoma	Hanby *et al.,* 1993; Taupin *et al.,* 1996
	Breast	Carcinoma	Theisinger *et al.,* 1996; Poulsom *et al.,* 1997
	Skin	Mucinous carcinoma	Hanby *et al.,* 1998

In addition to their expression in carcinomas, trefoil factors are also expressed in a wide variety of ulcerative conditions of the GI tract, such as Barrett's oesophagus (Hanby *et al.*, 1994; Labouvie *et al.*, 1999), a model of gastric ulceration in the rat (Alison *et al.*, 1995), and in duodenal ulcers (Hanby *et al.*, 1993a; Khulusi *et al.*, 1995). Both TFF1 and TFF2 showed an expression pattern resembling that of the native gastric epithelium. Furthermore, they were found also in the small and large intestine in Crohn's disease (Wright *et al.*, 1990a; Wright *et al.*, 1993), which only can serve to underscore their importance as molecules involved in the repair of gastrointestinal mucosa.
In hyperplastic polyps and adenomas of the colon, all three trefoil peptides have been detected (Hanby *et al.*, 1993c; Taupin *et al.*, 1996), but only the TFF1 peptide could be demonstrated immunohistochemically, whereas both TFF2 and TFF1 showed mRNA within the polyps.

To summarise, although trefoil peptides in human and other mammals are expressed in a tissue-specific manner within the normal GI tract (major site of expression), the situation is somewhat different in the case of damage to the GI tract and other organs, respectively. The biological role of these peptides in tumorigenesis is less clear. Experimental evidence for a role in tumour suppression comes from studies using knockout mice lacking the TFF1 gene (Lefebvre *et al.*, 1996). All such animals developed antral adenomas and 30% of them developed carcinomas. Whether the TFF1 gene acts directly in tumour suppression at a genetic level or indirectly via maintenance of normal mucin production remains to be ascertained. However, a study of trefoil expression in the developing gut of the rat established that trefoil peptides are early markers of epithelial cell maturation, thus, suggesting a possible role for these peptides in gut epithelial cell differentiation (Familari *et al.*, 1998).

1.5 Function of trefoil peptides

A functional characterisation of trefoil peptides was first carried out with native purified porcine TFF2 (Jørgensen *et al.,* 1982). The availability of this peptide enabled several studies of the function of pTFF2, e.g. aiding in peristaltic movement (Jørgensen *et al.*, 1982b). However recent studies on the human and porcine peptides have not supported these earlier findings (McKenzie *et al.*, 1997; Playford *et al.*, 1995).

With the production of recombinant trefoil peptides by *E. coli* or yeast (Thim *et al.*, 1995), extensive characterisation of trefoil peptide function was made possible.

The close association of TFF expression with mucous secreting cells also suggested that trefoil peptides could possibly be important for the stabilisation of the mucous lining of the GI tract (Kindon *et al.*, 1995). Although the physiological functions are not completely understood, there is growing evidence, both *in vitro* and *in vivo*, which supports the participation of TFFs in maintaining a normal mucosal barrier (protection) and reconstitution / wound healing (Mashimo *et al.*, 1996; Playford *et al.*, 1996). For example, when TFF1 are co-packaged within mucous cell granules (Wright *et al.*, 1993), the stimulation of restitution by TFF2 is enhanced in a cooperative fashion by the addition of mucin glycoproteins purified from the intestine of either rat or man and achieve up to a fifteen-fold enhancement in restitution (Dignass *et al.*, 1994) or uptake of oral trefoil peptides (recombinant hTFF2 and rTFF3) protect against ethanol- and indomethacin-induced gastric injury in rats (Babyatsky *et al.*, 1996). Another *in vivo* study showed that the upregulation of trefoil expression is prominent during the repair phase after injury (Itoh *et al.*, 1996).
In general, over-expression of trefoil peptides in the GI tract is most often found in connection with ulceration and particularly in epithelial cells migrating across the base of such lesions (UACL, see below; Wright *et al.*, 1990; Rio *et al.*, 1991; Wright *et al.*, 1993).

In various chronic ulcerative conditions in man, glandular structures develop within the mucosa derived from the ulcer associated cell lineage (UACL)(Wright *et al.*, 1990b). These UACL glands display an up-regulation of all trefoil peptides (Hauser *et al.*, 1993; Wright *et al.*, 1990a). They express EGF (Wright *et al.*, 1990), TGF-α (Ahnen *et al.*, 1991) and lysozyme (Stamp *et al.*, 1992), all of which are potentially able to contribute to the healing process. In fact, local goblet and endocrine cell types may also be recruited to secrete TFF1 into the local environment (Wright *et al.*, 1993). It seems that this lineage, budding out from bases of crypts adjacent to ulcers, plays a key role in the restoration of mucosal integrity.

With regard to the association with growth factors, there is some evidence for synergism in intestinal mucosal protection between the trefoil peptides and growth factors, including EGF and fibroblast growth factor (Szabo *et al.*, 1994). Experiments point to a synergy between rTFF3 and EGF resulting in better protection against indomethacin-induced gastric damage (Chinery and Playford, 1995).

Studies on mouse breast carcinoma cell lines transfected with human TFF1 cDNA have also indicated that the peptide has the ability to effect morphogenesis, causing the cells to grow as branched rather than spheroid structures in collagen gels (Williams *et al.*, 1996). This might be induced by rapid phosphorylation of β-catenin and the down-

regulation of adhesion molecules, particularly E-cadherin (Hanby *et al.*, 1996; Liu *et al.*, 1997). TFF2 also promotes branching morphogenesis in MCF-7 cells (Lalani *et al.*, 1999).

Moreover, TFF2 and TFF3 act as motogens, and are able to promote epithelial cell migration when cell monolayers are wounded in a TGF-β independent manner (Dignass *et al.*, 1994; Playford *et al.*, 1995).
Besides the protection of mucosa and restitution after epithelial damage, TFFs have been shown to inhibit apoptosis (Chen *et al.*, 2000; Kinoshita *et al.*, 2000; Lalani *et al.*, 1999; Taupin *et al.*, 2000).

The mechanisms by which trefoil peptides exert their main functions – gastric and intestinal protection and healing – have formed the subject of several articles (e.g., Poulsom, 1996; Sands and Podolsky, 1996; Plaut, 1997; Podolsky, 1997; Wright, 2001)

One of the theories on how protection is accomplished maintains that trefoil peptides, together with mucin glycoproteins, form stable gel complexes (Thim, 1994). Indeed, a strong association exists between the expression of TFF and mucins, which both contribute to mucosal defence. In general, TFF1 is associated with MUC5AC expression, TFF2 with MUC6, and TFF3 with MUC2 (Longman *et al.*, 2000). Several trefoil motifs in *X. laevis* are encoded by genes that also bear mucin core sequences (Hauser and Hoffmann, 1992). In mammals, however, no such homologues are known thus far.
It has been also suggested that the interaction between TFF2 and mucin inhibits, both *in vivo* and *in vitro*, proton permeation through the mucus layer (Tanaka *et al.*, 1997). In fact, the addition of TFFs to purified mucin preparations leads to a rapid increase in optical density and viscosity (Sands and Podolsky, 1996). The three dimensional structure of the trefoil domain also contains a binding pocket that may accumulate the sugar side-chain of mucin glycoprotein (Gajhede *et al.*, 1993; De *et al.*, 1994). When considering the *in vitro* study by Dignass *et al.* (1994), in which the motogenic effect of trefoil peptides could be enhanced by mucin (motogenic effect demonstrated for TFF1; Williams *et al.*, 1996) and the fact that mucins and trefoil peptides are produced in the same cells and often up-regulated simultaneously upon injury (Wright *et al.*, 1993; Longman *et al.*, 2000), the hypothesis that these components act together to protect and heal the GI tract seems well-founded. However, it should be emphasised that no direct proof of such a complex formation has yet been established.

Another hypothesis aimed at explaining the function of trefoil peptides concerns their binding to extracellular receptors. The presence of basolateral receptors which mediate the action of trefoil peptides has been claimed (Chinery *et al.*, 1993; Chinery and Cox, 1995; Frandsen *et al.*, 1986). While some reports illustrate the binding of trefoil peptides to GI membrane proteins of different sizes (Tan *et al.*, 1997; Thim and Mortz, 2000), they were unable to determine, whether these receptor were significant on a functional level.

In accordance with this theory, TFFs provoke actions *in vitro* quite similar to a receptor mediated response. For instance, application of TFF3 to epithelial cells leads to phosphorylation of β-catenin on tyrosine within 10 seconds (Chinery and Cox, 1995). The well-known growth factor, EGF, also phosphorylates β-catenin, and during this process, recombinant rTFF3 causes tyrosine phosphorylation of the EGF receptor itself (Efstathiou *et al.*, 1998). This finding has also been confirmed (Taupin *et al.*, 1999). This activation, however, does not appear to be direct, suggesting that TFF3 enhances cell migration through modulation of E-cadherin/catenin complex function (Liu *et al.*, 1997).
Furthermore, TFF3 has been shown, again *in vitro*, to decrease extracellular signals related to protein kinase (ERK) activity and the mitogen activated protein kinase (MAPK) pathway, probably through induction of a tyrosine phosphatase (Kanai *et al.*, 1998).

Although trefoil peptides seem to play an important role in the defence of the GI tract (Plaut, 1997), other protective factors such as immunoglobins, cytokines, and growth factors must be taken into consideration, when evaluating the overall defence mechanism (Podolsky, 1997; Cook *et al.*, 1999; Wong *et al.*, 1999).

1.5.1 Transgenic and knockout studies

The possibility of inserting DNA from any source directly into the mouse was first demonstrated by Gordon *et al.* (Gordon *et al.*, 1980), who devised the strategy of pronuclear microinjection. Studies for knockout models were performed a few years later (Lin *et al.*, 1985; Thomas and Capecchi, 1987).

Three transgenic and knockout models were generated for studying the possible function of trefoil peptides. To explore the function in the mucosal repair process, a transgenic mouse overexpressing human TFF1 specifically in the villi of the jejunum was constructed (Playford *et al.*, 1996). In this model, a marked difference in the amount of damage caused by indomethacin compared with the controls was found, thus, supporting the data assumption that trefoil peptides play an essential role in stimulating the gastrointestinal repair process (Playford *et al.*, 1996).

An additional question concerns the degree to which TFF1 plays an autocrine role in the development of mammary tumours (approximately 50% of human breast tumours produce TFF1). Those transgenic mice which overexpressed human TFF1 and secreted it into milk did not seem to have any mammary gland dysplasia, and TFF1 expression had no cognisable effect upon the physiology and/or development of the suckling young or the transgenic mother (Tomasetto *et al.*, 1989).

In a study of mice overexpressing TGF-α in the gastric mucosa, a marked increase in the mucosal surface cells producing murine Tff1 was observed (Goldenring *et al.*, 1996). When TGF-α was knocked out (Cook *et al.*, 1997), the normal increase in trefoil peptide expression following gastric ulceration was not detected, indicating that induction of trefoil peptide expression after injury depends on the presence of TGF-α (Cook *et al.*, 1997).

In the last two studies the genes encoding Tff1 and Tff3, respectively, were deleted by gene-targeting techniques (refer to 1.6).
Disruption of the *Tff1* gene (Lefebvre *et al.*, 1996) did not affect Tff3 expression. In contrast, Tff2 expression was not detected in the $Tff1^{-/-}$ stomach samples but was detected in the $Tff1^{-/-}$ pancreas samples (Lefebvre *et al.*, 1996), indicating coordinated expression of Tff1 and Tff2 in the stomach as already suggested after their genomic analysis (Gött *et al.*, 1996). According to this model, the absence of Tff1 did not seem to be compensated by increased expression of the other trefoil peptides. The mice lacking Tff1 showed adenomatous hyperplasia in the antropyloric part of the stomach, and after five months, 30 % of them developed multifocal intraepithelial or intramucosal

carcinomas. The small intestine was characterised by enlarged villi and an abnormal infiltrate of lymphoid cells. These results prompted researchers to suggest that apart from being responsible for normal differentiation of the antral and pyloric gastric mucosa, TFF1 may also function as a gastric-specific tumour suppressor gene. However, further studies are needed to fully evaluate this suggestion.

In the other gene knockout experiment, mice lacking Tff3 had impaired mucosal healing and died from extensive colitis after oral administration of dextran sulfate (Mashimo *et al.*, 1996). The *Tff3*$^{-/-}$ suffered from poor epithelial regeneration after injury. In contrast, gastric genes *Tff1* and *Tff2* showed reduced expression, once again suggesting that trefoil peptides may individually regulate transcription of the entire family (Taupin *et al.*, 1999).
In addition, luminal administration of recombinant TFF3 to the *Tff3*$^{-/-}$ mice resulted in normal healing; a clever proof for the protective role of TFF3 in maintaining a normal mucosal barrier. Ultimately, these findings demonstrate that TFF3 is a factor that participates in healing and repair mechanisms in the intestinal mucosa (Mashimo *et al.*, 1996).

1.6 Gene targeting

Mice are increasingly being used as models for the study of various human diseases. This is primarily because among mammalian models, they are most amenable to genetic manipulations.
Until recently, mutants were produced by mutagenesis followed by selection for a particular phenotypic change. Gene targeting by homologous recombination is an especially effective technique in generating mouse strains with defined mutations in their genome. These genetically modified, 'designer' animals permit us to ask questions about elaborate and complex biological systems (Mansour, 1990; Zimmer, 1992).

Two key technologies have facilitated this experimental system. First, the isolation of embryonic stem (ES) cells were isolated as permanent *in vitro* cell lines (Evans and Kaufman, 1981; Martin, 1981) that can repopulate the blastocyst embryo and possess the capacity to contribute to the germline tissue of mice (Bradley *et al.*, 1984). Second, mammalian cells were found to recombine introduced vector DNA with a homologous chromosomal target. The possibility of homologous recombination between foreign DNA and existing homologous sequences in the mammalian genome was studied from the beginning of the 1980s, and evidence was presented by two groups in 1985 (Lin *et al.,* 1985; Smithies *et al.*, 1985). Mutations introduced into these cell lines were transmitted to the offspring (Doetschman *et al.*, 1987; Robertson *et al.*, 1986). Thereby, the concept of ‘designer’ mutations became a reality.

The process of generating ‘knockout’ animals demands several key steps to obtain animals in which both alleles of the gene of interest are non-functional (Galli-Taliadoros *et al.*, 1995).

Gene targeting in mice involves:

- Engineering a gene targeting construct
- Transfection of embryonic stem (ES) cells
- Assessing ES cells for homologous recombination and screening
- Generating chimeric mice by blastocyst injection
- Breeding chimeric mice to homozygosity

1.6.1 Engineering a gene targeting construct

The first step toward a construct design involves the availability of a genomic clone or sequences containing the gene of interest. Two basic configurations of constructs are used for homologous recombination: insertion constructs and replacement constructs.

The *insertion* construct contains a region of homology to the target gene cloned as a single continuous sequence and is linearised by cleavage of a unique restriction site within the region of homology. Homologous recombination introduces the construct into the homologous site of the target gene and interrupts it by adding sequences. This leads to gene duplication, with the marker gene as well as the plasmid DNA being inserted. While similar or higher gene targeting efficiency has been achieved using this strategy in comparison to the replacement construct (see below)(Deng and Capecchi, 1992; Hasty *et al.*, 1991a), it has not been adopted for knockouts. As a result, the normal gene can be regenerated from the mutated target gene by a random intrachromosomal recombination event. This phenomenon was used to create a new method termed the 'hit and run' procedure (duplications, with subtle mutations, are introduced into the target gene and then removed ; Hasty *et al.*, 1991b).

The *replacement* construct is the more commonly used construct. It contains two regions of homology to the target gene, located on either side of a positive selectable marker, the plasmid DNA sequences, and a linearisation site outside of the homologous sequences of the vector (Fig. 4). Nearly all constructs rely on the positive selection of an antibiotic resistance gene (e.g., neomycin or neo^r)(Thomas and Capecchi, 1987).
This insertion serves a dual purpose: it interrupts the targeted gene and provides a tool for selection. In some cases, a negative selectable marker may also be used to enrich the transfected cells against random integration events (Mansour *et al.*, 1988). Homologous recombination proceeds by a double cross-over that replaces the homologous target gene sequences with the replacement construct sequences. Any heterologous sequences at the ends of the vectors are excised from the vector and are not recovered as stable genomic sequences in the recombinant allele following targeting.
Because no duplication of sequences occurs, the normal gene cannot be regenerated.

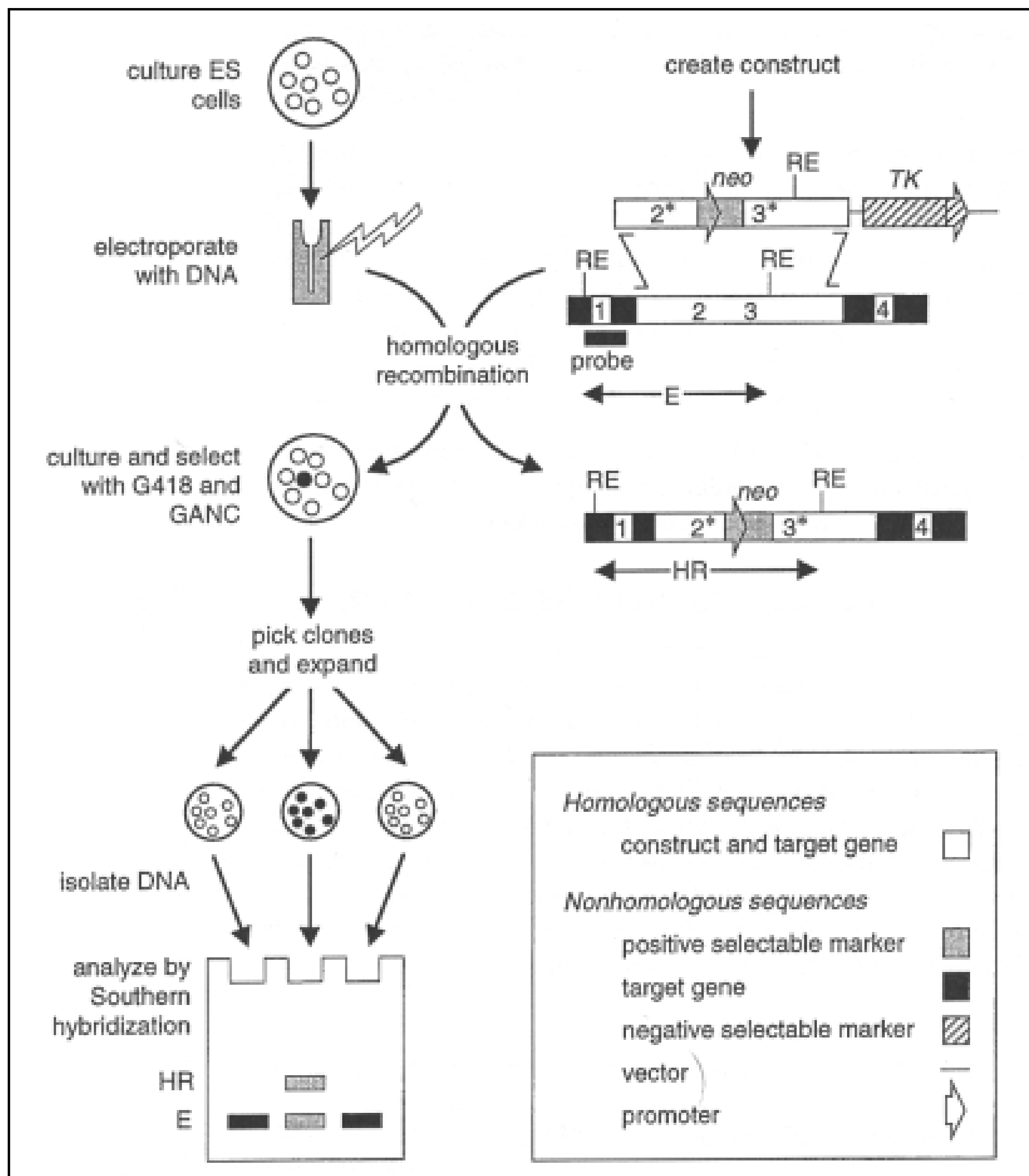


Figure 4: A replacement targeting construct. Production, selection, and identification of targeted gene disruption by homologous recombination. An example of a restriction enzyme site (RE) and hybridisation probe that can be used to identify cells in which homologous recombination has occurred (shaded) is shown (E: endogenous, HR :homologous recombinant). Adapted after Mortensen (2000).

The DNA used to construct the targeting vector must be from the same species as the cell in which the mutation is introduced. It should also be isogenic with the target cell (this is not absolutely required, but increases the probability of success), because animal strains may differ just as individual outbred animals differ. A single DNA mismatch is sufficient to decrease the rate of homologous recombination (Deng and Capecchi, 1992; Te Riele *et al.*, 1992).

Cre-loxP system

To study genes which are essential for murine development and therefore, lethal if disrupted, a system was developed, in which such genes can be inactivated in a tissue or cell-specific manner (Gu *et al.*, 1993; Gu *et al.*, 1994). This technique is based on the Cre-*loxP* recombination system of the bacteriophage P1 (Sternberg *et al.*, 1986). Cre recombinase is an enzyme which catalyses the site specific between the 34-bp motifs termed *loxP* site. If two *loxP* sites in the same orientation next to each other exist, Cre can act to loop out the sequence between the two sites, leaving a single *loxP* site in the original DNA and a second *loxP* in a circular piece of DNA, containing the intervening sequence.

Thus, a properly designed targeting construct containing *loxP* can be used either for introducing subtle mutations or for a temporally or spatially controlled knockout (Sauer, 1993).

1.6.2 Transfection of embryonic stem (ES) cells

Embryonic Stem Cells

Gene targeting is achieved in murine embryonic stem (ES) cells. ES cells are derived from 3.5 days postcoitum (dpc) mouse embryos and arise from pluripotent, uncommitted cells of the inner cell mass (ICM), the part of the mouse blastocyst that usually gives rise to the embryo proper (Evans and Kaufmann, 1981; Martin, 1981).

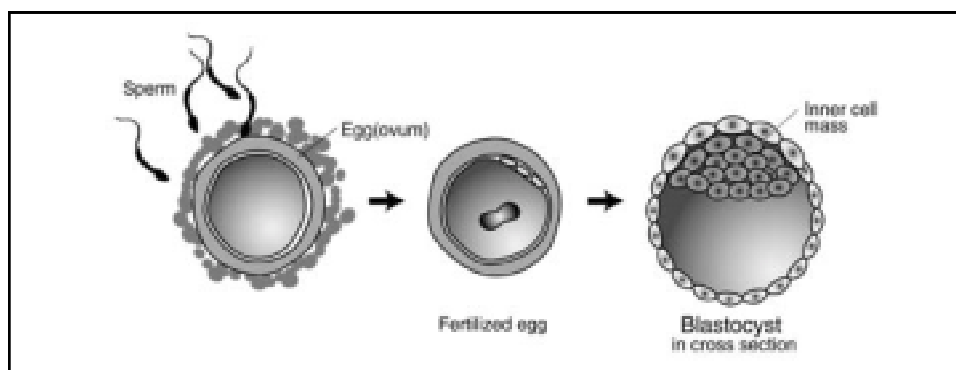


Figure 5 : Early development stages of the mouse (by D. Leja, NHGRI).

The ES cells can grow *in vitro* and retain the potential to contribute extensively to all of the tissues, including the germline tissue of an animal, when injected back into a host blastocyst, which is allowed to develop in a surrogate mother.

In the vast majority of these gene targeting experiments, ES cells derived from "129 mouse" substrains are an essential component (Galli-Taliadoros *et al.*, 1995; Simpson *et al.*, 1997). The most common method for introducing DNA into ES cells is by electroporation (Fig. 4)(Thomas and Capecchi, 1987).

1.6.3 Assessing ES cells for homologous recombination and screening

Not all ES cells actually take up the targeting construct. Those that do can be identified by growing cells in the selective medium (Fig. 4). A small proportion of integration may occur by homologous recombination, precisely at the endogenous counterpart of the transfected DNA.

To be sure, one of the most important aspects of any gene targeting experiment is to confirm that the desired genetic changes have occurred. Following the transfection of a replacement vector, colonies which survive selection should be clonally isolated and screened by PCR and Southern blot analysis for a specific recombination allele. First, a PCR-based screening for the rapid detection of candidate targeted colonies was performed (Kim and Smithies, 1988), followed by a second screening to verify the correct targeting by Southern blot analysis (Fig. 4).

1.6.4 Generating chimeric mice by blastocyst injection

The determining factor in the production of a gene knockout mouse strain is the generation of ES cell embryos that transmit the desired alteration to the next generation. Until recently, chimeric mice were almost exclusively produced by injection into the blastocyst of ES cells displaying the desired genotype and subsequently introduced into pseudo-pregnant recipient females (techniques are described in detail in Joyner, 2000; Fig. 6).

After blastocyst injection, ES cells become incorporated within the developing ICM of the embryo and contribute to the development of different (chimeric) embryonic lineages, in competition with host cells.

Most ES cell lines have an XY/male genotype (female derived XX cell lines are reported to be unstable). This has two advantages. First, XY ES cell lines, when injected into female/XX blastocysts, tend to bias the development of the resulting chimera toward a male phenotype (Iannaccone *et al.*, 1985). In male chimeras, only XY-bearing germ cells (i.e., those derived from the ES cells) will form functional gametes. XX primordial germ cells (i.e., those derived from the host blastocyst) will not form functional gametes and are lost. Therefore, the development of gametes derived from the ES cells is favoured.

Second, a male chimera can produce more offspring over its reproductive life span (~1.5 years) than a female (~8 litters), so that even chimeras with low percentage contribution of ES cells to the germline can be detected.

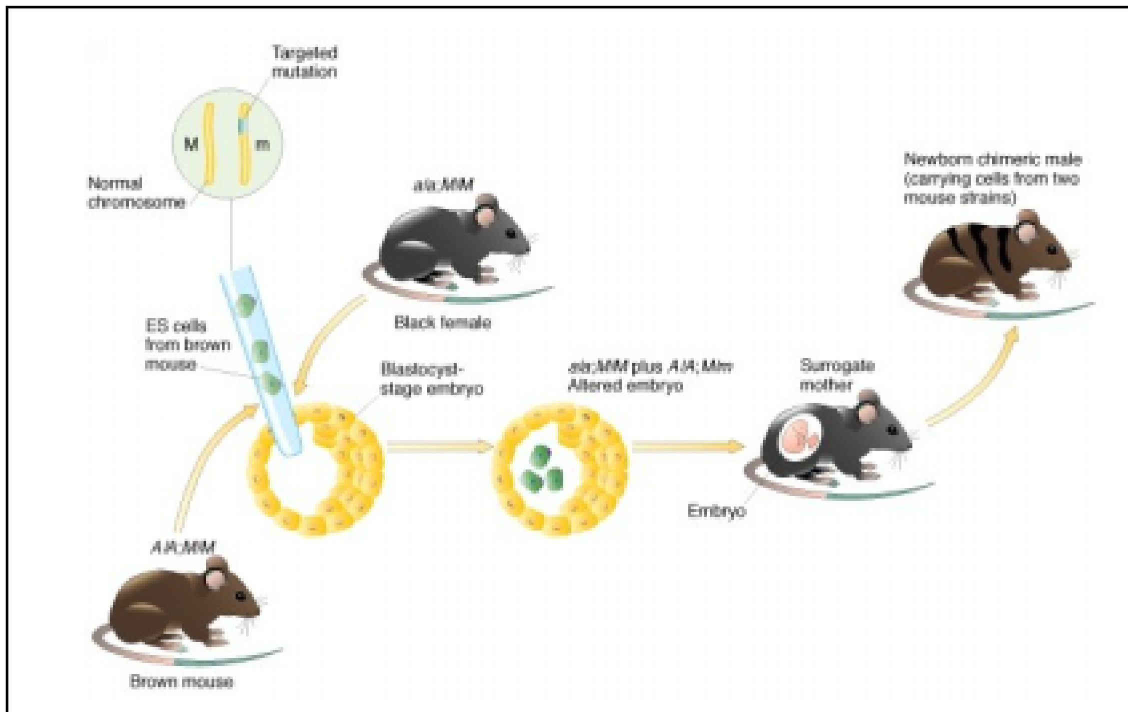


Figure 6: Production of a chimeric mouse. ES cells (green at top left) are isolated from an agouti mouse strain and altered to carry a targeted mutation (m) in one allele (M: wild-type allele). The ES cells are then injected into blastocysts. The embryos are transferred into surrogate mothers. Newborn chimeras indicate that the ES cells have survived and proliferated. *A* represents agouti, *a* black. Adapted after Griffiths *et al.* (1999).

1.6.5 Breeding chimeric mice to homozygosity

Most ES cell lines are derived from the "129 mouse" substrains (e.g., 129/J, 129/Sv). When injected into embryos of the C57BL6/J strain, the 129 ES cells tend to predominate in the chimeras (Schwartzberg *et al.*, 1989), thus affecting the likelihood that the ES cells will contribute to the germline. The most practical and readily apparent genetic marker of chimerism is coat colour. Most of the 129 ES cells have an agouti coat colour genotype (besides pink-eyed chinchilla for 129/Ola & 129/J and light chinchilla/albino for 129/SvJ; Simpson *et al.*, 1997), whereas that of C57BL6/J embryos is black. Therefore, the chimeric mice produced from this combination will display a colour mix of agouti/black (Fig. 6 and 37), which provides strong evidence for the survival of ES cells in the embryo.

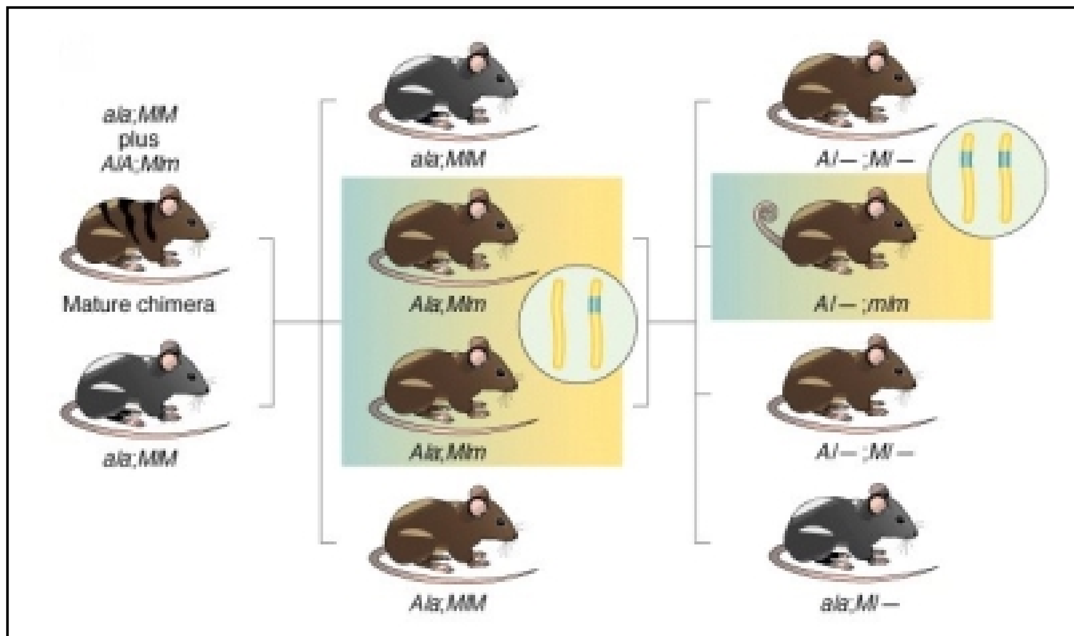


Figure 7: Breeding scheme to obtain homozygous mutant (m) animals. Chimeric males are mated to C57BL6/J female. Litters are screened for evidence of the targeted mutation (m) (green in inset). Males and females (concentric box) carrying the mutation are mated to one another to produce F2-mice whose cells carry the mutation (m/m) in both alleles (knockout animal in a curly-tail phenotype). Such animals (box on the right) are characterised by genomic analysis of their tail DNA. A represents agouti, a black and M wild-type allele. Adapted after Griffiths *et al.* (1999).

Only some of the animals carry the targeted gene in their germline. The criterion, for being able to transmit the ES genotype through the germline, is that more than 50% of the offspring born should be chimeric. Thus, chimeric males are generally test bred to ascertain contribution of the ES cells to germline (Fig. 7). A good test breeding scheme for this common combination is to breed chimeras back to C57BL6/J mice. Agouti animals are derived from the 129 ES cells and will be 129xC57BL6/J F_1 (Fig. 7, centric); non-agouti animals will be pure C57BL6/J.
If test breeding is done, the first available heterozygous animals are likely to be F_1 mice between two inbred strains (a 129 substrain and C57BL6/J). These mice will be genetically uniform except for the targeted allele and any alleles that were segregating in the 129 substrain from which the ES cells were derived (Simpson *et al.*, 1997). Once they have been characterised by Southern blot analysis of the tail DNA, heterozygotes for the targeted allele can be mated together (Fig. 7, centric box) to produce F_2 litters representing wild-type, heterozygous, and homozygous mutant animals.

Once a suitable breeding scheme is established for maintaining the mutation and the timing of any mutant effect, the challenge is to determine the cause of phenotypic effects, whether they are mild or severe.
The more is known about the biological function of the gene and its expression pattern, the easier it is to predict possible causes of any given phenotype. Histological analysis and gene marker analysis at embryonic or later stages (prior to death) can then be used to characterise the progression of the defect and in a causal way.

1.7 Aims of the project

The creation of knock-out / transgenic mice is of major interest to understanding the function of the TFF peptides.
At the beginning of this work Tff1 and Tff3 knockout mice had been constructed, but a Tff2 null-mouse did not exist. Mice lacking Tff1 showed decreased gastric mucus production, all developed gastric adenomas, and 30% of them developed multifocal intraepithelial or intramucosal carcinomas. At the same time *Tff3*$^{-/-}$ mice revealed a normal phenotype, but exhibited impaired mucosal healing and died from extensive colitis after oral administration of dextran sulfate sodium (Lefebvre *et al.*, 1996; Mashimo *et al.*, 1996). Remarkably, the expression of Tff2, which is normal in stomach, duodenum, and pancreas was not detected in the stomach samples of the *Tff1*$^{-/-}$ mice (Lefebvre *et al.*, 1996).
The present work was elaborated in order to elucidate the function of Tff2 *in vivo*. Therefore, the primary object was first the construction of a targeting vector to knockout the *Tff2* gene and second the characterisation of this gene targeting event in heterozygous (*Tff2*$^{+/-}$) mice.

For basic information of this project, the genomic structure of the mouse *Tff* gene cluster needed to be determined. Since the three known human genes are clustered in 21q22.3 (Gött *et al.* 1996), a similar arrangement in a homologous region of the mouse was to be expected. All three *Tff* genes were described (Lefebvre *et al.* 1993; Tomasetto *et al.*, 1990; Mashimo *et al.*, 1995) and the mouse *Tff3* gene was mapped to 17q by FISH (Burmeister and Meyer, 1997; Chinery *et al.*, 1996b).
To investigate this arrangement, four genomic recombinants were aligned into a contig and analysed in detail by restriction mapping and long range PCR.

MATERIAL

2. Material

2.1 Enzymes and reaction kits

[α^{32}P] dCTP	ICN, Eschwege, D
CIAP (calf intestine alkaline phospatase)	Roche Molecular Biochemicals, Mannheim, D
DNA Sequencing Kit	Perkin Elmer, Weiterstadt, D Amersham Pharmacia, Freiburg, D
Expand™ Long Template PCR System	Roche Molecular Biochemicals
Jetstar Plasmid Purification Kit	Genomed GmbH, Bad Oeynhausen, D
Klenow enzyme	Roche Molecular Biochemicals
pGEM®-T Vector System I	Promega, Mannheim, D
Proteinase K	Merck, Darmstadt, D
QIAquick Gel Extraction Kit	Qiagen, Hilden, D
Random Primed Labelling Kit	Roche Molecular Biochemicals
Restriction endonucleases	NEB, Franfurt, D; Takara, Taufkirchen, D; Roche Molecular Biochemicals MBI Fermentas, St.Leon-Rot, D
T4 DNA Ligase	Roche Molecular Biochemicals
Taq-Polymerase	Promega, Mannheim, D

2.2 Chemicals

The chemicals included in these various studies were purchased from different companies, such as Merck, Sigma-Aldrich, ICN, Difco, Roth, Pharmacia, Roche, Biomol, or Bio-Rad.

2.3 Vectors

pL2-neo	kindly provided by Prof. Dickson, ICRF, London, UK
pBK-CMV	Stratagene, Amsterdam, NL
pGEM-T®	Promega, Mannheim, D

2.4 Filter, membranes, columns, film material, tubes

0.2 ml PCR-stripes	Biozym, Hess. Oldendorf, D
12-ml tubes	Greiner, Frickenhausen, D
Butterfly membrane	Schleicher&Schuell, Dassel, D
Developing cassette	Siemens, Nixdorf, D
Eppendorf tubes (0.5 ml and 1.5 ml)	Eppendorf, Hamburg, D
Falcon tubes (50 ml)	Greiner
Fuji Medical X-ray film RX	Hirrlinger, Baden-Baden, D
Kodak GBX Developer	Hirrlinger
Kodak GBX Fixer and Replenisher	Hirrlinger
Mitsubishi Video Copy Processor	Hirrlinger

Nitrocellulose filter, 0.45 µm	Nalgene, Rochester, NY, USA
Nitrocellulose filter, 0.22 µm	Schleicher&Schuell
Nylon membrane, Nytran N	Schleicher&Schuell
Parafilm	American National Can™, Menasha, WI, USA
Storage phosphor imaging screen plus exposure casette	Bio-Rad, Munich, D
Whatman 3MM paper	Whatman, Maidstone, UK

2.5 Instruments

Biofuge 15 R	Heraeus, Osterode, D
Electrophoresis apparature	Gibco BRL, Karlsruhe, D
Electroporation System ECM 399 (and cuvettes)	BTX, San Diego, CA, USA
Gel Documentation System	LTF Labortechnik, Wasserburg, D
Gel-Dryer Model 583	Bio-Rad, Munich, D
Heat block	Barnstead, Dubuque, IA, USA
Heating plate	Bachofer, Reutlingen, D
Hybridisation oven	Bachofer
Incubator	Bachofer
LKB 2015 pulsator electrophoresis unit	Pharmacia
Micro centrifuge	Eppendorf, Hamburg, D
Microwave	Siemens
Mini centrifuge	Roth
Personal Molecular Imager FX	BioRad
Photometer	LKB-Pharmacia
Shaker	Pharmacia
Shaker-incubator	Bachofer
Sorvall-Centrifuge RC-5C	DuPont, Newton, CT, USA
Sterile bench	BDK, Sonnenbühl, Genkingen, D
Thermocycler, PCR	Biozym PTC-100 / PTC-200 Hybaid, Heidelberg, D Perkin Elmer 9600, Rodgau-Jügesheim, D
UV-Stratalinker™ 1800	Stratagene
UV-Table	LTF-Labortechnik
Vacuum Concentrator	Bachofer
Water bath	B. Braun, Melsungen, D

2.6 Computer software, Internet

- NCBI (GenBank): http://www.ncbi.nlm.nih.gov/
- European Bioinformatics Institute (EBI): http://www.embl-ebi.ac.uk/
- The Jackson Laboratory: http://www.jax.org/
- Mouse Genome Database: http://www.informatics.jax.org/mgihome/
- Primer: http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi
- DNAsis (Hitachi Software Engineering)
- Quantity One for Personal Molecular Imager FX (Bio-Rad, Munich, Germany)
- Gel documentation analyse software “Bio 1D” (LTF, Wasserburg, Germany)
- Paint Shop Pro 5 / MS Office

2.7 Oligonucleotides

The oligonucleotides (primer) syntheses were made by following companies:

- MWG-Biotech, Ebersberg, D
- Gibco BRL, LIFE Technologies, Karlsruhe, D
- BIG-Biotech, Freiburg, D

Table 4: List of primer pairs for PCR and primers for sequencing (f-r-for-rev refer to forward and reverse - orientation of the primers).

Name	*Sequence*	*Anneal.-temp.: °C*	*Purpose/characteristica*
m2-pro m2-end	TAGGTCTGCTCTTTAGAGCCTG AAATGAAGCTTCCTTTCGGCAAC	59°C	Tff2 gene locus
ClaI-for BamHI-rev	GG**ATCGAT**CCCCATCCCATACCC CG**GGATCC**TATAGGGTCTGAGGTC	57 > 72°C	generation of restriction sites (bold) for target vector
XbaI-for XbaI-rev	GC**TCTAGA**GCCTGGGATGGGGC GC**TCTAGA**GCCTGGACCGACAG	57 > 72°C	generation of restriction sites (bold) for target vector
m2/251-for m2/1189-rev	GGCACTGTGTGAGGGAAGTGA AATCTGTCACCCTGCCTTTGTC	59°C	5'-flanking DNA including exon 1 of *Tff*2 and a part of intron 1
Tff2f 5'-flanking Tff2r intron1	GGCACTGTGTGAGGGAAGTGA AATCTGTCACCCTGCCTTTGTC	59°C	5'-flanking DNA containing exon 1 of *Tff*2 (P2*)
Tff1f 5'-flanking Tff1r exon1	TCCGCTCAGCTCTACCTGAG CAGATCACCTTGTGCTCCATG	59°C	5'-flanking DNA containing exon 1 of *Tff*1 (P1)
Tff3f 5'-flanking Tff3r 5'-UTR	TGCTAGGCCCTGCTGCTGC CAGGACCACTGCACAGGATG	59°C	5'-flanking DNA of *Tff*3 (P3)
Tff1f 3'-UTR Tff2r 5'-flanking	TGACACAGTTCAACCCCTCAG GAGTCCATGACTCGGATACAG	59°C	distance between *Tff*1 and *Tff*2
Tff2f 3'-UTR Tff3r 5'-flanking	TTGGGAAGTCACCCTGAACTG GAGTAAAGAGCCGAAGACAGG	59°C	distance between *Tff*2 and *Tff*3
m2-pro m2-250rev	TAGGTCTGCTCTTTAGAGCCTG ATGCCCACCCATGAAATGCCC	61°C	PCR screening of BAC library (P2)
Reko-for WT-rev	GGTACCCCGGGTTCGAAATC ACTGTGATCAGCAGTCATGCG	59°C	screening for homologous recombination
Tff2ex2for WT-rev	GTCCCYTGGTGTTTCCACCC ACTGTGATCAGCAGTCATGCG	59°C	wild-type fragment (as a standard control)
M13-for	GTAAAACGACGGCCAGTGAATT	50°C	sequencing of ligation boundary
M13-rev	CAGGAAACAGCTATGACCATGA	50°C	sequencing of ligation boundary
SP6-r pL2	GCCAAGCTATTTAGGTGACAC	50°C	sequencing of ligation boundary
neo-rev	TTGGGTGGAAACATTCCAGGC	50°C	sequencing of ligation boundary
SeqEco-r	ACGAATTCCCCAGTGGCA	50°C	sequencing of ligation boundary
Pst-m2-for	GATTCTCTCAAGACCTGCT	50°C	sequencing unknown data from 5'-flanking DNA of *Tff2*

2.8 Bacterial cloning systems and bacteria stocks

2.8.1 Bacterial Artificial Chromosomes (BACs)

(Shizuya *et al.*, 1992)

BAC recombinants harbouring genomic mouse DNA from strain 129/Sv (new nomenclature: 129/S1; Simpson *et al.*, 1997) were obtained from Research Genetics, Huntsville, AL, USA. Primers (m2pro/m2-250rev) were designed from genomic sequences (GenBank database accession No. U78770) to screen recombinants by PCR. This PCR product yielded 3 BAC recombinants.

Facts about mouse BACs from CalTech:-*http://informa.bio.caltech.edu/idx_www_tree.html*-

- vector: pBeloBAC 11,
- host: HS996, modified DH10B
- average insert size: 130 kb

Table 5 : Mouse BAC clones containing the trefoil genes obtained by PCR screening.

BAC name	insert size	vector size
123M10	97 kb	7.4 kb
517E16	135 kb	7.4 kb
566K2	150 kb	7.4 kb

2.8.2 Bacteria strains

In the context of this work following strains were used:

- *E. coli* XL1-Blue MRF´strain:

Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F´proAB $lacI^qZ\Delta M15$ Tn10 (Tetr)]

Reference: Stratagene

- SURE strain:

e14– (McrA–) Δ(mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5 (Kanr) uvrC [F' proAB $lacI^qZ\Delta$ (M15 Tn10 (Tetr)]

Reference: Stratagene

- GM 272

Sex: F-
Mutation: (fhuA2 or fhuA31) (lacY1 or lacZ4) (tsx-1 or tsx-78) glnV44(AS) galK2(Oc) LAM- dcm-6 dam-3 mtlA2 metB1 thi-1 hsdS21

Reference: http://cgsc.biology.yale.edu/cgi-bin/sybgw/cgsc/Strain/12910

2.9 Gene targeting ("mouse") facility

The work with ES (embryonic stem) cells and their respective cell lines, blastocyst injections and mice breeding procedures (chimeras, heterozygous, and homozygous) were performed by colleagues (I. Rosewell *et al.*) from the ICRF at the Transgenic Lab in Clare Hall, UK. Upon establishment of the chimeric mice, the breeding programme was transferred to the ICSM at the Hammersmith Campus, London, UK.

2.10 Media for bacteria

- LB-(Luria Bertani) medium (modified)

bacto Tryptone:	10 g
bacto Yeast extract:	5 g
NaCl:	5 g
Glucose:	1 g

Dissolve the above components in double distilled water. Adjust the volume to a liter and autoclave it. Cool the solution and add the appropriate amount of antibiotic just before use. Typically Ampicillin (for T/A-vector and the targeting vectors) at 100 µg/ml, Kanamycin (for the pBK-CMV vector) at 50 µg/ml or Chloramphenicol (for BAC-recombinants) at 12.5 µg/ml.

- LB agar

Per one liter LB media (without glucose !) add 16 g Bacto agar and autoclave.

- SOB medium (per liter)

To 950 ml H_2O_{dd}, add:

Bacto tryptone:	20 g
Bacto yeast extract:	5 g
NaCl:	0.5 g

Shake until the solutes have dissolved and ad 10 ml of 250 mM KCl. Adjust the pH to 7.0 with 5 N NaOH and the volume to 1 liter. Sterilise by autoclaving. Just before use, add 0.01 Vol. of 2 M $MgCl_2$ (autoclaved).

- SOC medium

It's identical to SOB, except it contains glucose. Before use add 0.01 Vol. of a sterile (by filtration through a 0.22µm filter) 2 M solution of glucose (final conc.: 20 mM)

- X-gal for agar-plates

Stock solution: 20 mg/ml in dimethylformamide (use a glass or polypropylene tube), should be wrapped in aluminium foil to prevent damage by light (stored at –20°C).

- IPTG for agar-plates

Stock solution: 100 mM (23,8 mg/ml) in 70% Ethanol (stored at -20°C).

Before making the agar plates, add 2 ml of X-gal and 2 ml IPTG to 1 liter LB agar (after autoclave, allow it to cool to 50-55°C), if needed.

2.11 Buffers, solutions, and DNA markers

Electrophoresis Buffers:

- 10x TAE: 400 mM Tris-HCl, pH 8.0
  50 mM Sodium acetate
  10 mM EDTA

- 10x TBE: 890 mM Tris-HCl
  (pH 8.3) 890 mM Boric acid
  25 mM EDTA

- Ethidium Bromide: 0.1 mg/ml in 1xTAE

- 10 % APS:
  Place 1.0 g of ammonium persulfate (APS) into a tube and add distilled water to 10 ml volume. Invert to dissolve the APS and store at 4°C for up to one week.

- TE-Buffer: 10 mM Tris-HCl
  (pH 8.0) 1 mM EDTA

- Gel-loading buffer (10x): 0.25 % bromphenol blue
  0.25 % xylene cyanol FF
  30 % glycerol in water (storage temperature: 4°C)

Blotting solutions:

- Depurination solution: 0.25 M HCl
- Denaturing solution: 0.5 M NaOH
- Transfer solution: 0.5 M NaOH
  1.5 M NaCl
- Neutralising solution: 50 mM Phosphate buffer, pH 7.0
  50 mM Na_2HPO_4 (dibasic) titrated with 50 mM NaH_2PO_4 (monobasic)

Hybridisation solutions:

- Hyb-mix(ture), also called CHURCH buffer
  1 % BSA
  7 % SDS
  1 mM EDTA
  0.5 M Phosphate Buffer, pH 7.0

Store on bench indefinitely. If room gets too cool, it will precipitate out. Simply heat it at 65°C to redissolve.

- Wash solution I: 2x SSC
- Wash solution II: 2x SSC / 1 % SDS
- Wash solution III: 2x SSC / 0.1% SDS

- 20x SSC: 3 M NaCl
  0.3 M sodium citrate

  Mix, adjust pH to 7.0 with HCl, and store at 4°C

Isolation solutions:

- 3 M Sodium acetate (NaAc), adjusted with acetic acid to pH 5.2
- $Ethanol_{abs}$ ($EtOH_{abs}$) — Isopropanol —70% Ethanol
- E1 solution (cell resuspending)
  50 mM Tris-HCl, pH 8.0
  10 mM EDTA
  ▸ containing RNase: 100 µg/ml (stored at 4°C).
- E2 solution (cell lysis)
  200 mM NaOH
  1.0 % SDS
- E3 solution (neutralisation)
  3.1 M potassium acetate
  with acetic acid adjusted to pH 5.0

Sequencing:

- 5 % sequencing gel (36 cm glass plates):

21.0 g	urea (7 M)
8.4 ml	Acrylamide/Bis, 30% (29:1)
6.0 ml	10x TBE
20 ml	H_2O_{dd}
350 µl	10 % APS
15 µl	TEMED

  Buffer: 1x TBE

Transformation solution:

- DMSO
- FSB-buffer:

10 mM	potassium acetate,
(adjusted to pH 7.5 with 2 mM acetic acid)	
45 mM	$MnCl_2 \times 4H_2O$
10 mM	$CaCl_2 \times 2H_2O$
100 mM	KCl
3 mM	Hexaminecobalt chloride
10 %	glycerol

Sterilise the solution by filtration through a 0.45 µm (pore size) filter, dispense the solution into 50 ml aliquots and store at 4°C.

DNA-Markers:

- kb-ladder: (100 ng/µl): (MBI Fermentas)
  20.0 µl stock solution (0.5 mg/ml)
  16.6 µl loading buffer
  63.4 µl H_2O_{dd}
- λ-*Hind*III marker: (50 ng/µl): (MBI Fermentas)
  10.0 µl stock solution (0.5 mg/ml)
  16.6 µl loading buffer
  73.4 µl H_2O_{dd}
- PFGE marker: (NEB) A mixture of λ-DNA-*Hind*III fragments and λ concatemers embedded in 1% LMP agarose and 1-2 mm discs are applied to gels.

METHODS

3. Methods

3.1 Electrophoresis and DNA recovery from agarose gel

Agarose gel electrophoresis
0.7-2.0% gels were prepared by melting 0.7-2.0 g agarose in the presence of 1xTAE until the solution was clear. Electrophoresis was carried out in a horizontal configuration (Gibco) at 10 V/cm for about one hour.

Purification of DNA fragments using QIAquick Gel Extraction Kit
Several reactions such as ligation, sequencing, or PCR require purified fragments. An easy way is gel separation and recovery from agarose gel. The *QIA*quick Gel Extraction Kit was used; the desired DNA fragment was excised from the agarose gel and treated following manufacturer's protocol.

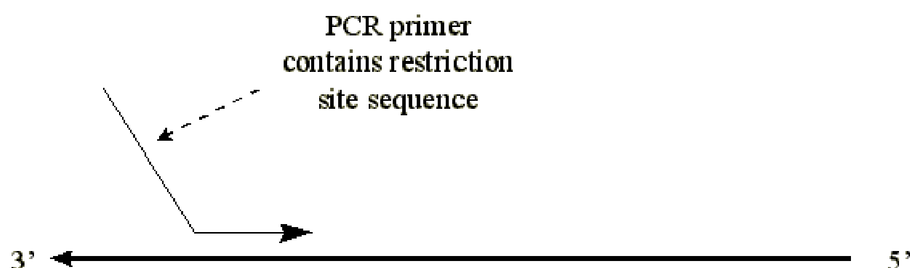
3.2 Restriction endonucleases

Restriction enzymes have the ability to recognise short DNA sequences and cut dsDNA at specific sequence dependent positions. Therefore, these proteins provide an essential tool in molecular biology. Generally these endonucleases are produced in bacteria and belong to a restriction/modification system.
For this work several different enzymes from different suppliers for several studies (restriction analysis to confirm results, Southern blot, cloning,...) were used.

3.3 PCR

PCR is an abbreviation for Polymerase Chain Reaction first reported in the mid-1980s by Mullis and colleagues (Saiki *et al.*, 1985; Saiki *et al.*, 1988). The technique enables the amplification of a specific DNA region within a known DNA sequence. This is achieved by the use of (oligonucleotide) primers which are complimentary to the defined sequence of the DNA template. A DNA polymerase enzyme, in suitable conditions in the presence of dNTPs, catalyses the synthesis of new DNA strands which are direct replicas of the target sequence.

Generation of restriction sites
Restriction sites can be used in different ways for cloning PCR fragments. The simplest way is to employ natural sites within the target DNA (should be in suitable positions). Another method involves the use of PCR to change the DNA sequence by introducing a new restriction endonuclease site (see diagram below).



In this way it is possible to introduce restriction site sequences into PCR products by having these sequences incorporated into the 5' end of the PCR primer(s). The short restriction site sequence on the 5' end of the PCR primer will not hybridise, and therefore may decrease the specificity of the PCR reaction at the beginning. As long as the 3' hybridising region is long enough, the primer will specifically bind to the appropriate site. Thus, the PCR product will have an additional DNA sequence at the 5' end which will contain the restriction enzyme site. A different restriction site sequence can be added via the other PCR primer. For ligation reactions, the PCR fragment can be inserted in a directional dependent manner in a host plasmid.

The following primers, containing the restriction sites for subcloning, were used to amplify homologous sequences for the targeting vector:

(1) XbaI-for and XbaI-rev (containing exon 1 and interrupted exon 2: size: 1.8 kb)
(2) ClaI-for and BamHI-rev (containing exon 3 and 4: size 2.4 kb)

and isogenic DNA derived from 129/Ola mice as template DNA.

Reaction:
PCR was performed in many cases. Important reactions will be illustrated later in the methods, where they are integrated. General protocol per 25 µl reaction:

mixture:	- PCR-buffer (10x)	1x (final concentration)
	- dNTPs (5 mM each)	0.8-1 mM (final conc.)
	- Primers (20 mM each)	0.8 µM each (final conc.)
	- *Taq*-DNA Polymerase	1-1.5 U
	- genomic DNA	150-300 ng or
	- BAC or plasmid DNA	15-25 ng

> for fragments up to 1.5 kb:
Cycle conditions were 30 sec at 94°C, 30 sec at the appropriate annealing temperature, and 1-2 min (regarding to the size) at 72°C for 30-35 cycles. PCR negative controls were performed by replacing the template DNA with H_2O_{dd}. The $MgCl_2$ concentration was between 1.5-1.75 mM.

> for fragments between 1.5 and 5 kb:
Cycle conditions were 30 sec at 94°C, 30 sec at the appropriate annealing temperature, and 2-3 min at 72°C for 10 cycles. The next 20-25 cycles contained additional 20 sec per cycle (e.g. n=11: 2`20", n=12: 2`40" etc) at 72°C. The $MgCl_2$ concentration was higher (2-2.25 mM)(negative controls see above).

Long distance PCR (up to 20 kb)
Primer pairs Tff1f 3'-UTR/Tff2r 5'-flanking and Tff2f 3'-UTR/Tff3r 5'-flanking (Table 4) were used for inter-gene PCR amplifications to determine the distances between the trefoil genes in mouse using the Expand™ Long Template PCR System (Roche Molecular Biochemicals) following manufacturer's protocol.

3.4 Ligation

During the ligation reaction, both the linearised vector and the insert of interest were mixed and ligated by the enzyme T4-DNA ligase. The vector was digested with the considered restriction enzyme and recovered after gel electrophoresis. The resulting fragment was dephosphorylated (if only one enzyme in use) to prevent self-ligation. The insert was prepared with the same restriction enzymes and purified by gel electrophoresis.
20-50 ng of the pre-treated vector were used per ligation reaction. The amount of insert was calculated individually. For ligation with overhangs, a 3-5 (BAC-DNA shot-gun: up to 10) to 1 molar insert to vector ratio was used. The entire ligation mix contained in addition to the mentioned components, 10x T4-DNA ligase buffer and 1.5 Weiss-units of the enzyme T4-DNA ligase. Ligation happened O/N at 4°C and the linearised vector without insert served as negative control.

TA-cloning of PCR products
Taq and other polymerases seem to have a terminal transferase ("extendase") activity which results in the non-template addition of a single nucleotide to the 3'-ends of PCR products (Clark, 1988). In the presence of all 4 dNTPs, dATP is preferentially added (Mead *et al.*, 1991). This terminal transferase activity is also the basis for the TA-cloning strategy: *Taq* polymerase is used to add a single dT to the 3'-ends of a blunt-cut cloning vector such as pGEM-T (Promega; Fig. 8A), and simple ligation of the PCR product into the "sticky-ended" plasmid.
This method was used to clone the primarily "XbaI-sited" PCR fragment and the "ClaI/BamHI- sited" PCR fragment of the mouse *Tff2* gene.

Verification of the clones
The insert or the sequence of interest was determined by restriction analysis and sequencing using the dideoxy chain termination method (refer to 3.13).

BAC-DNA shot-gun cloning
This method is used to subclone large BAC inserts to select the fragments of interest. For that reason 5 µg of BAC DNA (here: 123M10) was digested with 10 Units of *Pst*I+*Xho*I and *Xba*I+*Xho*I, respectively, with the appropriate buffer. After the electrophoresis (restriction check) the purified BAC DNA fragments were ligated with the pBK-CMV vector (Fig. 8B), which was prepared with the same restriction enzymes and purified by gel electrophoresis. For further preparation see 3.11 (colony hybridisation). After identifying of the desired fragments, they were transferred into the polylinker sites of pL2-neo (Fig. 8C) by an additional cloning step to complete the targeting construct.

3.5 Screening BAC library

BAC recombinants harbouring genomic mouse DNA from strain 129/Sv were obtained from Research Genetics, Huntsville, AL, USA. Oligonucleotides (Table 4) were designed from cDNA or genomic sequences (GenBank database accession No. Z21858, U78770, AJ 271002-271004; Chinery *et al.*, 1996b) to screen recombinants by PCR. These PCR products were also used as ^{32}P labelled probes to verify and characterise the recombinants by Southern blotting and hybridisation as described (Church and Gilbert, 1984).

PCR
Thirty cycles of PCR amplification were undertaken at 95°C for 30 sec and at the appropriate annealing temperature for 30 sec, as well as 72°C for 3 min. The purified PCR amplification products of 884 bp for *Tff1*, of 928 bp for *Tff2* and of 1190 bp for *Tff3* were then labelled with the random primed reaction (refer to 3.7).

BAC DNA Isolation
From an O/N culture of the BAC clones 0.5 ml was inoculated into 100 ml of LB medium with 12.5 µg/ml chloramphenicol and cultured at 37°C with vigorous shaking (~250 rpm) for 16 h.
For BAC DNA purification, the Qiagen Plasmid Midi Kit was used (alkaline lysis, (Birnboim and Doly, 1979) and the manual provided by manufacturer was followed. Quantity and quality of DNA are dependent on the BAC clone itself such as insert size etc.. DNA quality and quantity were tested by digesting it with restriction enzymes and running a 1% agarose gel for 1 h.

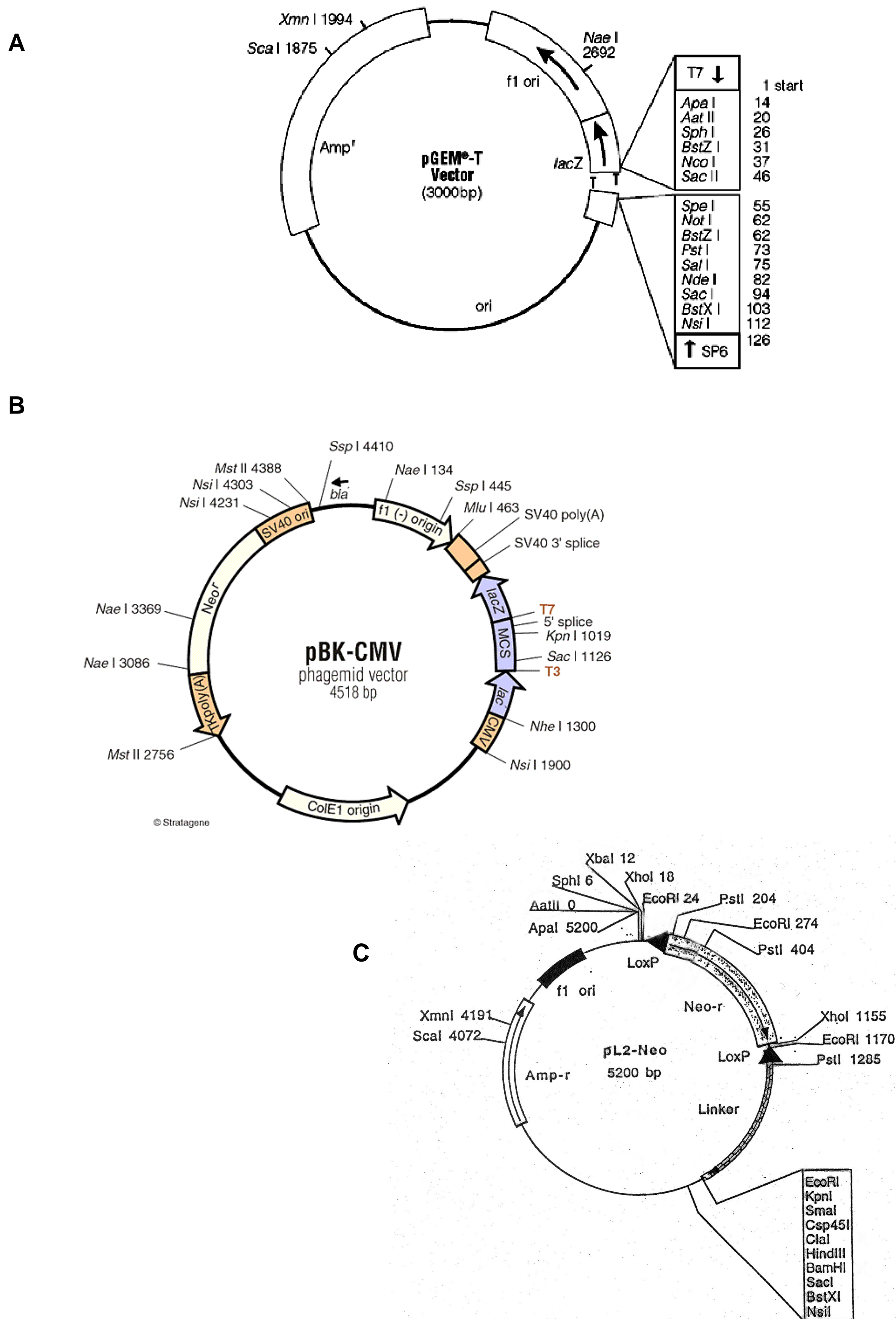


Figure 8: Vector maps of the T/A vector pGEM-T (A), the pBK-CMV vector (B), and pL2-neo vector used for the targeting construct (C).

3.6 Transformation

Electroporation

The preparation of cells for electroporation is easier than preparing chemical competent cells. Bacteria are grown to mid-log phase, chilled, centrifuged, and then washed three times with autoclaved H_2O_{dd} to reduce the ionic strength of the cell suspension. The bacteria (SURE strain, XL1-Blue MRF´strain) for electroporation (Dower *et al.*, 1988; Taketo, 1988) were prepared according to the protocol of "Current Protocols in Molecular Biology" detailed in `section 1.8` (Introduction of Plasmid DNA into Cells) (Ausubel *et al.*, 1995). One cell aliquot of 50 µl per transformation was used.

Chemical transformation

For this procedure, which was developed by Hanahan (Hanahan, 1983), the frozen competent cells were prepared and transferred according to the laboratory manual "Molecular Cloning" (Sambrook *et al.*, 1989) detailed in `section 1.76` using the FSB-buffer. A cell aliquot of 100 µl was used here.

Guiding principles:

Keep the cells as cold as possible during the whole procedure and manipulate the cells as gently as possible. The final ligation reaction, which combines the PCR or BAC-DNA fragments and the vectors (pGEM-T or pBK-CMV or pL2-neo), occurs in a very small volume containing ng amounts of vector plus insert product all mixed together. Electroporation uses strong, brief pulses of electric current to punch holes in the cell membranes of the *E. coli*. If 1 to 2 µl of the final ligation mix is added to the cells just prior to electroporation, a fraction of the cells will absorb some of the DNA through these holes.

Preparation and procedure for electroporation:

- Check that the cuvette holder has been placed in the freezer. Also check that the settings on the electroporator are set to 1.6 kV and a pulse duration of 5 msec.
- Fill ice box and place cuvette (1 mm-gap cuvettes are used), ligation mix, and electro-competent cells in box.
- Add 5-10 ng of the ligation mix to 50 µl of electro-competent cells and incubate for 5 min.
- Transfer the mixture of cells and ligation to the cold cuvette with a pipette; mix gently. Cap the cuvette and shake it down once to be sure that the cells are evenly along the bottom of the cuvette. Place it into the holder.
- Push the button for electroporation. The time constant should be close to 5 milliseconds. (If the cuvette pops it is bad. Add electro-competent cells to one of your extra cuvettes and try again. If it still pops there is probably too much salt).
- Immediately withdraw the cuvette and add the pre-warmed (37°C) 950 µl SOC to it; mix it directly, but gently; and transfer the cell mix into the 12-ml tubes.
- Leave the cells to recover in the 37°C shaking incubator for 60 min.
- Plate ~100 µl or a certain volume of the cells on pre-warmed selective LB-plates (with IPTG and X-gal, if blue/white selection is possible)

3.7 Random oligonucleotide-primed synthesis

Random oligonucleotide-primed synthesis is the method for producing uniformly radioactive DNA of high specific activity (Feinberg and Vogelstein, 1983) and for carrying out the labelling procedure. The DNA fragment containing the sequence of interest is purified by gel electrophoresis. Denature the resulting linear DNA molecules by boiling, anneal to random-sequence hexanucleotides, and then incubate with the Klenow fragment in the presence of dNTP. In this way, the hexanucleotides prime the DNA of interest at various positions along the template and are extended to generate double-stranded DNA allowing labelling of both strands.

Standard assay:

25 ng denatured DNA
3 µl dATP, dGTP, dTTP mixture (0.5 mmol/l)
2 µl reaction mix (containing random hexanucleotide oligos)
5 µl =50 µCi [α^{32}P] dCTP, 3000Ci/mmol
1,5 µl Klenow enzyme (2 U/µl)
make up to 20 µl with H_2O_{dd}

and incubate for 45 min at 37°C and stop the reaction by adding 2 µl 0.2 M EDTA (pH 8.0) and/or by heating to 65°C for 10 min.

When using labelled DNA for hybridisation, removal of non-incorporated dNTPs was performed by ethanol precipitation. Followed as below:

- add 2 µl tRNA (25 mg/ml) and 1/10 Vol. 3 M NaAc and 2 Vol. cold $EtOH_{abs}$
- precipitate for 45 min at –70°C
- centrifuge for 15 min at 13000 rpm, RT
- remove the fluid phase and wash the pellet with 70 % EtOH
- centrifuge for 5 min at 13000 rpm, RT
- remove the EtOH and resuspend in 100 µl 1xTE buffer

Before hybridisation, the probe was heated at 100°C for 10 min and cooled on ice.

3.8 Southern Blot

Southern blotting has been one of the milestones of DNA analysis since it was first described by E.M. Southern (Southern, 1975). This technique is used to transfer DNA from its position in the agarose gel to a nylon membrane placed directly above the gel. The DNA is denatured, neutralised, and transferred in a high-salt buffer by capillary action through the gel and into the papers above. The denatured DNA in the gel moves passively with the buffer and is stopped by the membrane. The denatured, single-stranded DNA binds to the membrane, is permanently bonded by cross-linking with UV-light, and is hybridised to a radio-labelled probe to detect hybridising DNA samples.

UV-crosslinking is recommended for a nylon membrane as this leads to covalent attachment and enables the membrane to be reprobed several times. The membrane must be completely dry before UV-crosslinking.

Procedure:
- Photograph gel with ruler
- Depurinate gel 15 min RT in 0.25 M HCl
- Denature gel 15 min RT in 0.5 M NaOH
- While denaturing, cut nylon membrane to fit gel, cut 3MM (Whatman-) paper (3-4 sheets) to size of the gel, prepare appropriate size paper towels for blotting.
- Set up blot:
  - two layers of 3MM paper as wicks into the transfer buffer
  - gel on top of wicks, remove any bubbles, pre-wetted membrane on top of gel
  - 3MM paper on top of membrane
  - paper towels on top of 3MM paper
  - 250 g weight on top of paper towels
  - 6-18 h to blot (size dependent)
- Neutralise: 3-5 min , RT, shaking with the 50 mM phosphate buffer
- Air dry membrane, UV-crosslinking
- Ready for the hybridisation experiments

Southern blot was used for different experiments. Generally, 10-15 µg DNA was digested with appropriate restriction enzymes, if genomic DNA was investigated, and 5 µg in the case of BAC recombinants.

3.9 Spot hybridisation

In comparison with Southern blot anaylsis, this method provides a more rapid and less time-consuming determination of the presence of specific DNA in samples (i.e. clones). First, a direct transfer of denatured DNA to "spots" or "dots" on filters or nylon membranes is made and the spotted membrane is then hybridised with a labelled probe. The dot blots should be prepared on nylon membrane or filter from the batch that was used for Southern transfers.

Reagents:
2M NaOH, nylon membrane, positive control DNA, radio-labelled probe (mainly 928 bp-fragment of 5'-flanking region of *Tff*2), solutions are needed for hybridisation.

Procedure:
- Add 2 µl 2M NaOH to 1 µl (0.5-1 µg) DNA sample, leave it for 10 min at RT (denaturing phase)
- Cut nylon membrane and make marks as usual
- Spot the DNA solution onto the membrane marked with circular shapes
- When the solution is spotted down, let the spots air dry, rinse the membrane in phosphate buffer and then cross-link the DNA to the membrane
- The filters are ready for the hybridisation process

3.10 Hybridisation

All hybridisation methods depend upon the ability of denatured DNA to re-anneal when complementary strands are present in an environment near but below their T_m. In a hybridisation reaction with double stranded DNA on a nylon membrane and a single-stranded DNA probe, three different annealing reactions occur. First, there are the desired probe-template DNA interactions which result in a signal. Second, there are mismatch interactions that take place between related but non-homologous sequences. These mismatch hybrids are the ones that will be eliminated during the washing of the membranes. Non sequence-specific interactions also occur as background noise signals. The ability to extract information from a particular hybridisation experiment is a function of signal-to-noise ratio.

Procedure:
To prevent the non-specific binding pre-hybridisation of the nylon membrane should be carried out at 65°C for a few hrs. Add labelled probe to 20 ml Hyb-mix and hybridise O/N at 65°C in the hybridisation oven. Hybridisations are performed in the Hyb-mix buffer.

High stringency conditions are necessary for washing the blots:
-2 times for 5 min in Wash solution I at RT
-2 times for 15 min in Wash solution II at 65°C
-2 times for 5 min in Wash solution III at 65°C (optional)

The membranes must be exposed for autoradiography at –70°C with an intensifying screen and developed after 1-5 days (depending on the experiments). Or radioactivity on the membranes is visualised by phosphor imaging on Personal Molecular Imager FX (refer to 3.16). Quantity One analysis software should be used to quantify bands on images generated with the Imager System.

Stripping probes from blots:
-incubate in the hybridisation oven at 80°C for a few hrs in 20 ml of Wash solution II
-rinse blot thoroughly in distilled water tap.

3.11 Colony hybridisation

BAC-DNA shot-gun cloning
The detection of desired fragments out of complex genomes requires shotgun cloning (or subcloning) of BACs into vectors that carry smaller inserts and that can serve as templates for generating targeting vectors.

By using the colony hybridisation method (Grunstein and Hogness, 1975), a very large number of colonies of *E. coli* carrying different plasmids can be rapidly screened to determine which plasmid contains the specific DNA fragment. The colonies to be screened are formed on nitrocellulose filters or nylon membranes. After a reference set of these colonies is prepared by replica plating, the cells are lysed and their DNA is denatured and fixed to the filter *in situ*. The resulting DNA-prints of the colonies are then hybridised to a radioactive probe that defines the sequence of interest, and the result of this hybridisation is assayed by autoradiography. Colonies whose DNA-prints exhibit hybridisation can then be picked from the reference plate.

This method is employed to isolate and characterise, respectively, clones of the BAC 123M10 shot-gun plasmids that contain the desired 5'-part sequences (PstI-XhoI fragment or XbaI-XhoI fragment) of *Tff2*. For screening/hybridisation of the clones, including the 5'lanking DNA of *Tff2*, the 928 bp probe (5'-DNA and exon 1 of *Tff2*) is used.
Probe preparation: The DNA is labelled by random primed method (refer to 3.7) with [α^{32}P] dCTP (3000 Ci/mmol).

Material & Method:
- Butterfly membrane for colony lifts (with write-on tabs and orientation notches)
- Prepare the solutions as follows:
  - *solution 1:* 1.5 M NaCl, 0.5 M NaOH, 0.1 % SDS (denaturing & lysis)
  - *solution 2:* 1 M Tris-HCl, pH 7.4 (neutralisation)
  - *solution 3:* 1.5 M NaCl, 0.5 M Tris-HCl, pH 7.4 (neutralisation)
- Cut three pieces of Whatman 3MM filter paper to the appropriate size and fit them into the bottoms of three glass trays.
- Saturate the pieces of 3MM filter paper in each of the solutions prepared before.
- Place the (82 mm disc) membrane on the surface of the agar plate until completely wet (don't forget to mark the orientation).
- Afterwards the membrane is placed colony side up on filter paper saturated with solution 1 for 7 min.
- Transfer the membrane to 3MM filter paper saturated with solution 2 for 5 min, then place on filter paper saturated with solution 3 for 5 min.
- Place the membranes, colony side up, on a dry sheet of 3MM filter paper. Allow the filters to dry for a minimum of 30 min at RT.
- DNA was fixed to the membranes by UV irradiation with a Stratalinker™ (120 mJoules/cm^2 at 254 nm for 3-5 mm diameter colonies).

3.12 DNA extraction

Ethanol Precipitation
Procedure:
- Measure the volume of DNA containing sample. If the volume is less than 450 µl, place it in a 1.5 ml eppendorf tube.
- Add 0.1 Vol. of 3M NaAc pH 5.2 (mix) and afterwards 2 Vol. of cold $EtOH_{abs}$; mix.

- Place the sample at –20°C (if there is a lot of DNA, you can proceed almost immediately. If there is only a few hundred ng, it is best to wait at least one hour or O/N) or at –70°C for 1 hr.
- Centrifuge at 14000 rpm, 4°C for 20 min
- Decant the supernatant, and tap the tube (upside down) on a paper towel to remove as much EtOH as possible. Keep the tube upside down (take care: you may loose the pellet, if detached)
- Wash it with 70 % EtOH, centrifuge (see above)
- Remove the EtOH again, dry in Speed-Vac for 5 min
- Resuspend in an appropriate volume of TE buffer or H_2O_{dd}

Isolation of plasmid-DNA

➢ *Mini-prep*

Bacterial cells are lysed in NaOH/SDS to denature proteins and chromosomal DNA (Birnboim and Doly, 1979).

Procedure:

- Grow 3 ml over night (O/N) culture with appropriate antibiotics
- 1.5 ml into microfuge tube, centrifuge 1 min max. speed. Keep pellet
- Add 250 µl of E1 solution, vortex briefly
- After pellet is in solution, add 250 µl of E2 solution, mix by inverting, RT 5 min
- Add 250 µl ice cold of E3 solution, mix by inverting, and for 10 min on ice
- Centrifuge (14000 rpm), 20 min at 4°C
- Transfer supernatant to new tube, add 0.7 Vol. isopropanol, vortex
- Centrifuge (14000 rpm), 20 min at 4°C
- Remove all isopropanol and keep visible DNA pellet
- Wash: add 150 µl of 70% EtOH
- Centrifuge (14000 rpm), 20 min at 4°C
- Dry down
- Resuspend in 30 µl TE pH 8.0 or H_2O_{dd}

➢ *Midi-prep*

This procedure is the method of choice for isolating large amount of double stranded plasmid-based templates for restriction or sequence analysis. It employs a modified alkaline/SDS method to prepare the cleared lysate. After neutralisation, the lysate is applied onto a column and the plasmid DNA is bound to the anion exchange resin. Washing the resin removes RNA and all other impurities. Afterwards, the purified plasmid DNA is eluted from the column and finally concentrated by an alcohol precipitation.
For this purpose, the Jetstar Plasmid Purification System was used and following the manual provided by manufacturer (Genomed, D).

3.13 Sequencing analysis

Two main sequencing techniques are in common use: the chemical degradation method of (Maxam and Gilbert, 1977) and the enzymatic (dideoxy chain termination) method (Sanger *et al.*, 1974).
The enzymatic method is by far the most popular and widely used technique for DNA sequence determination. Thermal cycle sequencing is a method of dideoxy sequencing in which template DNA is amplified in linear fashion using a single primer and thermostable DNA polymerase. This technique has several advantages as compared with conventional dideoxy sequencing. Since the primer extension reaction is performed not just once but 20-35 times, a smaller amount of template DNA is required (Krishnan *et al.*, 1991).

Various plasmid DNAs were sequenced using different primers (Table 4), generally starting approximately 30-50 bp away from the sequence of interest. The reaction was prepared as following:

1 µg plasmid-DNA or
0.3-0.5 µg BAC-DNA
mix:
0.5 µl Primer (10 pmol/µl)
4.0 µl Sequence-mix
make up to 20 µl with H_2O_{dd}

Cycle sequencing:
Program: 96^oC 10 sec
50^oC 5 sec
60^oC 4 min
cycle: n=25
4^oC ∞

Samples were run under denaturing conditions in the presence of 1x TBE.
Running conditions: pre-run: 2500 V, 50 mA, 40 W for 30 min at 51^oC
main-run: 2500 V, 50 mA, 40 W for 8 h at 51^oC

3.14 Gene targeting

3.14.1 Construction of a replacement vector

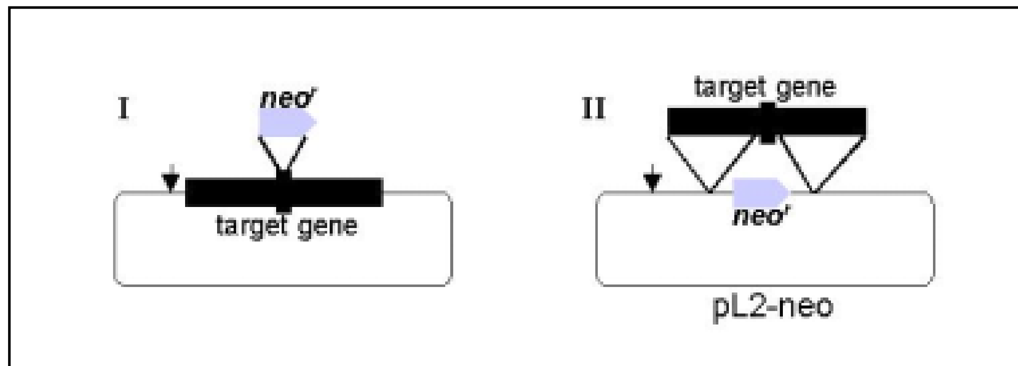


Figure 9: Two strategies to create a replacement construct. The arrows indicate the point to linearise.

The targeting vector is designed to recombine with and mutate a specific chromosomal locus. The minimal components are sequences (availability of a genomic clone containing the gene of interest) which are homologous with the desired chromosomal integration site and a plasmid backbone.

In the first method (Fig. 9; **I**), the target gene fragment is subcloned into a plasmid vector; the neo is inserted into a rare restriction enzyme site in the target gene fragment. In method **II**, which was used for this work, the target gene fragment was cleaved into two pieces that were subcloned into the polylinker sites of pL2-neo (Fig. 9). Note that the relative orientation of homologous fragments in the construct must retain the one found in the target gene.

The following are guidelines for the construction which generates an easily identifiable null allele:
* use a fragment of isogenic homologous sequences of 5-8 kb (here: 5.45 kb)
* insert the *neo*^{r} gene into an upstream exon or important 5' exons
* interrupt or delete exons without a unit number of codons to avoid generating a protein product with partial or novel function after RNA splicing
* for PCR: clone the selection marker so that one arm of homology is 0.5-2 kb
* linearise the vector outside the homologous sequences
* design a diagnostic Southern screening strategy (here: *BamH*I and *Xba*I)

Figure 10 describes procedures involved in gene-targeting and the generating of knockout mice (an overview).

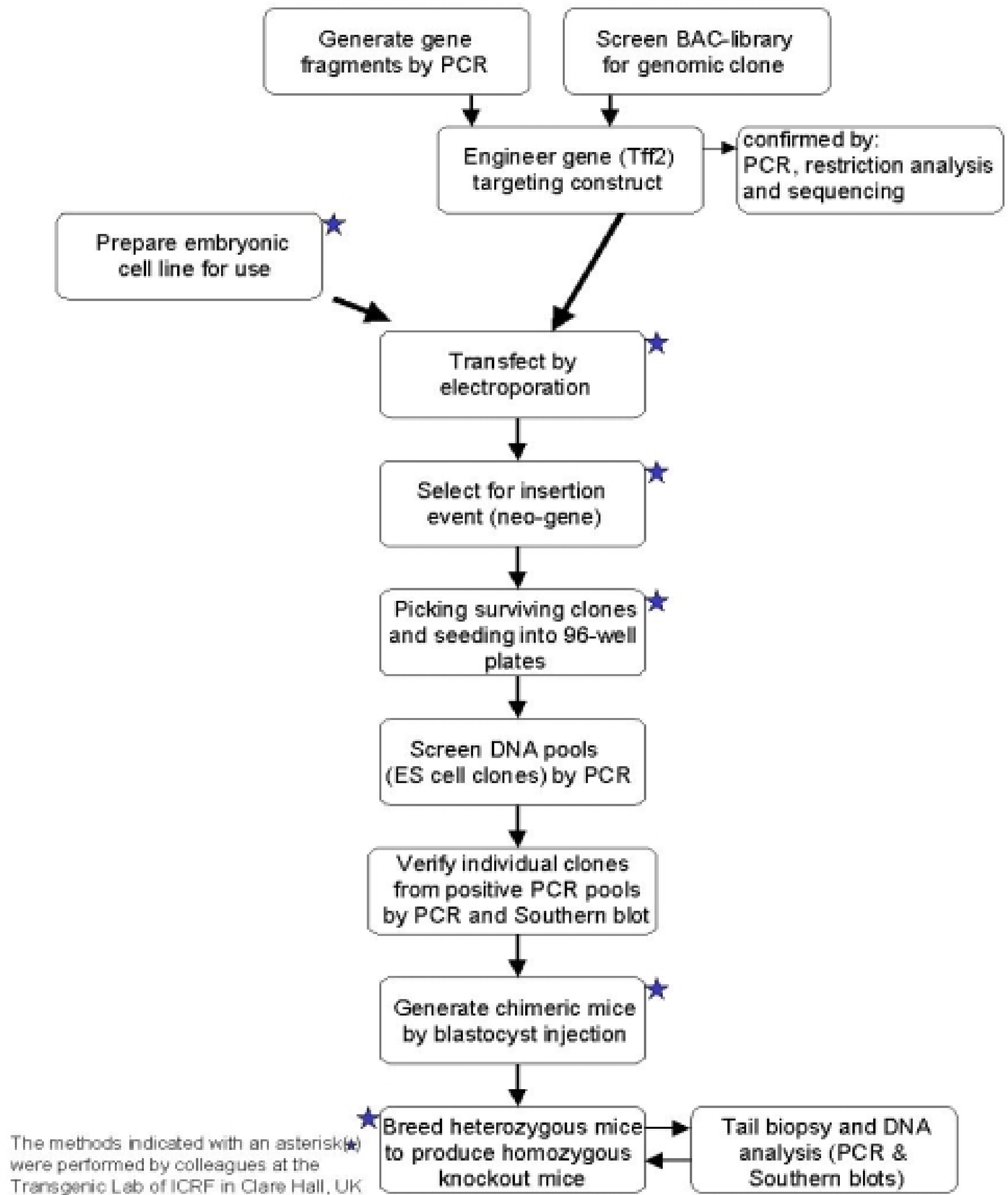


Figure 10 : Overview; procedures involved in the generating of knockout mice.

3.14.2 Detect gene targeting events in ES cells

Embryonic stem cells

The embryonic stem (ES) cells initially used were derived from 129/Ola mice and are commonly called GK129 cells. A second ES cell line, called AB2.2, was also employed, derived from 129/Sv mice (new nomenclature: 129/S1; Simpson *et al.*, 1997). AB2.2 cells are commercially available from Lexicon Genetics, TX, USA.

Positive Selection

The neo^r cassette, placed internally into the targeting vector, serves as a marker for positive selection as well as a potential inactivator of the targeted gene (DeChiara *et al.*, 1990). The construct, which had been transfected into ES cells in culture, positively selected cells, were further analysed. The neo^r gene, encoding a bacterial aminoglycoside phosphotransferase, renders a given cell resistant to the antibiotic G418 (Colbere-Garapin *et al.*, 1981). To isolate cells carrying a targeted mutation, all the cells are cultivated in a G418 media (200-275 µg/ml G418 for 250 ml ES media).

Picking and Expansion of Colonies after Electroporation

After electroporation and selection, the ES cell colonies take 8-12 days of growth to become visible with the eye and can be picked at this time. The ES cells clones are seeded into 96-well plates and grown for 5 days. Then, one-half of the cells are frozen in the original plate while remainders are grown in a replica plate for PCR and Southern blot analysis.

3.14.3 Detection of homologous recombinants

PCR analysis using DNA prepared directly on the 96-well plates

For the DNA extraction, the cells are rinsed with PBS and lysed by proteinase K in the plate. The rest of the extraction steps are carried out in the same plate, discarding the solutions by gentle inversion and leaving the nucleic acids attached to the plate. DNA is precipitated by adding 50 µl isopropanol to each well and by allowing the plate to stand for 30 min at RT. After centrifugation (1200g, 5 min at RT) empty the wells and wash the DNA three times by filling and emptying the wells with 70% EtOH. Finally, allow the wells to dry for 20 min at RT and resuspend DNA in 30 µl H_2O_{dd}.

Screening for homologous recombinants was first undertaken with pools of ES cell clones. The screening method of choice was PCR (very sensitive and rapid). It could also have been performed on crude cell lysates (described above). For the detection of a gene targeting event, the homologous recombined DNA region was amplified with two primers: one specific for the mutation (targeting vector: Reko-for) and the other for the endogenous (non-vector) DNA (WT-rev). In this way, a correctly sized PCR product (2.4 kb) would only be generated if two primer sites were juxtaposed by homologous recombination.

A PCR screening has the advantage that it can be conducted with minimal expansion of ES clones from cells in 96-well plates (Ramirez-Solis *et al.*, 1992). Possible recombinant clones are further verified for the correct integration by Southern blot analysis. Homologous recombination should lead to the generation of a additional restriction fragment, readily distinguishable from a fragment of the endogenous copy. Sufficient DNA can be obtained for this analysis from colonies of 12- or 24-well plates.

PCR *Taq* polymerase was used in the presence of 2.25 mM $MgCl_2$. Cycle conditions were 94°C-4 min; 94°C-30 sec, 59°C-30 sec, 72°C-3 min (10 cycles); 94°C-30 sec, 59°C-30 sec, 72°C-3 min plus an additional 20 sec per cycle (25 cycles); 72°C-7 min. Products were analysed on 0.7% agarose gels after ethidium bromide staining. To detect the homologous recombined events the primer pair Reko-for/WT-rev (expected size: 2.4 kb) was used. By using WT-for/WT-rev (~2.65 kb) primers the quality of DNA was tested, as well as a control.

For Southern blots, 15 µg of digested (I: *BamH*I and II: *Xba*I) ES cell DNA was subjected to agarose (0.7%) gel electrophoresis and transferred to nylon membranes. Probes were labelled by random priming using [α^{32}P] dCTP. The 928 bp fragment (Fig. 24) of the *Tff2* gene was used as a probe.

3.14.4 ES cell injection

Clones identified as positives were thawed, expanded, and micro-injected into C57BL6/J day 4 embryos (blastocysts). The colleagues at the transgenic lab routinely inject 36 blastocysts per clone and usually inject two positive clones per project.
Injected blastocysts were transferred into day 3 pseudo-pregnant female mice (CBAx C57BL6/J females for pseudo-pregnants were used).

3.14.5 Chimeras

The generation of chimeras between ES cell clones and embryos is an essential step in these processes which, when successful, leads to the derivation of new strains of mice with the desired altered genome. Approximately 18-21 days after transfer pups will be born; about 7-10 days after birth they can be phenotyped for coat colour contribution and sexed.
The evaluation of chimeric animals is necessarily subjective, but in general, the degree of coat colour chimerism of a particular animal correlates with the degree of germline contribution. To determine whether a chimera shows germline transmission, the 129/Sv-C57BL6/J chimeras are backcrossed with C57BL6/J mice (breeding of heterozygous strains).

3.14.6 Mouse tail biopsies

Germline transmission was determined by DNA typing (Southern blot analysis) of tail biopsies.

DNA extraction:

- Remove 0.5 cm of tail into polypropylene microfuge tube (do not mince).
- Add 0.5 ml digestion buffer with 25 µl of a 10 mg/ml proteinase K

➢ digestion buffer:	50 mM	Tris-HCl
(pH 8.0)	100 mM	EDTA
	100 mM	NaCl
	1 %	SDS

- Incubate overnight at 50-55 °C with gentle shaking.
- Add 0.7 ml neutralised phenol/chloroform/isoamyl alcohol (25:24:1).
- Mix fairly vigorously (do not vortex).
- Spin in microfuge at top speed 10 minutes and transfer the upper phase to new microfuge tube.
- Add 2 Vol. $EtOH_{abs}$ at RT and invert until DNA precipitate forms.
- Spin in microfuge 10 minutes and carefully remove and discard supernatant.
- Add 0.5 ml 70% EtOH (-20 °C) and invert several times.
- Spin in microfuge 10 minutes and carefully remove and discard supernatant.
- Air dry at room temperature.
- Add 100 µl TE buffer and incubate at 65°C for 15 minutes to resuspend DNA.

The total yield is approximately 20-50 µg DNA. An amount of 10-15 µg was used for restriction enzyme (*BamH*I) digestion to confirm the germline transfer.

3.15 Pulsed-Field Gel Electrophoresis (PFGE)

DNA molecules larger than 25 kb in size must squeeze to get through even the largest pores and cannot be sieved by a gel. Molecules in the size range all migrate at about the same rate, called “limiting mobility”.
A solution to this problem was found in 1984 when Schwartz & Cantor reported the development of pulsed-field gel electrophoresis (Schwartz and Cantor, 1984). Here, the CHEF-PFGE was used (contour-clamped homogeneous electric fields; Chu *et al.*, 1986).
Procedure:

Day 1: setting up the pulsed field gel

BAC DNA digested by rare cutters like *Mlu*I or *Not*I or *Sal*I was separated with 1 % agarose gel in 0.25x TBE using LKB 2015 pulsator electrophoresis unit (Pharmacia).

Program electrophoresis parameter: a linear pulse time gradient from 5 to 15 sec was applied for 20 h at 180 V constant voltage.

Day 2: staining and documenting the gel and preparing of Southern blot .

3.16 Autoradiography

3.16.1 Exposure to X-ray film

Preparing the membrane for autoradiography:
In a darkroom, place the nylon membrane (wrapped in Saran wrap) in a light-tight exposure cassette and cover it with a sheet of X-ray film. Expose the film for an appropriate length of time. When intensifying screens are used, the film must be exposed at –70°C. The low temperature stabilises the silver atoms and ions that form the latent image of the radioactive source.

X-ray film may be developed by hand as follows:

Developer:	3-5 min
water (stop bath):	1 min
Fixer:	5 min
Running water:	10 min

The temperatures of all solution should be 18-20°C.

3.16.2 Using storage phosphor and phosphor imaging

In this case the use of photostimulable storage phosphor imaging plates was described as an alternative to film for recording and quantifying autoradiographic images (Personal Molecular Imager FX from Bio-Rad). The scanner can scan at resolutions of 50, 100, 200 and 800 microns. The advantages of storage phosphors over film are a linear dynamic range covering five orders of magnitude and a sensitivity 10 to 250 times that of an X-ray film.
The imaging plates are composed of fine crystals of BaFBr: Eu^{+2} in an organic binder (Amemiya and Miyahara, 1988). High energy radiation (e.g., X-rays, ultraviolet light, gamma rays or beta radiation from a hybridised membrane) will excite an electron of the Eu^{+2} ion into the conduction band. This electron is then trapped in an "F-center" of the $BaFBr^-$ complex with a resultant oxidation of Eu^{+2} to Eu^{+3}. The excited $BaFBr^-$ complex exhibits a distinct absorption band centred around 600 nm. By exposing the excited complex to light from a laser (635 nm), the electrons are liberated back to the conduction band reducing Eu^{+3} to Eu^{+2*}. Eu^{+2*} then releases a photon at 390 nm as it returns to ground state (Sonoda *et al.*, 1983). Trapped energy is released as photons that are in turn captured by a photomultiplier tube.
The analysis software (Quantity One, BioRad) permits control of the scanning system and accurate analysis of the captured image or data.

RESULTS

4. Results

4.1 Construction of the targeting vector for disruption of *Tff2*

The main aim of this project was not only to establish a targeting vector, that disrupts the *Tff2* gene by removing a large part of exon 2, but to characterise a successful germline transmission of this mutant allele to heterozygous (*Tff2*$^{+/-}$) mice.
The target gene was cleaved into two pieces, a proximal part containing exon 1 with small part of exon 2 and the distal part with exon 3 and 4 (strategy I, Fig. 11). Afterwards they were subcloned into the vector pL2-neo, containing a neomycin cassette flanked by *loxP* sites for further selection (G418) and interrupting the desired gene. It was not intended to use Cre-mediated recombination of loxP to delete the neo-gene. However, if it were found to affect the expression (Olson *et al.*, 1996; Pham *et al.*, 1996) of neighbouring *Tff* genes, the neo gene could be simply removed by crossing the *Tff2* knockout strain with a Cre-mouse. Figure 11 shows the genomic arrangement of the *Tff2* gene, which contains four exons (exons 2 and 3 coding for the trefoil motifs), and the strategy to develop the *Tff2* knockout by removing a sufficient part of exon 2.
The sequence data was obtained from GenBank (accession No. U78770) and personal communication (O. Lefebvre).

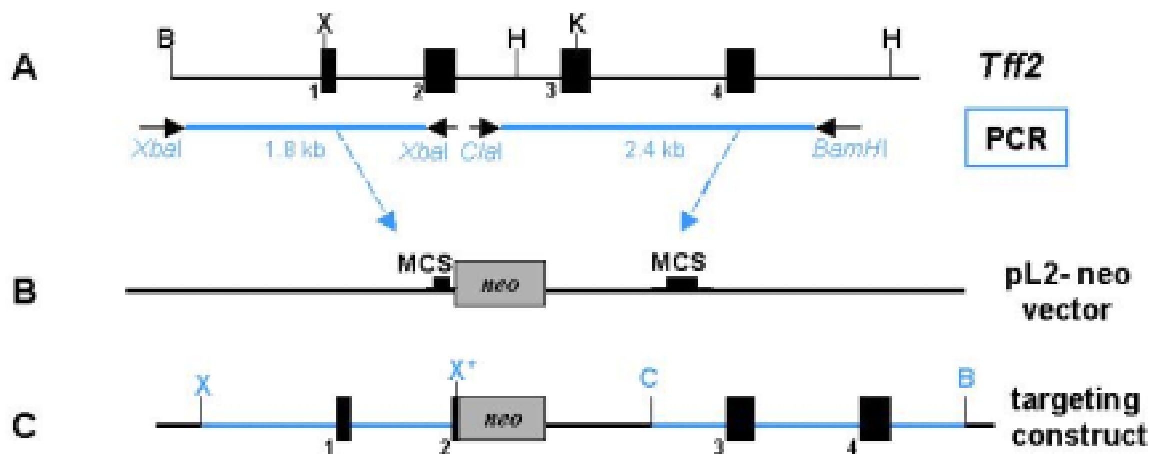


Figure 11: Construction strategy of the first targeting vector. (A) The *Tff2* gene locus. Ligation of PCR-products using modified primers with pL2-neo (B) results in targeting construct (C). The horizontal arrows indicate primer orientation. Solid black boxes indicate the four exons of *Tff2*. B, BamHI; X, XhoI; X/X* (blue) XbaI; E, EcoRI; H, HindIII; K, KpnI; MCS, multiple cloning site.

The *Tff2* gene was amplified by using the designed primers m2-pro and m2-end (4.9 kb: Fig. 12 A). A 2.4 kb genomic DNA fragment (Fig. 12 B) was generated by PCR with the use of ClaI-for and BamHI-rev primers, containing corresponding restriction sites. A second fragment (1.8 kb) was obtained with the primers XbaI-for and XbaI-rev (Fig. 12 C).

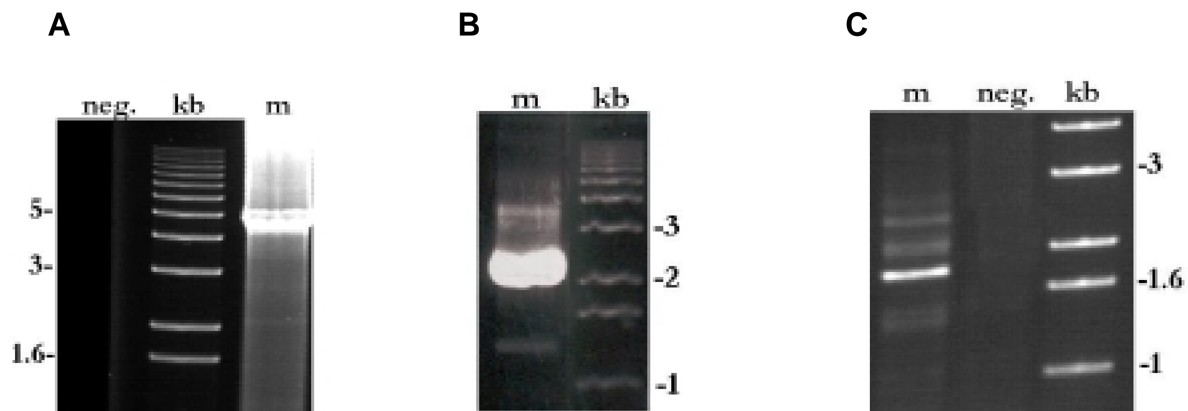


Figure 12: PCR amplification of the *Tff2* gene (A), the ClaI-BamHI fragment (B) and XbaI-XbaI product using mouse genomic DNA (m) derived from 129 strain (neg.: negative control, kb: molecular marker).

The ClaI-BamHI (CB) fragment was subcloned into the T/A-cloning site of the pGEM-T vector (Promega), generating the plasmid pT1-CB. After restriction with the appropriate enzymes (*Cla*I and *BamH*I), it had to be subcloned into the ClaI and BamHI sites of the pL2-neo vector. Initially, it was not possible to excise this insert. Detailed analysis revealed restriction by *BamH*I but not by the enzyme *Cla*I (see circular and supercoiled form in Fig. 13).

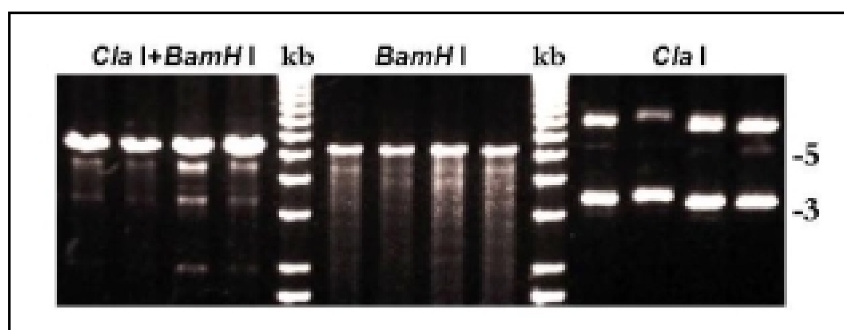


Figure 13: Restriction of the plasmid pT1-CB using *Cla*I and *BamH*I enzymes.

However, sequencing of the plasmid demonstrated the existence of both restriction sites (Fig. 14).

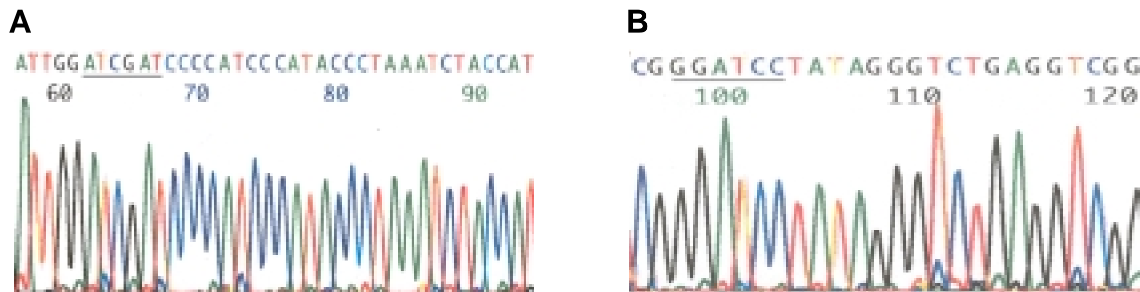


Figure 14: Sequence of the ligation boundaries of pT1-CB using M13-primers. The location of restriction sites for ClaI (A: ATCGAT) and BamHI (B: GGATCC) are shown underlined.

The answer to this problem was the *dam* gene (Dam methylase) of the *E.coli* host strain. It transfers a methyl group from S-adenosylmethionine to the N^6 position of the adenine residues in the sequence "GATC" and, therefore, restriction enzyme cleavage by *Cla*I (methylation sensitive) is blocked. By using Dam^- strains (GM272) to grow the pT1-CB shuttle vector, the modification-dependent restriction was avoided (Fig. 15).

Figure 15: Restriction enzyme cleavage of pT1-CB (v-vector; i-insert).

After excision and preparation of the insert, the CB-fragment was then cloned into the pL2-neo vector, generating the plasmid pL2-CB (size: 7.6 kb). This was confirmed by restriction analyses using different enzymes and sequencing (Fig. 16). The sequence data revealed the expected sites of ligation and the sizes of the restriction fragments.

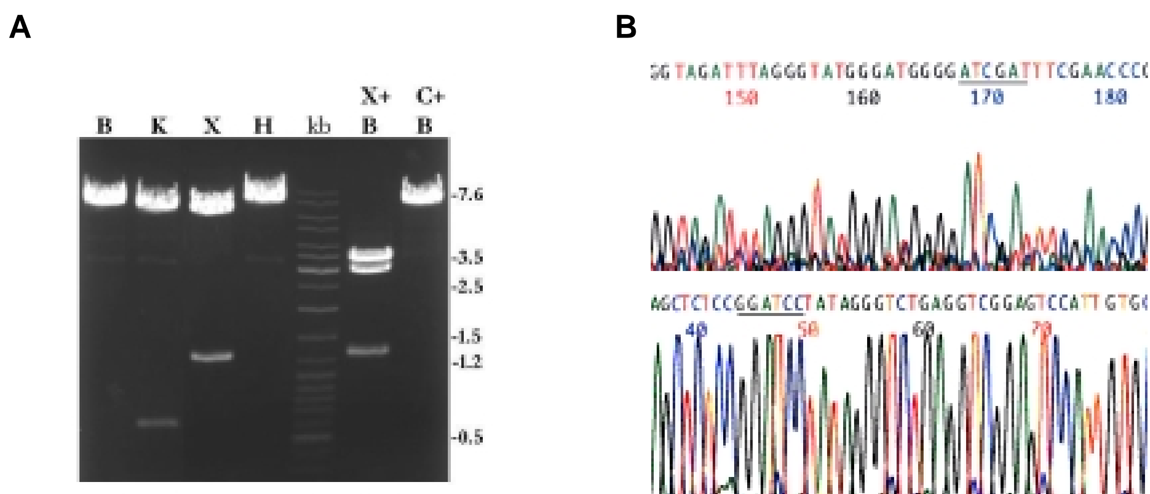


Figure 16: Restriction (A) and sequence (B) analysis of pL2-CB. The clone was digested with *BamH*I, *Kpn*I, *Xho*I, *Hind*III and double digested with *Xho*I+*BamH*I as well as *Cla*I+*BamH*I. The underlined sequences belong to the restriction sites of ClaI (ATCGAT) and BamHI (GGATCC). Note that in host XL1-Blue MRF´ digestion with *Cla*I is not visible because of dam-methylation.

Another difficulty arose during the cloning of the upstream part of the target gene. Despite different modifications of the ligation and/or transformation parameters, the XbaI-XbaI fragment (1.8 kb) could not be subcloned into the pGEM-T vector (Promega). Interestingly, instead of the desired clone, plasmids with changed insert sizes appeared. We concluded that this was due to a recombination event triggered by the pyrimidin-rich repeat cluster in intron 1 (total array length: ~340 bp; repeat unit: CTTT and CTTC). This region probably leads to genetic instability in *E.coli*. In support of this idea is the fact that it was easy to clone a 928 bp-fragment adjacent to the repeat cluster containing 5'-flanking DNA, exon 1 and part of intron 1 (Fig. 17).

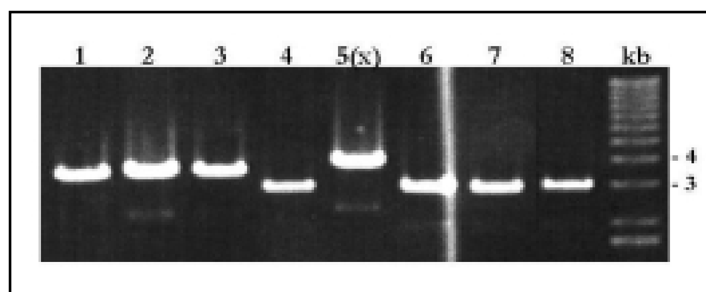


Figure 17: Restriction of different plasmid-DNA by *Sal*I. Lanes 1-3 correspond to the positive control (512 bp fragment: total size 3.5 kb) supplied by the manufacturer and lanes 4, 6-8 to the vector only (3 kb). Lane 5 (**x**) shows the ligated 928 bp fragment of *Tff2* (total size: ~4 kb).

To obtain more sequence data from *Tff2*, two pairs of primers (no. 1: m2-pro&m2-250rev and no. 2: m2/251-for&m2/1189-rev) were designed to amplify exon 1 or 5'-flanking DNA, respectively. These primers were then used for PCR screening of a 129/Sv genomic DNA BAC library. Three BAC recombinants, 123M10 - 517E16 - 566K2, harbouring genomic mouse DNA were obtained from Research Genetics.
Oligonucleotides were designed from cDNA or genomic sequences (GenBank accession No. Z21858, U78770, AJ 271002-271004; Chinery *et al.*, 1996b) to screen the recombinants by PCR (Fig. 18)(for a more detailed investigation and characterisation of BACs see 4.2).

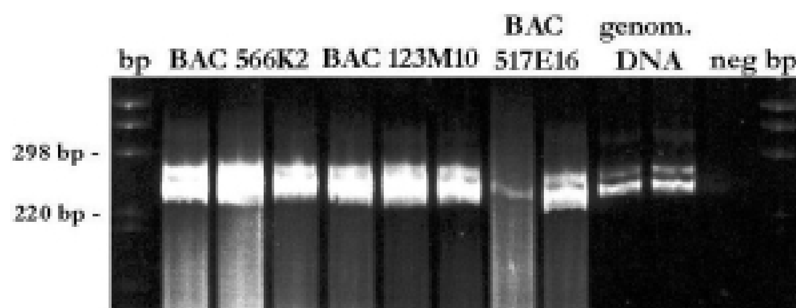


Figure 18: PCR amplification of a 250 bp product from 5'-flanking DNA of *Tff2* using different BAC- and mouse genomic DNA (neg, negative control).

The 928 bp PCR product (Fig. 17) was used as a ^{32}P labelled probe to verify and characterise the BAC recombinants by Southern blotting and hybridisation (Fig 19). BAC123M10 and BAC517E16 were double-digested with different enzymes to locate the desired 5'-flanking DNA fragment of the trefoil gene *Tff2* to create the revised construct as outlined in strategy II, Fig. 24.
Because more sequence structure data were obtained for the 5' flanking region of *Tff2*, the new targeting vector would exclude exon 2, intron 1 (containing the repeat cluster) and would take advantage of the XhoI-site in exon 1 (Fig. 11).
Fragments produced through the *Pst*I-*Xho*I or *Xba*I-*Xho*I (Fig. 19) digestion were of major interest, since the sizes (3.0 and 5.0 kb) were suitable for achieving a total size of homologues sequences of at least 5 kb.

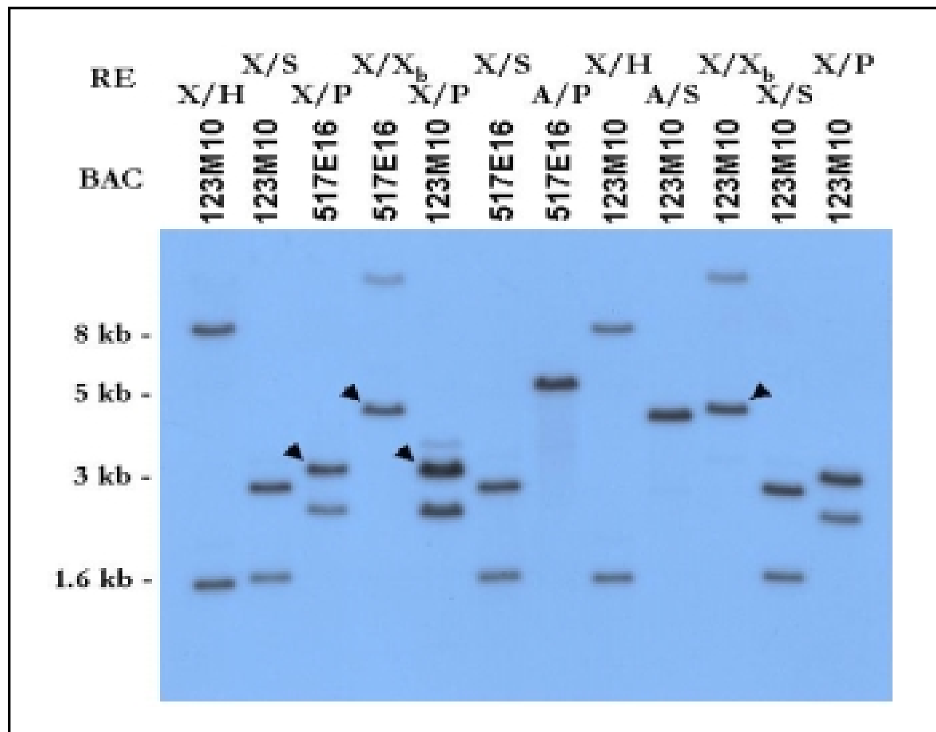


Figure 19: Autoradiograph of the Southern blot of two BACs (123M10 and 517E16), double digested with restriction enzymes (RE), showing the hybridisation with the 928 bp probe. Arrowheads indicate the signal of fragments useful for the targeting vector (X *Xho*I; H, *Hind*III; S, *Sac*I; P, *Pst*I; X_b *Xba*I; A, *Apa*I).

In a first assay the PstI-XhoI fragment (~3.0 kb) containing the first three codons of exon 1 was subcloned by shotgun cloning into the PstI-XhoI sites pBK-CMV (Stratagene) vector (Fig. 20). Two plasmids were generated, which were confirmed by restriction analysis and spot hybridisation.

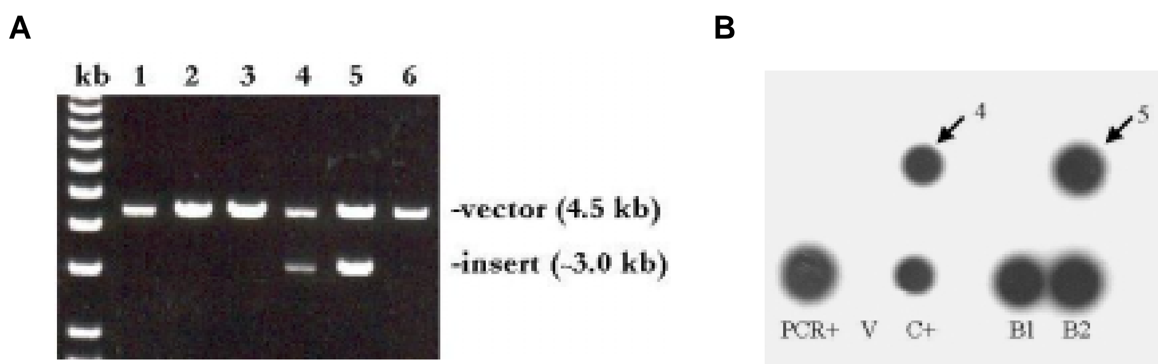


Figure 20: (A) pBK-CMV vectors digested with *Pst*I and *Xho*I. Lanes 4 and 5 display the desired insert. (B) Spot hybridisation: after miniprep extraction, DNA was denatured and blotted onto nylon membranes and hybridised with the 928 bp probe (5'-flanking region of *Tff2*). *PCR+* and *C+* represent the 928 bp product (positive control), *V* vector (negative control), and *B1/B2* two BAC recombinants containing *Tff* genes.

The insert was also sequenced, confirming the presence of the exon 1 and its start-codon (ATG) as well as the PstI boundary at 5' end (Fig. 21).

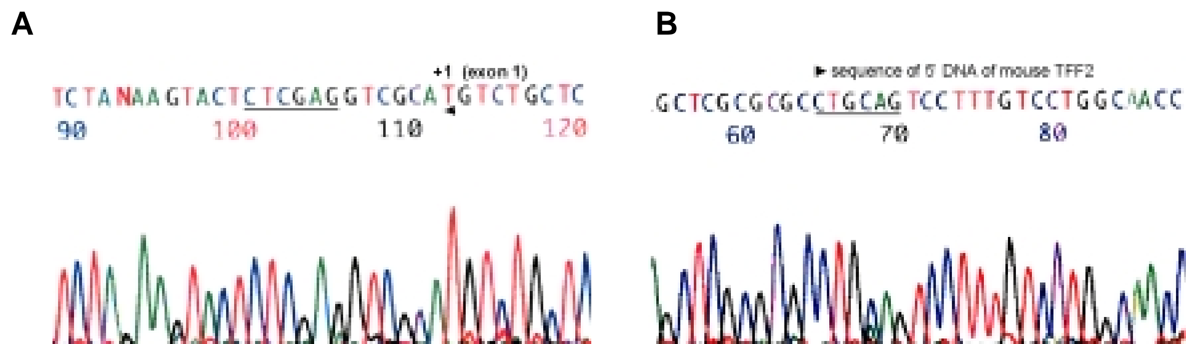


Figure 21: Partial sequence of the insert. (A) XhoI site (underlined); arrowhead indicates the start codon ATG. (B) PstI site (underlined) and unpublished sequences of 5' part of the gene.

As a next step, the PstI-XhoI fragment was excised with *Aat*II and *Xba*I from the pBK-CMV plasmid and then subcloned into the AatII-XbaI (insertion in a directional dependent manner) of pL2-CB, generating the plasmid "pL2-Δm2", also called the targeting vector. This vector was confirmed by sequencing of ligation boundaries (like in Fig. 21; the whole sequence see Fig. 23), PCR, spot hybridisation, and restriction analysis (Fig. 22).

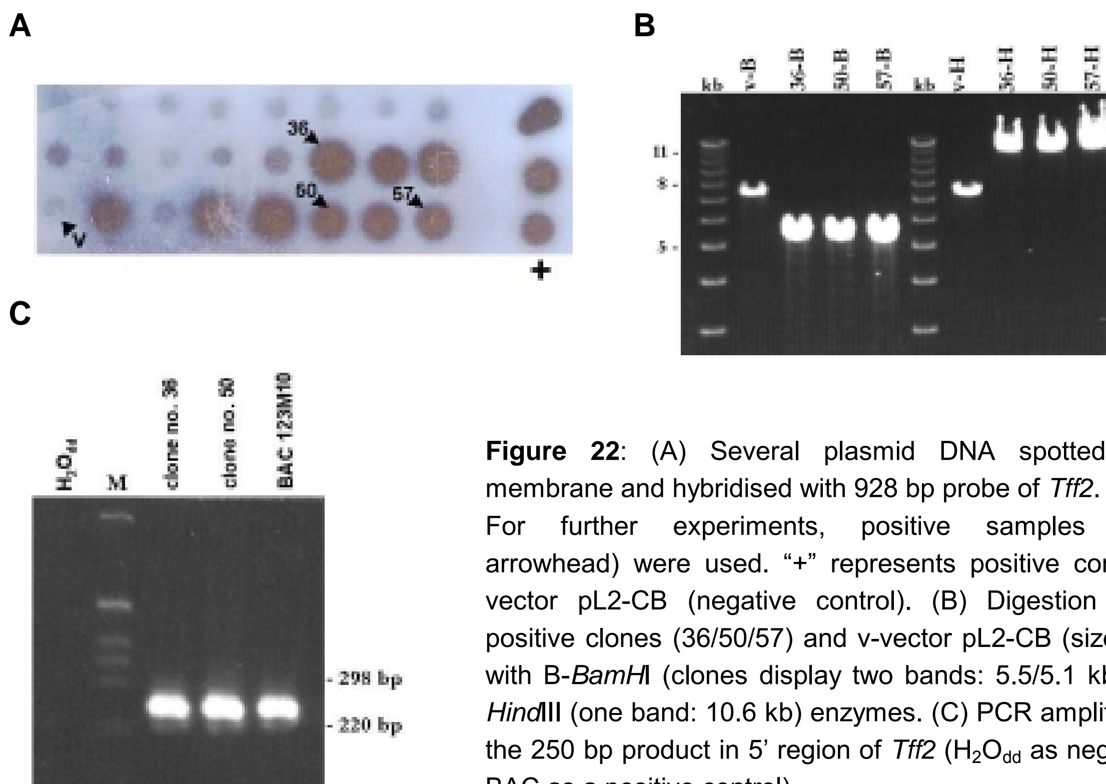


Figure 22: (A) Several plasmid DNA spotted on the membrane and hybridised with 928 bp probe of *Tff2*.
For further experiments, positive samples (marked arrowhead) were used. "+" represents positive controls, "v" vector pL2-CB (negative control). (B) Digestion of three positive clones (36/50/57) and v-vector pL2-CB (size: 7.6 kb) with B-*BamH*I (clones display two bands: 5.5/5.1 kb) and H-*Hind*III (one band: 10.6 kb) enzymes. (C) PCR amplification of the 250 bp product in 5' region of *Tff2* (H_2O_{dd} as negative and BAC as a positive control).

```
5'(PstI)-
ctgcagtcctttgtcctggcaaccatagcatggtaatttggggcagctaatttacttgcaaaatggcttg
attctctcaagacctgctttgaaattcagctagattgggcctggggcacctccactcttggccagttag
cctgttcctgtggtgtgcctctgcgtctgtctgtaacccatccttgggttcctcgtagtctccttcctca
ggcagggtagagtccaccctcttagaaccactacaaaaccgggagtgtccggctgcctcctacctcgggt
tcagagcatgggggagctcggtcctctaagccttatgctgctgcttccacgcatgctcagcccgctgagc
atcccttccttcctcaatatccgatttgtctggccttcaggtcaaggcagtgaaggatttgactaagaac
tcaaggaactcacatttatggtctgcggcttcgttttcttccttctcccacccagctaccctggttctgg
gcctggttcttgtttctctactctgccttctttccctcctctggattcctgtcatttgacagggctttag
agtcccctacaggagagggtggtctggaccgtttttttctgtgtgctccgagaaagactggctagtacct
gaggaattccagtttgcctgaatagataaacaccatctcccccaaggtgtgcagggagactggggtgggg
ctggaggtgggagtggagagaagacatgagaaaaatcacaccacgaagtcacggctggtctgagagcatg
cgcagaagcagcagcgcgaggcttgcacatgcaagccctggaggaggcagcggggtgctcctggtgttgt
agtgggggggtggtgatgtgagccccctggaagtcgggctggaggagaagaggcacagacaaggagagct
gagctgctgtggagtcgagggttaggatttaaggaagcagctgctgtggtagccacagagaaagtgctag
aagcgagagtgtactgaacgggaaagaagcccgtgttttgtccgtggagagttcaggttctgagagagag
agagagagagagagagagagagagagagagagagagagagagagagagagagagaaggggaagttcctagag
tccaggagaagatacagaagatacagagcccggcagccatcagtctgcatgtgacctgcctggggttcag
aagccgtttatgcctgtaccaagatttgggccaatgaccttgggactgagtccttgtcatcagagacggg
tcagtggaccccagtgaggcaggtgcaatagacgtggcacagtgcagagtctgtatccacagcacccata
agttttcctataaagtggaaactgtagctgtggggttctcctgagcctctgtgattggcccacagcctgg
tcccacagatatggctgtctattgttgtgcagtacccttgactttatatggaatggggcgggggtgcagt
gctcataatatggatagcctgttcatcaggacacaggcttgtacaggggggaagggataatgctcaaagga
aagataaatgtctcttcaggggacaggtggctttggcggaacccctgatggggcaggggaacttgggttg
ggatgtgggccagagaggagggaggaatgaggatgttcaagaaaatgggttcttggcaagtcttatcatc
ttgatggctgtcctagttgactgccccttgctatgataaccaaaaggaacctggatagaaaagcaattat
ttatccaatacatccctatcagtcagtcattgagggaagtcagacaggaccaccaagcaggaacctggag
gcaggaaccaggaggcaggaactgaagcagagaccatggagaaatgctgcttactggcttcccccaccc
accccccatggcttactcagtctgctttccatctcttgccagatgaccctggcttgtgtcaagttgacac
caaaacagcacagcaacaacaacaacaaagatacgtaataaacacagagataaagaaccgctgggatgtg
ccacagagaaatagagccctgctgggtttgcaaggatccaggtgaggttttgaggatttaggtctgctct
ttagagcctgggatggggccaaggcttttaaattctgtatccgagtcatggactctgtagcctatctccc
acagctgggcagactccttcagccccaaaggacacacacacacacacacacacagtccctctccttctct
cctctgacatttctgttctgagttgtcctgatcccctgggttctgtcctcggtgtccattttacaaatcc
cctctccggggcatttcatgggtgggcatctcagtcctaggcactgtgtgagggaagtgaagtatcccca
cggggcaaagagtgtccagcttcagggatgaagaaagctgcctgccccatgacggctcccttaggccctg
gcggctggtctctgtgtacaacagtccttggggctggcagggagccaggtgaagtagggtgtgccttgca
tcaggaatgcagactgggtcccctggagatgctcccccctcccctgggaggctcctggatgtaatgacctt
ctctctctgtgctctttgtttctccatagcgtcattcacttcaaatatttaaaaaacaaaacaaaacaaa
caaacaagaaacaaaaaaagaaaacaaacaaaaaacaaaacctaagaagtctctacagtaggcacaaacc
ttacccagagtaacaccacgaggctaatcccttcctgggtgtgaccctgtgggtgccctattccagttgc
atacgtggcccattcccctttcagtaacctggcacgacccagcctgacccagcccatgtccggaggggtt
tttcccacctcctctgggcctccacttggagccatgtcagcttttctgcacaaagaacctaggaggcttg
gggccacgagtaagggacagttatcagggccattgagcaaacaaacaggagccaataaaatgccagctgg
tgggaactgggcattctgtaggctgcccaggtccagtggagcagacATGCGACCTCGAG–3'(XhoI)
```

Figure 23: Sequence of the PstI-XhoI fragment (3069 bp) used for the targeting construct. Underlined sequence data was known at the beginning of work; bold sequence was published in GenBank later by Ribieras *et al.* 2000 (accession No. AJ271003); and important parts (e.g. restriction sites) were confirmed by my data (except for red nucleotide). Blue coloured sequence containing PstI site (bold/underlined) is unpublished. At the end of the sequence, CDS (coding sequence, yellow) with start codon ATG (bold) and XhoI site (bold/underlined).

The targeting vector pL2-Δm2 was linearised with *Aat*II and sent to the colleagues at ICRF (Transgenic Lab, Clare Hall, UK), where the subsequent step of the work, the transfection into ES cells was undertaken.

Figure 24 summarises the revised strategy II to construct a targeting ("knockout") vector by using BAC recombinants as an additional source.

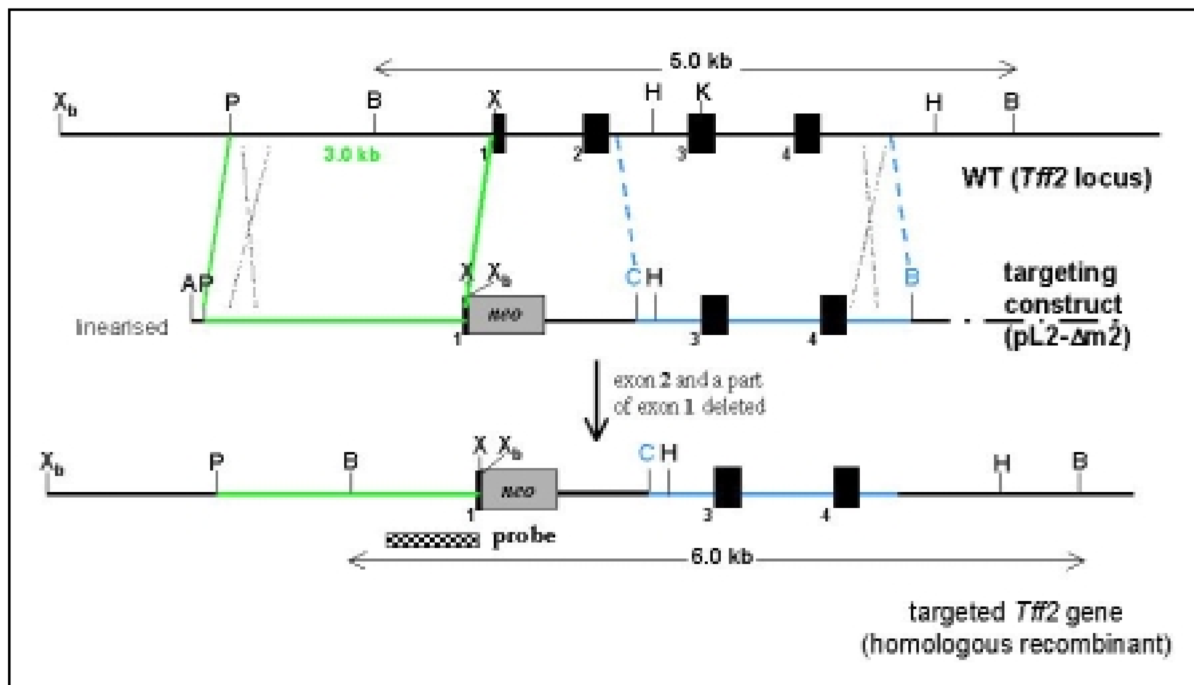


Figure 24: Construction of the targeting vector pL2-Δm2. Mutational inactivation of the *Tff2* gene by homologous recombination. Top: restriction map of *Tff2*; middle: 3 kb PstI-XhoI and 2.4 kb ClaI-BamHI fragments subcloned into pL2-neo vector. Solid black boxes indicate the four exons of *Tff2*. The dashed lines mark the targeted region for homologous recombination. Bottom: restriction map after recombination; a diagnostic BamHI site generates a 6.0 kb fragment using the probe (black&white box) indicated, in contrast to a 5.0 kb fragment in the wild-type (WT) mouse locus. A, AatII; B, BamHI; C, ClaI; E, EcoRI; H, HindIII; K, KpnI; X, XhoI and X_b, XbaI.

In order to further improve the final recombination event by increasing the size of homologous sequences, a second targeting vector, namely pL2-ΔTff2 (Fig. 25), with an enlarged 5'-flanking DNA (5 kb XbaI-XhoI fragment; Fig. 19) was constructed.
Shotgun cloning of XbaI-XhoI BAC fragments led to successful generation of pBK-CMV plasmids containing the desired 5'-flanking DNA fragment of *Tff2*. Several hundred colonies were investigated by colony hybridisation (data not shown). The insert was amplified by PCR using modified primers (additional XbaI site to the XhoI site) to take advantage of the XbaI site of pL2-CB and to accomplish the last cloning step. After the *Xba*I restriction of the amplified product, it was subcloned into the XbaI site of the pL2-CB vector, generating the gene targeting construct pL2-ΔTff2 (Fig. 25).

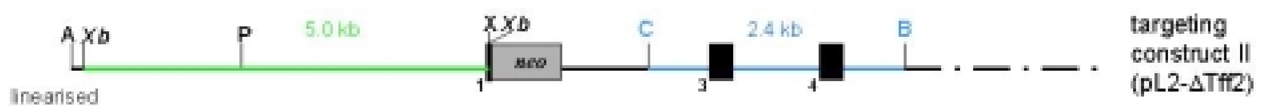


Figure 25: Targeting vector pL2-ΔTff2. For letter description see Fig. 23.

The reactions were confirmed as previously described by spot hybridisation, restriction analysis and PCR (Fig 26). The orientation of the insert was determined by PCR using the primers m2pro and neo-rev (Fig. 26C) and restriction analysis (fingerprinting-data not shown).

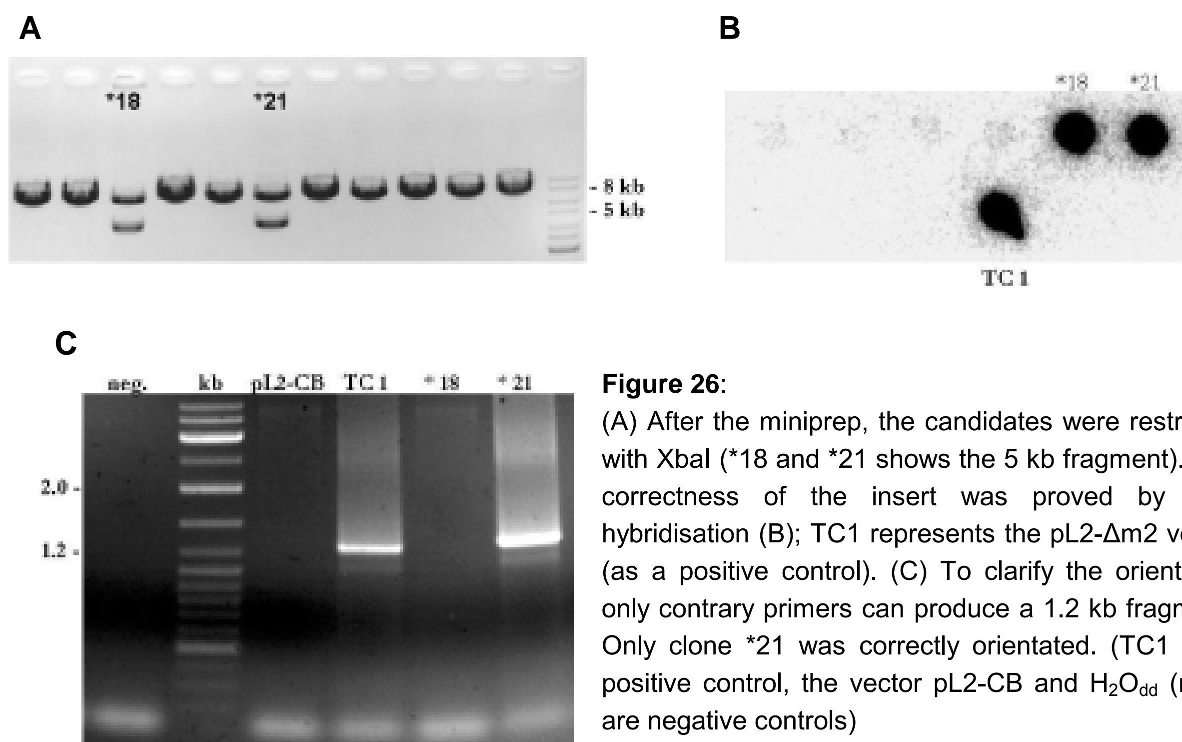


Figure 26:
(A) After the miniprep, the candidates were restricted with XbaI (*18 and *21 shows the 5 kb fragment). The correctness of the insert was proved by spot hybridisation (B); TC1 represents the pL2-Δm2 vector (as a positive control). (C) To clarify the orientation only contrary primers can produce a 1.2 kb fragment. Only clone *21 was correctly orientated. (TC1 as a positive control, the vector pL2-CB and H_2O_{dd} (neg.) are negative controls)

4.2 Genomic structure of the *Tff* gene cluster

4.2.1 Characterisation of the YAC/ BAC contig

By using *Tff2* (for BAC) and *Tff3* (for YAC) specific oligonucleotides for the BAC and YAC murine genomic libraries, the screening yielded one positive YAC and three BAC candidates which appeared useful for more detailed studies. PCR with primers representing all three *Tff* (see Table 4) genes disclosed that the YAC contained *Tff2* and *Tff3* sequences, but no *Tff1*. (Fig. 27).

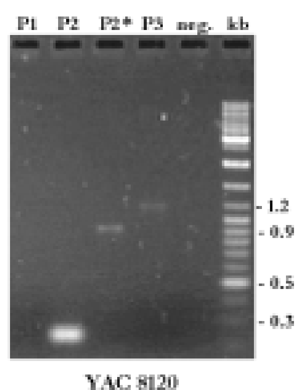


Figure 27: PCR amplification of YAC-DNA using *Tff*-primers. *Tff*3 primer (P3) produces a 1,14 kb fragment (arrowhead); two *Tff*2-primers generate the 250 bp (P2) and 928 bp (P2*) fragment, respectively. *Tff*1 (P1) was negative, expected size ~850 bp (neg., negative control)

All BACs showed products using the 5'-flanking DNA primers representing all three *Tff*. Southern blot analysis displayed, however, that two BACs (517E16 and 566K2) were carrying complete *Tff1* and *Tff2* sequences and only a part of exon 1 of the *Tff3* gene, while BAC 123M10 harboured all three *Tff* genes (Fig. 28 and 29).

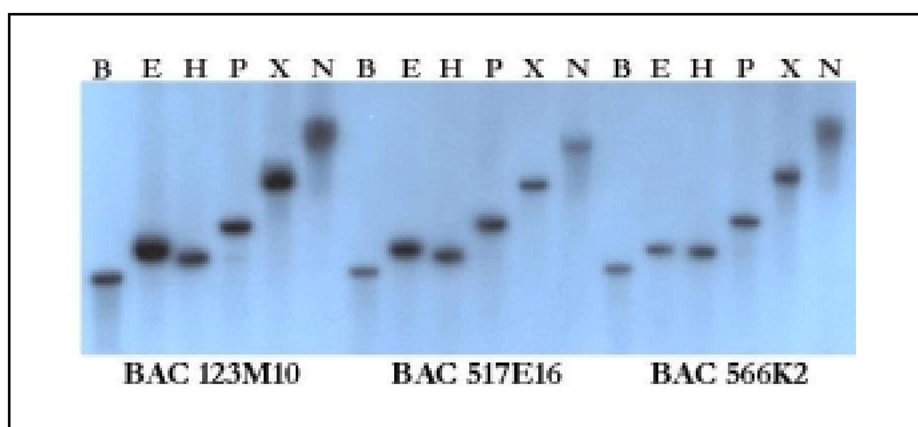


Figure 28: Autoradiograph of the Southern blot of all three BACs showing the hybridisation signals with the 884 bp probe of *Tff1*. All BACs demonstrate the same pattern. (B, BamH**I**; E, EcoR**I**; H, Hind**III**; P, Pst**I**; X, Xba**I** and N, Not**I**). When using the 928 bp probe of *Tff2*, again all BACs revealed the same pattern (data not shown).

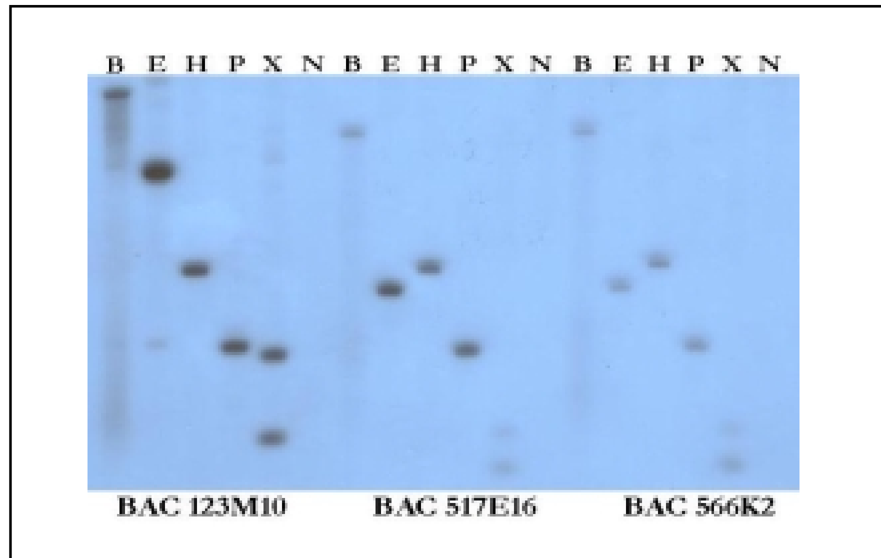


Figure 29: Autoradiograph of the Southern blot of all three BACs showing the hybridisation signals with the 1140 bp probe of *Tff3*. Here two BACs (517E16 and 566K2) present the same pattern. BAC 123M10 in particular displays signals of longer fragments. B, BamHI; E, EcoRI; H, Hind**III**; P, Pst**I**; X, Xba**I** and N, Not**I**.

Fingerprinting and PFGE of the BACs by restriction enzymes in analogy to the study of the human *TFF* contig (Gött *et al.*, 1996) resulted in the construction of the mouse contig (Fig. 30). The presence or absence of particular *Tff* genes within the recombinants allowed us to place the murine *Tff* cluster within a genomic fragment range of approximately 48 kb.

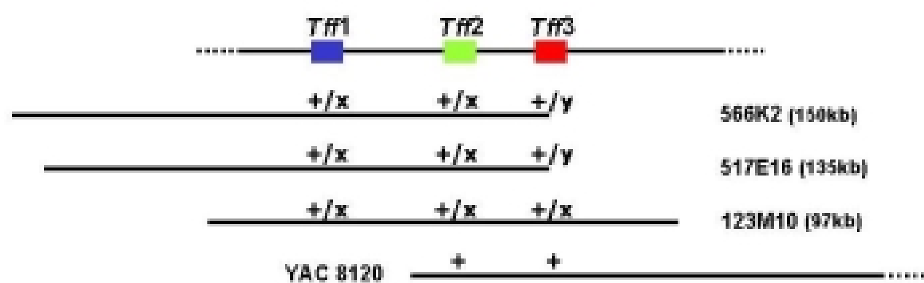


Figure 30: Organisation of the gene cluster of trefoil peptides (*Tff*) within a BAC/YAC contig. Presence of *Tff* genes were analysed by PCR (**+**) and hybridisation (**x**). y indicates hybridisation of smaller restriction fragments, suggesting an insert boundary in BAC 517E16 and BAC 566K2.

4.2.2 Mapping of the *Tff* cluster to mouse chromosome 17

Tff3 was previously mapped to mouse chromosome 17 (Burmeister and Meyer, 1997; Chinery *et al.*, 1996b) around 17.0 cM next to the genetic markers cystathione beta synthase (*Cbs*) and alpha A crystallin (*Crya1*)(Hamvas *et al.*, 1996). Therefore, FISH analysis was performed using BAC 123M10 and signals were obtained at 17q (Fig 31).

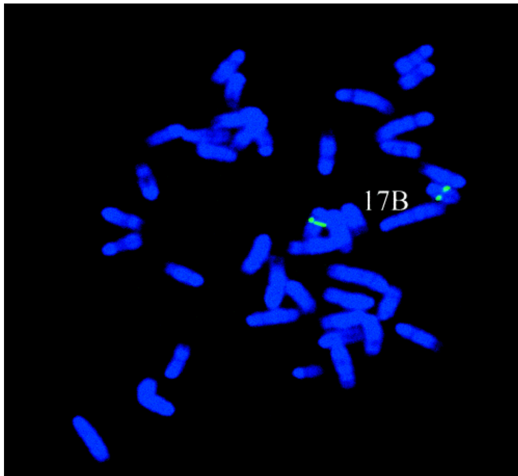


Figure 31: Fluorescence *in situ* hybridisation of the BAC 123M10 probe showing hybridisation to mouse chromosome 17 (17q).

4.2.3 Distances and orientation of *Tff* genes by long distance PCR display the murine *Tff* gene cluster within 40 kb

The three mouse genes of the trefoil factor family were mapped in detail within the contig (see Fig. 30) after employing PCR by using appropriate primers. By applying oligonucleotide primers chosen from the 5'-flanking and 3'-flanking region of each gene (Tff1f 3'-UTR/Tff2r 5'-flanking and Tff2f 3'-UTR/Tff3r 5'-flanking: see Table 4) to long range PCR experiments, the distance between the three *Tff* genes was assessed. Since outward-directed primers from known flanking regions or 3'-UTR were chosen, the data also disclosed the transcriptional orientation of all three genes.

While a PCR band of 17 kb represents a 18 kb distance between third exon of *Tff1* and the first exon of *Tff2*, a PCR product of 9.5 kb indicated a intergenic length of 11.5 kb between *Tff2* and *Tff3* (Fig. 32 and 33).

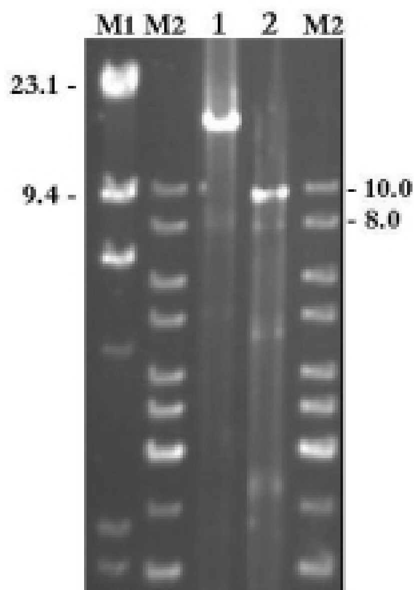


Figure 32: Intergenic distances of the *Tff* genes shown by PCR amplification with appropriate primer pairs.
Lane 1: the distance between *Tff*1 and *Tff*2; lane 2: distance between *Tff2* and *Tff3* (M1, λ-HindII-marker; M2, kb-ladder)(sizes calculated by Bio1D-Software, LTF, D).

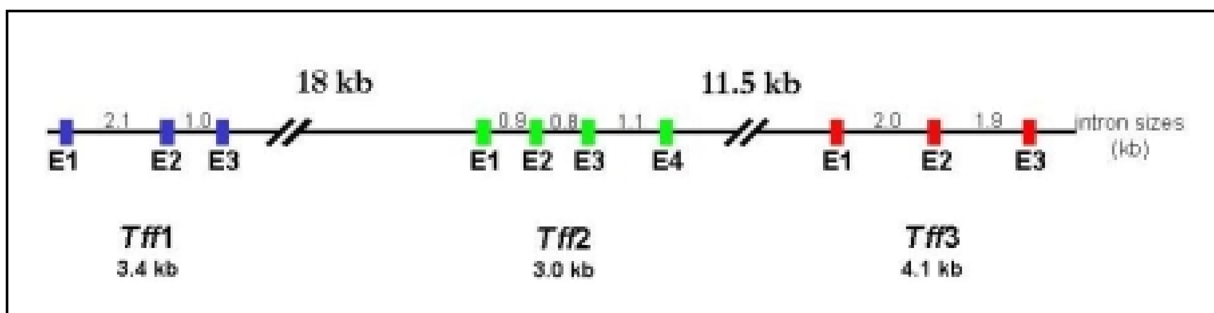


Figure 33: Genomic organisation of the mouse trefoil gene locus harbouring the genes *Tff1, Tff2,* and *Tff3* (E: exon).

The genomic structures exhibited a similar organisation in man and mouse. For example, the gene size of *Tff1* is equivalent to human *TFF1*. The clustered chromosomal location strongly suggests that these genes were created by a process of gene duplication and unequal crossing over. To confirm this hypothesis, I compared the amino acid sequences of the TFF domains, encoded by the central exons that are responsible for the modular character of the corresponding genes (Table 6).

Interestingly, the murine *Tff2* domains I and II (encoded by two separate exons), display more homology to the corresponding human TFF2 domains I and II (red: 83.7 % and 86.1%) than the domains I and II within each species (41.8 % and 44.2 %).

Table 6: Sequence relationships between amino acid sequences of TFF domains from *Homo sapiens* (TFFs) and *Mus musculus* (Tffs). Percent amino acid identities are shown for all pairwise combinations.

	TFF1	TFF2-I	TFF2-II	TFF3	Tff1	Tff2-I	Tff2-II	Tff3
TFF1	100							
TFF2-I	58.1	100						
TFF2-II	44.1	44.2	100					
TFF3	53.5	51.1	48.8	100				
Tff1	72.1	60.5	39.8	46.5	100			
Tff2-I	58.1	83.7	44.1	46.5	60.5	100		
Tff2-II	41.8	41.8	86.1	44.1	41.8	46.5	100	
Tff3	51.1	53.5	46.5	76.7	48.8	53.5	48.8	100

4.3 Screening of ES cells for homologous recombinant clones by PCR and Southern blot analysis

A reliable screening strategy is an essential component of targeting and should be developed during construct design. Consecutive screening was applied: a first screening (which involved PCR) for the rapid detection of candidate targeted colonies (several hundred), and a second one to verify correct targeting by Southern blot analysis. A PCR screening ideally employs a primer (WT-rev) adjacent to the region of homology (outside the shorter arm, namely the 2.4 kb fragment) and a second specific for the novel sequence within the vector sequences (for example the linker of the selectable marker; Reko-for). In this way, a correctly sized PCR product is only generated when the two primer sites are juxtaposed by homologous recombination (Fig. 34).

After electroporation and positive selection (G418, to select for *neo*), ES cell clones need to be screened for incorporation of the targeting construct into the correct genomic locus. The screening strategy was designed in such a way that a band normally seen on a PCR amplification or following Southern blot would be replaced by a band of a predicted size if and only if homologous recombination occurred. Since ES cells are diploid, only one allele is usually altered by the recombination event. If the desired targeting has occurred, this would be detected by two bands representing both, the wild-type and the homologous recombinant alleles.

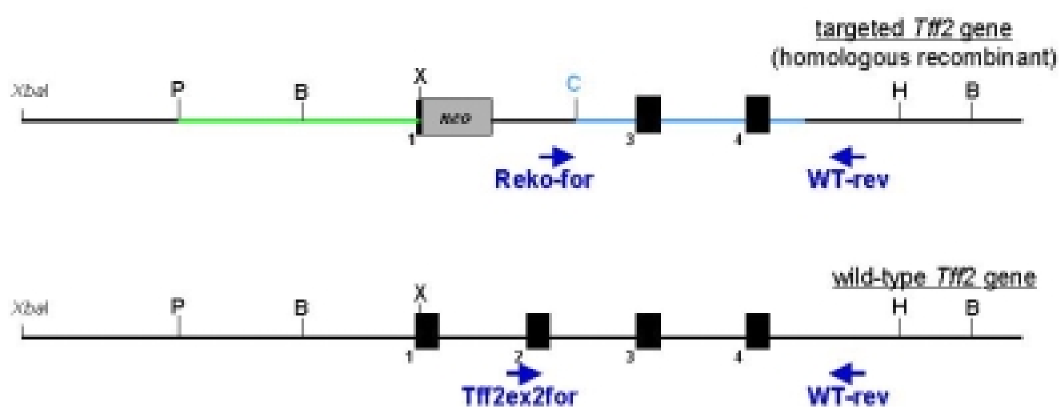


Figure 34: Sites of the primer used for the PCR screening (top) and testing (bottom). Only homologous recombinant clones will generate the genomic DNA fragment (for description see Fig. 24). Arrows indicate the site of primer annealing and solid black boxes the exons.

The ES cell clones were seeded into 96-well plates. One-half of the cells were frozen in the original plate, while the remainder were grown in a replica plate for PCR and Southern blot analysis.

The GK129 embryonic stem (ES) cells initially used derived from 129/Ola mice. In this first assay, DNA of eight 96-well plates were investigated (total: 768). In order to review practicability and DNA quality, the wild-type allele was tested using the primers Tff2ex2for and WT-rev (size: 2.65 kb). This PCR testing worked well, but the screening for homologous recombinants yielded no positive clones in these panels.

Subsequently, we changed our strategy and used AB2.2 ES cells derived from 129/Sv mice, since the cloned PstI-XhoI fragment part derived from a 129/Sv mouse genomic BAC library.

DNA was isolated from 552 ES cell clones and first analysed by PCR. This analysis identified one clone that had a correctly sized PCR product (size: 2.4 kb). The intensity of the fragment, however, in comparison with the corresponding wild-type allele was not equal (Fig. 35, c-B1). It was assumed, that a contamination with another ES cell clone(s) was responsible for this result. Additional subcloning of this clone (c-B1) generated clones #A1, #E1, and #G9 with an equal distribution of bands corresponding to the wild-type and recombinant allele (Fig. 35).

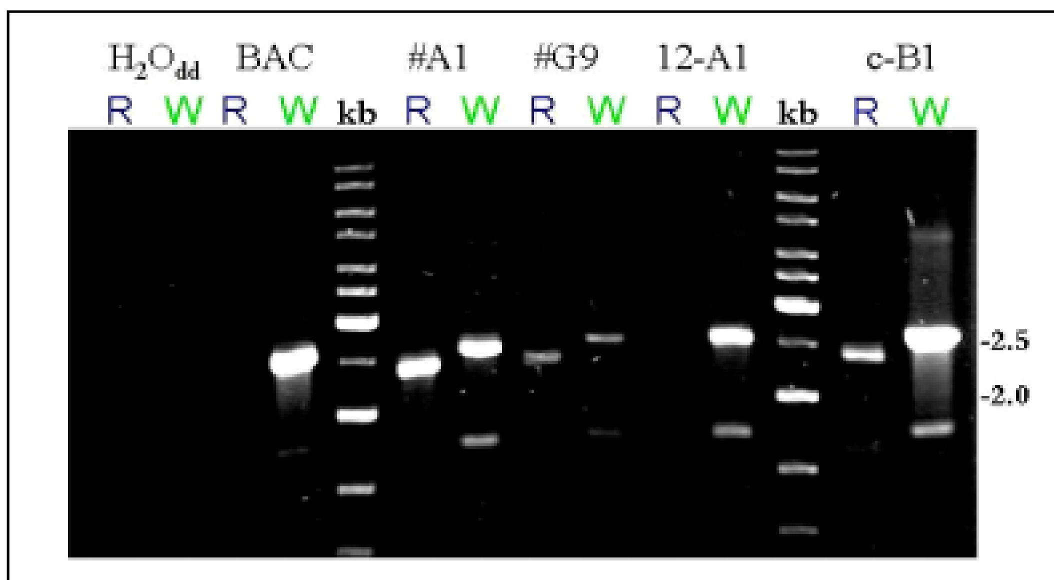


Figure 35: PCR amplification of correctly targeted ES cell clones (#A1, #G9, and c-B1; blue R) in comparison to the wild-type (green W) fragment (as a control). ES cell clone 12-A1 shows no homologous recombination. As expected, BAC contains just the wild-type fragment and water is used as a negative template control. ES cell clone c-B1 reveals an unequal band intensity. R represents recombinant.

Verification through Southern blotting requires that the colonies be expanded *in vitro* to provide a sufficient amount of DNA to carry out the necessary analysis.
This disruption by homologous recombination was confirmed through Southern blot analysis by using the "diagnostic" BamHI site shown in Fig. 24 and an additional site, namely XbaI. Previous studies of the restriction map of *Tff2* displayed a XbaI fragment above 12 kb using the 928 bp *Tff2* probe. Homologous recombinant clones would indicate a "diagnostic" 5 kb hybridisation signal, due to the XbaI site in the polylinker of targeting vector pL2-Δm2. Indeed, this could be demonstrated in subclones #A1, #E1 and #G9 (Fig. 36).

The additional DNA band of approximately 9 kb is unspecific since it also appears in the wild-type mouse.
These results were verified by *BamHI* restriction, resulting in bands corresponding to the expected sizes of wild-type and recombinant alleles.

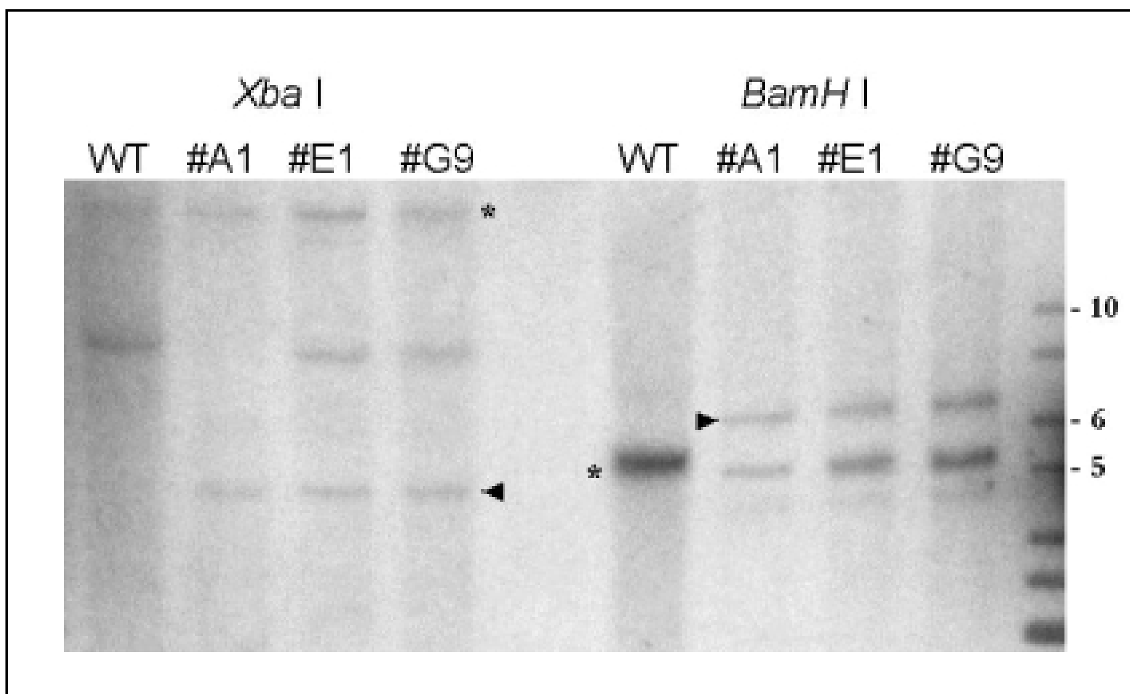


Figure 36: Autoradiograph of the Southern Blot. Genomic DNA of three ES cell clones (#A1-#E1-#G9) and wild-type (WT) were digested with *Xba*I or *BamH*I and hybridised with the probe shown in Fig. 24. The asterisks (*) indicate wild-type 5 kb BamHI or >12 kb XbaI fragment; the arrowheads demonstrate the signals of fragments arose by homologous recombination; *Xba*I generates a ~5 kb diagnostic fragment, BamHI a ~6 kb fragment.

Two ES cells (#A1 and #G9) from these positively identified clones were individually expanded and micro-injected into day 4 blastocysts of C57BL6/J embryos by our colleagues at the Transgenic Lab of ICRF, London. Injecting the ES cells into the blastocyst and transferring this embryo back into the uterus of a foster mother, namely, induce the differentiation of ES cells. 46 injected blastocysts (see Table 7) were transferred into day 3 pseudo-pregnant female mice (CBAxC57BL6/J females were used). Many of the embryos injected this way give rise to chimeric offspring that can be identified by coat colour contribution.

4.4 Breeding of chimeras and heterozygous mice

A chimeric mouse (Fig. 37) is created by mixing cells from a host embryo with ES cells. Coat colour markers are usually used for the detection of chimerism. These animals are mosaic for two cell lineages; the host embryo has a different genotype for coat colour than the ES-cells. The level of chimerism can easily be seen in the pattern formation of the coat. Mice are born naked; the coat becomes visible at 7-10 days of age. If one breeds the chimeric animal, some of the ES cell lineage will form functional gametes and be passed on to the next generation.
Twelve chimeras (nine male, three female) were identified by coat colour contribution. Table 7 shows the injection record of the ES cell clone #G9.

Table 7: Record including data of born and chimeric animals as well as sex and coat colour contribution.

Total number: Injected	**46**	Born / Injected	**65%**	Average % 129 coat colour contribution	**60%**
Total number: Born	**30**	Chimeric / Born	**40%**	Highest % 129 coat colour contribution	**85%**
Number of chimeras	**12**	Male / Female	**9:3**		

Figure 37: An example of a chimera, kindly provided by Ian Rosewell, ICRF.

High contribution male chimeras were backcrossed to C57BL6/J females. Germline transmission was determined by DNA typing of tail DNA.

4.4.1 Germline transmission confirmed by genotyping of F_1 offspring

The first available heterozygous animals were among F_1 mice (129 substrain and C57BL6/J) with brown coat colour. These mice are genetically uniform except for the targeted allele *Tff2.*

Chimeric males were mated to C57BL6/J females to obtain ES-derived offspring that would analysed by a Southern blot of tail DNA to identify the heterozygous (*Tff2*$^{+/-}$) mutants (Fig. 38).

18 F_1 mouse tails were digested to extract DNA for further analysis. For genotyping, Southern blot analysis was performed and at each case a sample of 15 µg DNA was used for digestion with *BamH*I. Germline transmission of the disrupted *Tff2* allele could be confirmed by this analysis of tail DNA from different progeny (Fig. 38).

The wild-type allele gave rise to a 5 kb BamHI fragment (refer to Fig. 36), and an additional 6 kb BamHI fragment indicated a heterozygous (*Tff2*$^{+/-}$) genotype. In this way, germline transmission was clearly determined. Table 8 summarises the hybridisation results of all tails based on sex and allele arrangement.

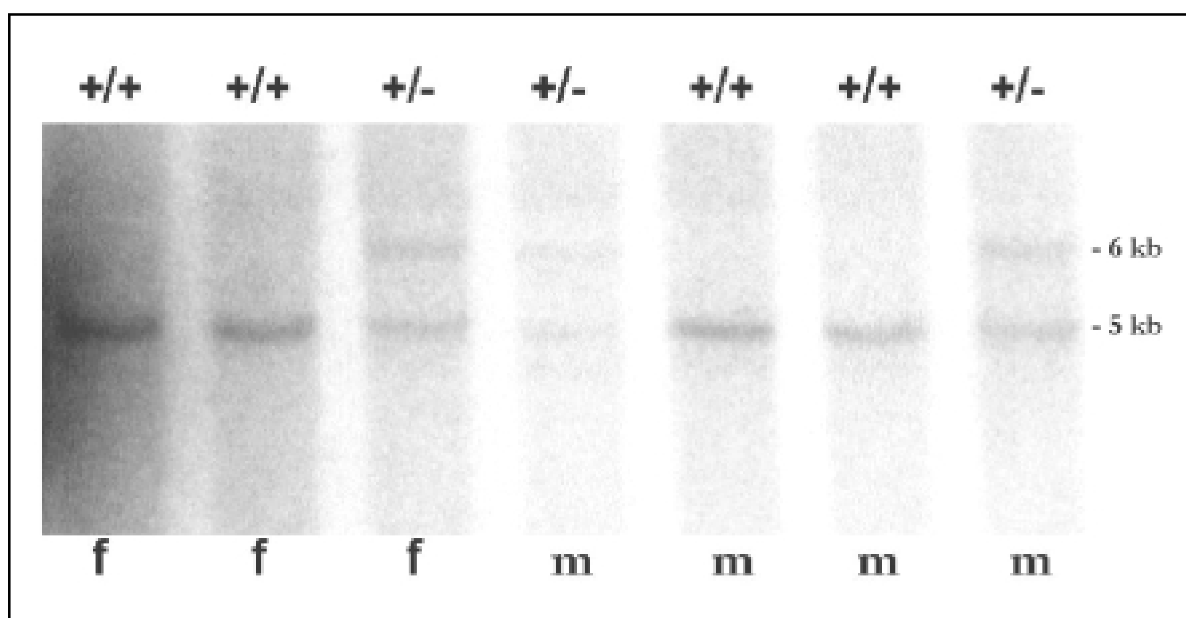


Figure 38: Mouse genotyping with Southern blot analysis. Genomic DNA was extracted from tail fragments of the F_1 litter. After BamHI digestion, DNA was fractionated on 0.7% agarose gel, blotted onto membranes, and hybridised with the probe described in Fig. 24 (f: female, m: male, +/+ wild-type, +/- heterozygous).

Table 8: *Tff*2 genotype (given in absolute numbers) regarding the sex in 18 tail biopsies.

	wild-type (*Tff*2$^{+/+}$)	heterozygous (*Tff*2$^{+/-}$)	Σ
female	6	3	9
male	4	5	9
Σ	10	8	18

Presently, heterozygous mutant mouse will be mated *inter se* to generate progeny with all three genotypes, mutant (-/-), heterozygous mutant (+/-) and wild-type mice (+/+) according to Mendelian rules. The new offspring could then be again analysed by Southern blot of tail DNA to identify the genotype. By this procedure, we will finally identify the correct (-/-) mice to study the affected phenotype.

DISCUSSION

5. Discussion

The trefoil factor family (TFF) is a family of peptides which bear the three-loop trefoil domain. They are a group of secreted proteins expressed by mucous-producing cells throughout the normal gastrointestinal (GI) tract in a site-specific manner.
Over the past ten years, our knowledge has increased concerning the involvement of trefoil peptides in protection or mucosal defence and healing/repair processes in the GI tract. In the case of our study object TFF2, for example, the messenger RNA levels increases within 30 min after mucosal injury was induced by a cryoprobe on the serosal surface of the rat stomach (Alison *et al.*, 1995). Recombinant TFF2 also increases cell migration *in vitro* models of cell wounding (Dignass *et al.*, 1994; Playford *et al.*, 1995) and acts as a cytoprotective agent in rats treated with indomethacin or aspirin (Cook *et al.*, 1998; Playford *et al.*, 1995). Among other things, the addition of trefoil factors to purified mucin preparations leads to a rapid increase in both optical density and viscosity and acts synergistically in cell migration assays (Sands and Podolsky, 1996). Thus, a key role of TFF peptides for maintaining the surface integrity of mucous epithelia is now widely accepted. Nevertheless, the underlying molecular mechanisms and the mode of TFF peptide action is still unknown.

Since the first description of homologous recombination in mouse embryonic stem (ES) cells (Doetschman *et al.*, 1988; Thomas and Capecchi, 1987), gene targeting has become a routine technique in mouse molecular and developmental genetics. Essentially, it allows not only the introduction of inactivating mutations into a gene of interest in ES cells, but also the passage of the targeted allele through the germline, enabling the assessment of gene function from the phenotypic consequences.

By studying these animals, researchers can be able to determine responses to special treatments on a cellular, detailed histopathological level and compare the results with wild-type animals or other available Tff knock out mice (Lefebvre *et al.*, 1996; Mashimo *et al.*, 1996). The example of the Tff1 knockout mouse demonstrates that the results of *in vivo* studies are not always predictable or consistent with the results from *in vitro* analysis.

The aim of the project was the construction of the targeting vector containing the mutated trefoil gene *Tff2* (pL2-Δm2) and the characterisation of the germline transmission of the disrupted *Tff2* allele in heterozygous progeny, since the final goal was to understand the function of the *Tff2* gene.

This construct was transfected first into embryonic stem (ES) cell line derived from 129/Ola, but it did not result in any positive clone. Upon using ES cells derived from 129/Sv, one clone that by homologous recombination contained a disrupted *Tff2* was identified. Subsequently, the successfully targeted ES cell line was injected into blastocysts and transplanted back to the pseudo-pregnant female mice for the breeding program of so-called knock-out mice.

Some general principles concerning the construction of the *Tff*2 gene targeting vector and the underlying strategy must be considered and are discussed in the following segments.

5.1 Replacement constructs substitute the endogenous target gene with the disrupted sequences

When considering the anatomy of targeting constructs, two distinct vector designs can be used: replacement and insertion vectors.
The insertion construct requires only a single crossover (the so-called “O-type” recombination), producing a duplication in the genome of the regions of homology present in the vector. This vector is linearised by cleavage of a unique enzyme within the region of homology. This type of construct is less frequently used for simple knock outs, in part because there is no deletion of any part of the gene; the result is in tandem duplication and disruption of the normal gene structure, not of the exons. In addition, a duplication of exon sequence without a frame-shift mutation does not exclude protein function, and if the insertion is targeted near the 5’ end of the gene, it is possible that it will regenerate a fully or partially functional promoter. To ensure that the exon duplication will create a mutation, it may be necessary to introduce stop codons into a single exon within the targeting vector.

For the purpose of homologous recombination, we decided to use the replacement construct, which leads to a double cross over event with a linearisation site outside of the homologous sequences of the vector. Because no duplication of sequences occurred, we can be sure that *Tff2* was not regenerated, and thus, can serve as an important feature for subsequent operations.
However, an insertion vector can target at a 5-to 20-fold higher frequency than replacement vectors when supplied with the same homologous sequences (Hasty *et al.*, 1991c). Nevertheless, several methods exist that enhance the frequency of homologous recombination or lower the non-targeted events, namely, positive-negative selection (refer to 5.3), promoterless resistance markers (expressed only if integrated properly

downstream of a promoter active in ES cell; Charron *et al.*, 1990), and markers bare of a polyadenylation signal produce stable transcripts only if inserted upstream of a genomic polyadenylation signal (Joyner *et al.*, 1989).

5.2 Use of increased size of homology and isogenic DNA improve the probability of correct targeting

Without a doubt, the size of recombining region and its homology have a profound effect on the frequency of targeting. A total of 5-10 kb is a size proven successful in many targeting experiments. The targeting vector we used for homologous recombination contains homology sequences with a total size of 5.4 kb, which is close to the lower limit.

In addition, the rate of homologous recombination grows with increasing length of the homologous DNA sequence (Deng and Capecchi, 1992). For example, increasing the total homology in a *Hprt* construct from 1.7 to 6.8 kb was found to result in at least a 20-fold increase in targeted events (Hasty *et al.*, 1991c). Although a greater length of homology is always desirable, it may become unwieldy when generating larger constructs through conventional restriction enzyme cloning methods. Another important result is the distribution of the sequences around the selectable marker. Reducing the shorter arm to a level below 1 kb results not only in diminished frequency, but also in reduced fidelity of recombination (Hasty *et al.*, 1991c; Thomas *et al.*, 1992).

Our shorter homology arm has a size of 2.4 kb, which would be sufficient. The longer the stretches of homology, however, the less sensitive a PCR screening procedure can be in identifying recombination events.

Moreover, sequence identity within the homology arms must be maintained. To avoid sequence differences between mouse strains, the DNA from the targeting construct should be isogenic with the target ES cell (this is not absolutely required, but increases the probability of success). It has been shown, that a single DNA mismatch is sufficient to dramatically decrease the rate of correct targeting. For example, homologous recombination at the retinoblastoma susceptibility gene (*Rb*) in embryonic stem cells derived from mouse strain 129 was 20-fold more efficient with a 129-derived targeting construct than with a BALB/c- derived construct (Te Riele *et al.*, 1992). In another study, however, non-isogenic DNA was found to perform equivalently and only minor differences could be demonstrated (Deng and Capecchi, 1992).

The difference between these two studies could be due to the individual genes used for the experiments. In a few extreme cases, the use of isogenic DNA has proven to be crucial (Wurst *et al.*, 1994), but the fact that non-isogenic DNA was used for the generation of knock-out mice demonstrates that this is not an absolute requirement.

In this procedure, the ES cells initially used were derived from 129/Ola mice in order to use isogenic DNA, since the short arm (ClaI-BamHI fragment) of the targeting vector was amplified from this 129/Ola strain. The longer arm (PstI-XhoI fragment), however, was subcloned from the 129/Sv mouse genomic BAC library. This non-isogenie could be the reason for the failure to detect homologous recombinants. In response, we changed to AB2.2 cells derived from 129/Sv mice just as the BAC library was derived from 129/Sv. The screening result revealed that of 552 stably transfected ES cell clones one had undergone homologous recombination.

Ultimately, since the frequency of homologous recombination is exponentially dependent on the length of homology, this factor is certainly of greater importance than the use of isogenic DNA (Deng and Capecchi, 1992; Thomas and Capecchi, 1987). For this reason, the second vector was constructed with increased size of homology (7.4 kb; pL2-ΔTff2, see Fig. 25) to improve the targeting frequency. In the meantime, a successful integration of an ES cell clone with the shorter construct pL2-Δm2 has been established.

5.3 Enrichment for homologous recombinant clones by positive-negative selection

Nearly all targeting constructs used for homologous recombination rely on the positive selection of a drug resistance gene, which is also used to disrupt and mutate the target gene.
Most of the cells surviving positive selection, however, have not undergone homologous recombination, but have only integrated the targeting construct DNA at some random location in the genome. In order to eliminate some of these random integration events, it is advantageous to employ a negative selection method (Mansour *et al.*, 1988). To induce a negative selection, a gene such as the Herpes Simplex Virus-thymidine kinase (HSV-TK) could be added to the end of the targeting construct. If homologous recombination occurs, the TK will get spliced out as it is not homologous to the endogenous gene.

If, however, a random integration event occurs the TK gene will remain. Selecting against the random integration events by adding gancyclovir to the medium will kill any cell that possesses a functional TK gene. Although in reality this positive/negative selection strategy is not foolproof, it allows a relatively efficient selection of rare homologous recombination events.

First, eight 96-well plates (768 colonies) were screened and none of them underwent any homologous recombination. A second screening with 552 ES cell (129/Sv strain cell line) clones yielded one homologous recombinant ES cell clone. Altogether, 1320 clones were investigated.
By using this strategy, at least the number of clones with random integration events that survived selection could be reduced, thus, facilitating the screening for the targeted event.

5.4 *Tff2* gene inactivation is achieved by deleting the signal peptide sequence and one trefoil domain

The generation of a null allele would require the disruption or deletion of critical sequences or functional domain(s). It is important to consider the probability that alleles may be produced that have residual activity or other functions. Targeting events that leave a part of coding sequences upstream or downstream of the insertion intact might have the potential to produce a truncated gene product. In addition, splicing can occur around an exon containing the selection cassette. Generally, null alleles are more likely to occur after deleting or recombining the *neo* gene into one of the 5' (upstream) exons rather than exons that encode C-terminus of the protein, since under these conditions minimal portions of the normal peptide are generated.

In our first strategy, a stop codon in the form of the XbaI site was inserted in exon 2, encoding the first functional trefoil domain, to avert this danger. The palindromic sequence of XbaI (TCTAGA) contains the stop codon TAG and was, therefore, placed in frame.
In the successfully applied targeting vector (pL2-Δm2; Fig. 24) the exon 2, which encodes one of the two functional trefoil domains, and the exon 1 were deleted. The deletion in exon 1 did not include the first 13 nucleotides up to the XhoI site. The following two points were the technical features to disable gene function: (1) to accomplish the gene function, *Tff2* needs the signal peptide sequence to be secreted, but this information was no longer available neither one functional trefoil domain and (2) in order to avoid translation of a truncated protein, a stop codon in the form of the XbaI

site (in frame) was inserted due to the polylinker site of the vector. Any transcription of this targeted gene will result in an oligopeptide (length: 7 amino acids) without function.
It is also possible that the presence of heterologous sequences, i.e. selectable marker, since homologous recombination can cause a number of unanticipated effects. For example, the presence of the *neo* gene, often with its own promoter, can alter the expression of closely neighboured genes (Olson *et al.*, 1996; Pham *et al.*, 1996). This can be a problem in gene clusters where neighboured genes are in the same family, since the genes affected may have similar or identical functions. This could be one possible explanation, why the majority of Tff1 deficient mice had lost Tff2 expression in the stomach (Lefebvre *et al.*, 1996).
To avoid similar effects, the targeting construct pL2-Δm2 already contains *loxP* sites flanking the *neo* gene. In the case of subtle influence of closely neighboured genes or trefoil genes, respectively, then *neo* can be removed after being targeted by transient expression of the Cre recombinase (Sauer, 1993). This will leave one *loxP* site in the genomic DNA. This method leads to a more exact investigation and avoids unanticipated effects.

5.5 A successful strategy: a combination of two screening procedures

An appropriate screening strategy should focus on to discriminating the desired homologous recombination event within the high background of random integration events.

A first screening for the rapid detection of (several hundred) candidate colonies is based on PCR (related to the high number of ES cell clones) which can be designed to detect the juxtaposition of the vector and the target locus (Joyner *et al.*, 1989). It works with small amounts of DNA, is fast and permits the screening of many clones. This is accomplished by using one primer which anneals to the heterologous marker DNA (e.g. *neo* gene) in the vector and a second one priming from the target chromosomal sequences just beyond the homologous sequences used in the construct (Fig. 34). Thus, amplification results in a unique product if homologous recombination occurs (Kim and Smithies, 1988).
The efficiency of such a PCR amplification is inversely related to the distance between the selected primers; for practical reasons, the product should be normally in the 0.5-2.0 kb range. Our PCR product size, used for screening, is 2.4 kb and at the beginning, it was technically difficult to amplify this size. By using an extend PCR program and other modifications, products up to 3 kb were amplified out of minimal amounts of low quality genomic DNA, prepared from less than 10000 ES cells.

In this way, homologous recombination is further characterised for correct integration with Southern blot analysis, the second screening strategy. The screening (of the wild-type and the knock-out or null allele) should produce a clear result, assuming appropriate restriction sites in the targeting construct are available to give a distinct difference in product size. Normally, a screening uses an enzyme that cuts at a defined site both outside the region of homology and within (or on the other side of) the construct, in combination with a probe outside the region of homology (external).
It is also possible to use a probe internal to the construct to detect the diagnostic restriction fragment. A similar-sized restriction fragment, however, may be formed in extremely rare cases by random insertion. The result then remains unclear, if only this screening method is used.
In this case, we used an internal probe, but decided to perform two separate digestions (BamHI and XbaI) to identify the structure of the targeted locus at its 5’ and 3’ ends. Both analyses showed the expected diagnostic restriction fragments (Fig. 36). These data in combination with the unique PCR result (Fig. 35), prove the discriminative power of the internal probe we used.
In addition to a correct band size, it is important to see that the intensity of the endogenous gene product after probe hybridisation is similar to that of the targeted gene, as long as each represents the gene on one chromosome. This is also valid for screening by PCR amplification. The screening of the 129/Sv ES cell clones revealed a positive clone, but it also showed unequal band intensity due to contamination with another ES cell clone (Fig. 35; c-B1). The subcloning of this mixed colony yielded pure clones, such as #A1 and #G9, each displaying the desirable equal intensity of the PCR product (Fig. 35).

After screening, the targeted ES cell clone was used to generate chimeric mice by injection into mouse blastocyst.

The high percentage chimeras were mated to C57BL/6 mice, and germline transmission of the disrupted *Tff2* allele was confirmed by genomic Southern blot analysis of tail DNA from this progeny. The wild-type allele generates a 5 kb BamHI fragment and the heterozygous mutant contains an additional 6 kb BamHI fragment (Fig. 38). The mating between heterozygous male and female will generate a progeny with all three genotypes: homozygous mutant, heterozygous mutant, and wild-type mice.

After genotyping of homozygous mutant *Tff*2 mice (again Southern blot analysis of tail DNA) detailed studies regarding the phenotypic changes is planned.

5.6 Genomic structure of the mouse *Tff* gene cluster

A few years ago two groups (Burmeister and Meyer, 1997; Chinery *et al.*, 1996b) assigned *Tff3* to chromosome 17 of the murine genome, a region homologous with the trefoil gene cluster on human chromosome 21q22.3, by fluorescence *in situ* hybridization (FISH). The other *Tff* genes, however, were not mapped in the mouse.

In order to investigate and detect the mouse *Tff* genes, three bacterial and one yeast artificial chromosome recombinants (BACs and YAC) were isolated and used for detailed characterisation.
Another reason for a detailed investigation involved the study of the mice lacking Tff1 (Lefebvre *et al.*, 1996). One must take into consideration that the interruption of a single *Tff* gene by knock-out experiments influences the regulation of nearby *Tff* genes. Expression of Tff2, which is normally expressed in stomach, duodenum, and pancreas, was not detected in the stomach samples (68%) of *Tff1*$^{-/-}$ mice (Lefebvre *et al.*, 1996; Tomasetto *et al.*, 1990). Lefebvre *et al*. was unable to answer, however, why disruption of *Tff1* specifically affects *Tff2* in the stomach.

Human *TFF* genes are known to be regulated in a coordinated fashion (Gött *et al.*, 1996). Taupin *et al.* have demonstrated, that mice deficient in the intestine-specific peptide intestinal trefoil factor (Tff3), in which colonic restitution is lethally impaired, show reduced expression of the gastric trefoil genes *Tff2* and *Tff1*, suggesting that trefoil peptides may individually regulate transcription of the entire family (Taupin *et al.*, 1999). Indeed, Otto and Patel suggested that there are differences in the coordinated response pathways for Tff peptides that vary with age (Otto and Patel, 1999). In accordance with this data, the 5'-flanking regions of the corresponding human and murine genes show common binding sites for developmental transcription factors such as HNF-3 or GATA (Beck *et al.*, 1999; Al azzeh *et al.*, 2000)

After having established the genomic trefoil gene cluster in the mouse, we now have the genetic material for further experiments to explore the regulation of *Tff* genes, for example, by reporter gene assays, or chromatin immunoprecipitation (ChIP), or Electrophoretic Mobility Shift Assay (EMSA).

5.6.1 Only one BAC recombinant contains all three *Tff* genes

PCR with primers representing all three *Tff* genes disclosed that the YAC contained *Tff2* and *Tff3* sequences, Southern blot analysis revealed that BAC 517E16 and BAC 566K2 were carrying complete *Tff1* and *Tff2* sequences and only a part of exon 1 of the *Tff3* gene, while BAC 123M10 harboured all three *Tff* genes (Fig. 28 and 29). Hybridisation signals of smaller restriction fragments (Fig. 29) and in particular HindIII and PstI digestion patterns, which were equal compared to BAC 123M10, suggest an insert boundary at an identical site within *Tff*3 in BAC 517E16 and BAC 566K2.
Since the HindIII site of pBeloBAC11 (the BAC vector) was used for cloning of those inserts, the HindIII site in exon 1 of *Tff3* must have been used.

Fingerprinting and PFGE of the BACs by restriction enzymes in analogy to the study of the human *TFF* contig (Gött *et al.*, 1996) resulted in the construction of the mouse contig (Fig. 30). Presence or absence of particular *Tff* genes within the recombinants allowed us to place the murine *Tff* cluster within a genomic fragment of approximately 40 kb. In contrast, the three human *TFF* genes are clustered within 50 kb in the chromosomal region 21q22.3, but are assembled in an identical transcriptional orientation (*TFF1-TFF2-TFF3*)(Gött *et al.*, 1996).

5.6.2 Mapping of the *Tff* cluster to mouse chromosome 17 using FISH

Tff3 was previously mapped to mouse chromosome 17 (Burmeister and Meyer, 1997; Chinery *et al.*, 1996b) around 17.0 cM next to the genetic markers cystathione beta synthase (*Cbs*) and alpha A crystallin (*Crya1*)(Hamvas *et al.*, 1998).

Therefore, FISH analysis was performed with BAC 123M10 and signals were obtained at 17q (Fig. 31). While larger parts of murine chromosome 17 are paralogous to human chromosome 6 (ca. 18 cM) and, to lesser extend, to 2p (6 cM), 18p (2 cM), 19p (1.5 cM) and 16p (1 cM), the human region 21q22.3 is paralogous to a small region of mouse chromosome 17 at position 17.2 to 17.4 cM (Hamvas *et al.*, 1996). The FISH data coincide quite well with this sequence arrangement.

5.6.3 Murine gene distances but not the overall *TFF* cluster architecture are different from the human counterpart

In order to explore the vicinity of the trefoil genes, long distance PCR was performed using DNA polymerase with proofreading activity. Since outward-directed primers from known flanking regions or 3'-UTR were chosen, the data also disclosed the transcriptional orientation of all three genes.
While a PCR amplification band of 17 kb represents a 18 kb distance between third exon of *Tff1* and the first exon of *Tff2*, a PCR product of 9.5 kb indicates a intergenic length of 11.5 kb between *Tff2* and *Tff3* (Fig. 32).

These distances are considerably different from the distances of the human *TFF* gene cluster (12 kb between *TFF1* and *TFF2*, 30 kb between *TFF2* and *TFF3*)(Gött *et al.*, 1996; Seib *et al.*, 1997). As reported for the human arrangement, the three murine *Tff* genes are also organised tandemly in a head-to-tail fashion (Fig. 33).

5.6.4 Highest homology levels of amino acid sequences are found between corresponding trefoil domains from mouse and man

The genomic structures exhibit a similar organisation in both man and mouse. For example, the gene size of *Tff1* is equivalent to human *TFF1*. *Tff2* would be also similar to the human gene, except for the larger size of human intron 2, which contains a 25 bp repeat that originated during simian evolution (Kayademir *et al.*, 1998).

The clustered chromosomal location strongly suggests that these genes have been created by a process of gene duplication and unequal crossing over. To support this hypothesis, the amino acid sequences of the TFF domains encoded by the central exons responsible for the modular character of the corresponding genes were analysed (Table 6).

Interestingly, the murine Tff2 domains I and II encoded by two separate exons, are more similar to the corresponding human TFF2 domains I and II (83.7 % and 86.1%) than the domains I and II within each species (41.8 % and 44.2 %). This cross-species homology is also evident for the domains of Tff1 (72.1 %) and Tff3 (76.7%), suggesting that this gene cluster was generated before the rodent-primate divergence about 100 million years ago (Novacek, 1992).

Since the Tff2-I domain exhibits considerable similarity to Tff1 (60.5 %; 58.1 % in the human counterparts) compared to Tff2-II (46.5%; 44.2% in the human counterparts), we propose that the Tff2-I domain evolved by a duplication event and/or unequal crossing over of the Tff1 domain.

A few years ago, a hypothetical evolutionary history of four 11p15.5 human mucin genes was presented by Desseyn *et al.* (Desseyn *et al.*, 1998). Three genes of this cluster (*MUC2*, *MUC5AC*, and *MUC5B*) contain several Cys-subdomains with ten cysteine residues that have a highly conserved position, a feature also characteristic for the TFFs. After comparing of the nucleotide sequences of the Cys-subdomains from the MUC genes, the authors argued that the three genes may have evolved from a common ancestral gene by two successive duplications. In addition, they postulated that *MUC5AC* and *MUC5B* evolved in a concerted manner via the MUC5ACB progenitor, while *MUC2* evolved separately (Desseyn *et al.*, 1998).
Based on the data in Table 6 and the conserved genomic organisation, it is possible that an ancestral *TFF* gene containing only one TFF domain was duplicated to form a TFF1 progenitor and an ancestral TFF2-II/TFF3 progenitor. The recent *TFF2* might have been generated in a later step after acquiring its first domain from TFF1.

Database searches done by our research group (Sommer *et al.*, 1999) revealed that an ancestral TFF-like domain initially arose before amphibian evolution. This is nicely demonstrated by the domains' presence in teleosts and tunicate.

5.7 Outlook

At this stage, once several male and female *Tff*2+/- mice have been successfully identified, a final breeding step is required to obtain homozygous animals exhibiting the phenotypic consequences of the loss of Tff2 in all cells. Thereafter, it may be possible to understand the functions of Tff2 by assessing the pathophysiological impact of the absence of this trefoil peptide in the mouse. Special attention must be given to a possible role in protection, healing, and tumorigenesis of the digestive tract.

Since the Tff1 and Tff3 knock-out mouse models manifested either reduced or no expression of the other remaining trefoil genes, Tff2 knock-out mice may also reveal aberrant expression of the other trefoil genes, which should help to elucidate their coordinated regulation.

For a more detailed analysis of mice lacking Tff2, gene inactivation may be controlled by using the Cre/*loxP* recombinase system. Therefore, an accurately designed, new targeting construct for conditional gene inactivation must be made, by means of a spatially (as in cell type- or tissue-specific) knockout or a temporally (through control of the activity or expression of the recombinase Cre) controlled knockout.

SUMMARY

6. Summary

Trefoil peptides are a group of small secreted proteins, which play an important role in the protecting and healing processes of the stomach and intestine.

The aim of these experiments was to create a basic foundation for understanding the exact function and/or the mode of action of the *Tff2* gene. Since in the case of human trefoil peptides, no naturally occurring chance mutations leading to constitutional diseases had been discovered, a straight forward experimental approach for new insight into this trefoil peptide function is to engineer a strain of so-called knock-out mice, in which Tff2 is missing. For Tff1 and Tff3, respectively, knock-outs were established already. The information gained from these experiments included: in case of *Tff1*$^{-/-}$ the development of gastric adenocarcinoma and mucosa abnormalities and for *Tff3*$^{-/-}$ decreased healing of the intestinal mucosa.

By employing a replacement strategy, one allele of *Tff2* was substituted in ES cells by homologous recombination in order to create the Tff2 knock-out mouse.
The primary aim of this project was to construct the targeting vector (pL2-Δm2; Fig. 24), whereby the mouse trefoil gene *Tff2* was disrupted, and to characterise the germline transmission of the mutant allele in heterozygous (*Tff2*$^{+/-}$) animals. Indeed, this goal was realised within the scope of this work. In the near future, the mice will be mated *inter se* to generate the mutant *Tff2* (-/-) knockout mice, completing the family of already existing trefoil knockout mice.

To obtain fundamental information about the mouse trefoil gene cluster, the genomic structure was determined. Three bacterial and one yeast artificial chromosome recombinants (BACs and YAC) were identified and used for detailed characterisation by PCR, restriction mapping, hybridisation, and fluorescence *in situ* hybridisation (FISH). In a similar fashion to the *TFF* gene cluster in humans, the mouse *Tff* genes cover a region of approximately 40 kb in the transcriptional order *Tff1-Tff2-Tff3* and are localised on chromosome 17. Based on this conserved genomic structure, a model of how mammalian TFF genes may have evolved by exon duplication and/or an unequal crossing over was proposed.

Zusammenfassung

Kleeblattpeptide sind kleine sezernierte Proteine, die eine wichtige Rolle bei Wundheilungsprozessen und dem Schutz des Verdauungstraktes spielen.

Das Ziel dieser Arbeit war es, Grundlagen für das Verständnis der Funktion des *Tff2* Gens zu schaffen. Für die humanen Kleeblattproteine sind keine natürlich vorkommenden Mutationen, welche zu konstitutionellen Krankheiten führen könnten, bekannt. Daher sind die sogenannten „Knockout-Mäuse“ für die Grundlagenforschung von großer Bedeutung, um neue Einblicke in die Wirkungsweise der Kleeblattproteine zu erlangen. In dieser Arbeit sollte das natürlich vorhandene *Tff2* Gen ausgeschaltet werden, um pathologische Situationen zu simulieren und dadurch die physiologische Funktion des Genproduktes aufzuklären. Für Tff1 bzw. Tff3 sind Knockout-Mäuse bereits etabliert. Bei der *Tff1*$^{-/-}$-Maus konnte eine vermehrte Bildung von Magenspezifischem Adenokarzinom beobachtet werden, wogegen die *Tff3*$^{-/-}$-Maus eine verminderte Wundheilung der Darmschleimhaut zeigte.

Durch homologe Rekombination wurde ein Allel des *Tff2* Gens in den embryonalen Stammzellen der Maus ersetzt. Dazu wurde ein Zielvektor (pL2-Δm2; Fig. 24) konstruiert, in welchem das murine Kleeblattgen *Tff2* deaktiviert vorliegt. Darüberhinaus wurde das mutierte Allel in die Keimbahn der heterozygoten (*Tff2*$^{+/-}$) Tiere erfolgreich integriert. Basierend auf diesen Charakterisierungen werden nun die heterozygoten (*Tff2*$^{+/-}$) Tiere untereinander gekreuzt, um letztendlich die *Tff2* (-/-) Knockout-Maus zu züchten. Die Erzeugung dieser Maus wird die Familie der bereits bestehenden Kleeblatt-Knockout-Mäuse komplettieren.

Als Grundlage für diese Arbeit wurde die genomische Struktur der murinen Kleeblattfamilie bestimmt. Drei bakterielle und eine von der Hefe abstammende Rekombinante (BACs und YAC) wurden identifiziert und für die detaillierte Charakterisierung durch PCR, Restriktionskartierung, Hybridisierung, und Fluoreszenz *in situ* Hybridisierung (FISH) verwendet. Daraus ergab sich, dass in der Maus die drei *Tff* Gene ähnlich wie im menschlichen *TFF* Gen-Cluster angeordnet sind. Sie erstrecken sich über eine Länge von ca. 40 kb in der transkriptionellen Anordnung *Tff1-Tff2-Tff3* auf Chromosom 17. Aufgrund dieser konservierten Genomstruktur wurde ein Evolutionsmodell, basierend auf Exon-Duplikation und ungleichem Crossing-over, für die *TFF* Gene bei Säugetieren entwickelt.

7. Bibliography

Ahnen, D. A., Gullick, W., and Wright, N. A. Multiple growth factor production by the cell lineage induced by mucosal ulceration in Crohn's disease. Gastroenterology, A512 (abstract). 1991.

Al-azzeh, E. D., Fegert, P., Blin, N., and Gött, P. (2000). Transcription factor GATA-6 activates expression of gastroprotective trefoil genes TFF1 and TFF2. *Biochim Biophys Acta* **1490** :324-332.

Alison, M. R., Chinery, R., Poulsom, R., Ashwood, P., Longcroft, J. M., and Wright, N. A. (1995). Experimental ulceration leads to sequential expression of spasmolytic polypeptide, intestinal trefoil factor, epidermal growth factor and transforming growth factor alpha mRNAs in rat stomach. *J Pathol* **175**:405-414.

Amemiya, Y. and Miyahara, J. (1988). Imaging plate illuminates many fields. *Nature* **336**:89-90.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidemann, J. G., Smith, J. A., and Struhl, K. (1995). Current Protocols in Molecular Biology. *Current Protocols in Molecular Biology*, John Wiley & Sons Inc., New York, NY, USA

Babyatsky, M. W., deBeaumont, M., Thim, L., and Podolsky, D. K. (1996). Oral trefoil peptides protect against ethanol- and indomethacin-induced gastric injury in rats. *Gastroenterology* **110**:489-497.

Baker, M. E. (1988). Oestrogen-induced pS2 protein is similar to pancreatic spasmolytic polypeptide and the kringle domain. *Biochem J* **253**:307-309.

Beck, S., Fegert, P., and Gött, P. (1997). Factors regulating pS2-reporter gene expression in MCF-7 breast cancer cell line. *Int J Oncol* **10**:1051-1055.

Beck, S., Sommer, P., dos Santos, S. E., Blin, N., and Gött, P. (1999). Hepatocyte nuclear factor 3 (winged helix domain) activates trefoil factor gene TFF1 through a binding motif adjacent to the TATAA box. *DNA Cell Biol* **18**:157-164.

Berry, M., Nunez, A. M., and Chambon, P. (1989). Estrogen-responsive element of the human pS2 gene is an imperfectly palindromic sequence. *Proc Natl Acad Sci USA* **86**:1218-1222.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl Acids Res* **7***:* 1513-1523.

Bonkhoff, H., Stein, U., Welter, C., and Remberger, K. (1995). Differential expression of the pS2 protein in the human prostate and prostate cancer: association with premalignant changes and neuroendocrine differentiation. *Hum Pathol* **26**:824-828.

Botzler, C., Oertel, M., Hinz, M., and Hoffmann, W. (1999). Structure of the Xenopus laevis TFF-gene xP4.1, differentially expressed to its duplicated homolog xP4.2. *Biochim Biophys Acta* **1489**:345-353.

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

Burmeister, M. and Meyer, G. E. (1997). The trefoil gene maps to mouse chromosome 17. *Mamm Genome* **8**:223-224.

Cappelletti, V., Coradini, D., Scanziani, E., Benini, E., Silvestrini, R., and Di Fronzo, G. (1992). Prognostic relevance of pS2 status in association with steroid receptor status and proliferative activity in node-negative breast cancer. *Eur J Cancer* **28A**:1315-1318.

Carr, M. D. (1992). 1H NMR-based determination of the secondary structure of porcine pancreatic spasmolytic polypeptide: one of a new family of "trefoil" motif containing cell growth factors. *Biochemistry* **31**:1998-2004.

Carr, M. D., Bauer, C. J., Gradwell, M. J., and Feeney, J. (1994). Solution structure of a trefoil-motif-containing cell growth factor, porcine spasmolytic protein. *Proc Natl Acad Sci USA* **91**:2206-2210.

Cavailles, V., Garcia, M., and Rochefort, H. (1989). Regulation of cathepsin-D and pS2 gene expression by growth factors in MCF7 human breast cancer cells. *Mol Endocrinol* **3**:552-558.

Chadwick, M. P., Westley, B. R., and May, F. E. (1997). Homodimerization and hetero-oligomerization of the single-domain trefoil protein pNR-2/pS2 through cysteine 58. *Biochem J* **327 (Pt 1)**:117-123.

Charron, J., Malynn, B. A., Robertson, E. J., Goff, S. P., and Alt, F. W. (1990). High-frequency disruption of the N-myc gene in embryonic stem and pre-B cell lines by homologous recombination. *Mol Cell Biol* **10**:1799-1804.

Chen, Y. H., Lu, Y., De Plaen, I. G., Wang, L. Y., and Tan, X. D. (2000). Transcription factor NF-kappaB signals antianoikic function of trefoil factor 3 on intestinal epithelial cells. *Biochem Biophys Res Commun* **274**:576-582.

Chinery, R., Bates, P. A., De, A., and Freemont, P. S. (1995). Characterisation of the single copy trefoil peptides intestinal trefoil factor and pS2 and their ability to form covalent dimers. *FEBS Lett* **357**:50-54.

Chinery, R. and Cox, H. M. (1995). Immunoprecipitation and characterization of a binding protein specific for the peptide, intestinal trefoil factor. *Peptides* **16**:749-755.

Chinery, R. and Playford, R. J. (1995). Combined intestinal trefoil factor and epidermal growth factor is prophylactic against indomethacin-induced gastric damage in the rat. *Clin Sci (Colch)* **88**:401-403.

Chinery, R., Poulsom, R., and Cox, H. M. (1996b). The gene encoding mouse intestinal trefoil factor: structural organization, partial sequence analysis and mapping to murine chromosome 17q. *Gene* **171**:249-253.

Chinery, R., Poulsom, R., Elia, G., Hanby, A. M., and Wright, N. A. (1993). Expression and purification of a trefoil peptide motif in a beta- galactosidase fusion protein and its use to search for trefoil- binding sites. *Eur J Biochem* **212**:557-563.

Chinery, R., Poulsom, R., Rogers, L. A., Jeffery, R. E., Longcroft, J. M., Hanby, A. M., and Wright, N. A. (1992). Localization of intestinal trefoil-factor mRNA in rat stomach and intestine by hybridization *in situ*. *Biochem J* **285 (Pt 1)**:5-8.

Chinery, R., Williamson, J., and Poulsom, R. (1996a). The gene encoding human intestinal trefoil factor (TFF3) is located on chromosome 21q22.3 clustered with other members of the trefoil peptide family. *Genomics* **32**:281-284.

Chu, G., Vollrath, D., and Davis, R. W. (1986). Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* **234**:1582-1585.

Church, G. M. and Gilbert, W. (1984). Genomic sequencing. *Proc Natl Acad Sci USA* **81**:1991-1995.

Clark, J. M. (1988). Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res. **16**, 9677-9686.

Colbere-Garapin, F., Horodniceanu, F., Kourilsky, P., and Garapin, A. C. (1981). Construction of a dominant selective marker useful for gene transfer studies in animal cells. *Dev Biol Stand* **50**:323-326.

Collier, J. D., Bennett, M. K., Bassendine, M. F., and Lendrum, R. (1995). Immunolocalization of pS2, a putative growth factor, in pancreatic carcinoma. *J Gastroenterol Hepatol* **10**:396-400.

Cook, G. A., Familari, M., Thim, L., and Giraud, A. S. (1999). The trefoil peptides TFF2 and TFF3 are expressed in rat lymphoid tissues and participate in the immune response. *FEBS Lett* **456**:155-159.

Cook, G. A., Thim, L., Yeomans, N. D., and Giraud, A. S. (1998). Oral human spasmolytic polypeptide protects against aspirin-induced gastric injury in rats. *J Gastroenterol Hepatol* **13**:363-370.

Cook, G. A., Yeomans, N. D., and Giraud, A. S. (1997). Temporal expression of trefoil peptides in the TGF-alpha knockout mouse after gastric ulceration. *Am J Physiol* **272**:G1540-G1549.

Dante, R., Ribieras, S., Baldassini, S., Martin, V., Benzerara, O., Bouteille, C., Bremond, A., Frappart, L., Rio, M. C., and Lasne, Y. (1994). Expression of an estrogen-induced breast cancer-associated protein (pS2) in benign and malignant human ovarian cysts. *Lab Invest* **71**:188-192.

De, A., Brown, D. G., Gorman, M. A., Carr, M., Sanderson, M. R., and Freemont, P. S. (1994). Crystal structure of a disulfide-linked "trefoil" motif found in a large family of putative growth factors. *Proc Natl Acad Sci USA* **91**:1084-1088.

DeChiara, T. M., Efstratiadis, A., and Robertson, E. J. (1990). A growth-deficiency phenotype in heterozygous mice carrying an insulin- like growth factor II gene disrupted by targeting. *Nature* **345**:78-80.

Deng, C. and Capecchi, M. R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol Cell Biol* **12**:3365-3371.

Desseyn, J. L., Buisine, M. P., Porchet, N., Aubert, J. P., Degand, P., and Laine, A. (1998). Evolutionary history of the 11p15 human mucin gene family. *J Mol Evol* **46**:102-106.

Devine, D. A., High, A. S., Owen, P. J., Poulsom, R., and Bonass, W. A. (2000). Trefoil factor expression in normal and diseased human salivary glands. *Hum Pathol* **31**:509-515.

Dignass, A., Lynch-Devaney, K., Kindon, H., Thim, L., and Podolsky, D. K. (1994). Trefoil peptides promote epithelial migration through a transforming growth factor beta-independent pathway. *J Clin Invest* **94**:376-383.

Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S., and Smithies, O. (1987). Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* **330**:576-578.

Doetschman, T., Maeda, N., and Smithies, O. (1988). Targeted mutation of the Hprt gene in mouse embryonic stem cells. *Proc Natl Acad Sci USA* **85**:8583-8587.

dos Santos, S. E., Ulrich, M., Doring, G., Botzenhart, K., and Gött, P. (2000). Trefoil factor family domain peptides in the human respiratory tract. *J Pathol* **190**:133-142.

Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**:6127-6145.

Efstathiou, J. A., Noda, M., Rowan, A., Dixon, C., Chinery, R., Jawhari, A., Hattori, T., Wright, N. A., Bodmer, W. F., and Pignatelli, M. (1998). Intestinal trefoil factor controls the expression of the adenomatous polyposis coli-catenin and the E-cadherin-catenin complexes in human colon carcinoma cells. *Proc Natl Acad Sci USA* **95**:3122-3127.

Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**:154-156.

Familari, M., Cook, G. A., Taupin, D. R., Marryatt, G., Yeomans, N. D., and Giraud, A. S. (1998). Trefoil peptides are early markers of gastrointestinal maturation in the rat. *Int J Dev Biol* **42**:783-789.

Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**:6-13.

Foekens, J. A., Rio, M. C., Seguin, P., van Putten, W. L., Fauque, J., Nap, M., Klijn, J. G., and Chambon, P. (1990). Prediction of relapse and survival in breast cancer patients by pS2 protein status. *Cancer Res* **50**:3832-3837.

Frandsen, E. K., Jørgensen , K. H., and Thim, L. (1986). Receptor binding of pancreatic spasmolytic polypeptide (PSP) in rat intestinal mucosal cell membranes inhibits the adenylate cyclase activity. *Regul Pept* **16**:291-297.

Gajhede, M., Petersen, T. N., Henriksen, A., Petersen, J. F., Dauter, Z., Wilson, K. S., and Thim, L. (1993). Pancreatic spasmolytic polypeptide: first three-dimensional structure of a member of the mammalian trefoil family of peptides. *Structure* **1**:253-262.

Galli-Taliadoros, L. A., Sedgwick, J. D., Wood, S. A., and Korner, H. (1995). Gene knock-out technology: a methodological overview for the interested novice. *J Immunol Methods* **181**:1-15.

Gmach, M., Berger, H., Thalhammer, J., and Kreil, G. (1990). Dermal glands of Xenopus laevis contain a polypeptide with a highly repetitive amino acid sequence. *FEBS Lett* **260**:145-148.

Goldenring, J. R., Poulsom, R., Ray, G. S., Wright, N., Meise, K. S., and Coffey, R. J., Jr. (1996). Expression of trefoil peptides in the gastric mucosa of transgenic mice overexpressing transforming growth factor-alpha. *Growth Factors* **13**:111-119.

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., and Ruddle, F. H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* **77**:7380-7384.

Gött, P., Beck, S., Machado, J. C., Carneiro, F., Schmitt, H., and Blin, N. (1996). Human trefoil peptides: genomic structure in 21q22.3 and coordinated expression. *Eur J Hum Genet* **4**:308-315.

Griffiths, A. J. F., Gelbart, W. M., Miller, J. H, and Lewontin, R. C (1999). "Modern Genetic Analysis," W.H. Freeman&Company, New York, NY, USA.

Grunstein, M. and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc Natl Acad Sci USA* **72**:3961-3965.

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

Gu, H., Zou, Y. R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73**:1155-1164.

Hamvas, R., Trachtulec, Z., Forejt, J., Williams, R. W., Arzt, K., Fischer-Lindahl, K., and Silver, L. M. (1996). Mouse Chromosome 17. *Mamm Genome* **6**:281-299.

Hamvas, R. M., Trachtulec, Z., Vernet, C., and Forejt, J. (1998). Encyclopedia of the mouse genome VII. Mouse chromosome 17. *Mamm Genome* **8** Spec No:S320-S342.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**:557-580.

Hanby, A. M., Chinery, R., Poulsom, R., Playford, R. J., and Pignatelli, M. (1996). Downregulation of E-cadherin in the reparative epithelium of the human gastrointestinal tract. *Am J Pathol* **148**:723-729.

Hanby, A. M., Jankowski, J. A., Elia, G., Poulsom, R., and Wright, N. A. (1994). Expression of the trefoil peptides pS2 and human spasmolytic polypeptide (hSP) in Barrett's metaplasia and the native oesophageal epithelium: delineation of epithelial phenotype. *J Pathol* **173**:213-219.

Hanby, A. M., McKee, P., Jeffery, M., Grayson, W., Dublin, E., Poulsom, R., and Maguire, B. (1998). Primary mucinous carcinomas of the skin express TFF1, TFF3, estrogen receptor, and progesterone receptors. *Am J Surg Pathol* **22**:1125-1131.

Hanby, A. M., Poulsom, R., Elia, G., Singh, S., Longcroft, J. M., and Wright, N. A. (1993a). The expression of the trefoil peptides pS2 and human spasmolytic polypeptide (hSP) in 'gastric metaplasia' of the proximal duodenum: implications for the nature of 'gastric metaplasia'. *J Pathol* **169**:355-360.

Hanby, A. M., Poulsom, R., Singh, S., Elia, G., Jeffery, R. E., and Wright, N. A. (1993b). Spasmolytic polypeptide is a major antral peptide: distribution of the trefoil peptides human spasmolytic polypeptide and pS2 in the stomach. *Gastroenterology* **105**:1110-1116.

Hanby, A. M., Poulsom, R., Singh, S., Jankowski, J., Hopwood, D., Elia, G., Rogers, L., Patel, K., and Wright, N. A. (1993c). Hyperplastic polyps: a cell lineage which both synthesizes and secretes trefoil-peptides and has phenotypic similarity with the ulcer-associated cell lineage. *Am J Pathol* **142**:663-668.

Hasty, P., Rivera-Perez, J., and Bradley, A. (1991a). The length of homology required for gene targeting in embryonic stem cells. *Mol Cell Biol* **11**:5586-5591.

Hasty, P., Ramirez-Solis, R., Krumlauf, R., and Bradley, A. (1991b). Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. *Nature* **350**:243-246.

Hasty, P., Rivera-Perez, J., Chang, C., and Bradley, A. (1991c). Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. *Mol Cell Biol* **11**:4509-4517.

Hattori, M., Fujiyama, A., Taylor, T. D., Watanabe, H., Yada, T., Park, H. S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D. K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., and Schudy, A. (2000). The DNA sequence of human chromosome 21. The chromosome 21 mapping and sequencing consortium. *Nature* **405**:311-319.

Hauser, F., Gertzen, E. M., and Hoffmann, W. (1990). Expression of spasmolysin (FIM-A.1): an integumentary mucin from Xenopus laevis. *Exp Cell Res* **189**:157-162.

Hauser, F. and Hoffmann, W. (1991). xP1 and xP4. P-domain peptides expressed in Xenopus laevis stomach mucosa. *J Biol Chem* **266**:21306-21309.

Hauser, F. and Hoffmann, W. (1992). P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic polymorphism. *J Biol Chem* **267**:24620-24624.

Hauser, F., Poulsom, R., Chinery, R., Rogers, L. A., Hanby, A. M., Wright, N. A., and Hoffmann, W. (1993). hP1.B, a human P-domain peptide homologous with rat intestinal trefoil factor, is expressed also in the ulcer-associated cell lineage and the uterus. *Proc Natl Acad Sci USA* **90**:6961-6965.

Hauser, F., Roeben, C., and Hoffmann, W. (1992). xP2, a new member of the P-domain peptide family of potential growth factors, is synthesized in Xenopus laevis skin. *J Biol Chem* **267**:14451-14455.

Henry, J. A., Bennett, M. K., Piggott, N. H., Levett, D. L., May, F. E., and Westley, B. R. (1991). Expression of the pNR-2/pS2 protein in diverse human epithelial tumours. *Br J Cancer* **64**:677-682.

Henry, J. A., Nicholson, S., Hennessy, C., Lennard, T. W., May, F. E., and Westley, B. R. (1990). Expression of the oestrogen regulated pNR-2 mRNA in human breast cancer: relation to oestrogen receptor mRNA levels and response to tamoxifen therapy. *Br J Cancer* **61**:32-38.

Higashiyama, M., Doi, O., Kodama, K., Yokouchi, H., Inaji, H., Nakamori, S., and Tateishi, R. (1994). Prognostic significance of pS2 protein expression in pulmonary adenocarcinoma. *Eur J Cancer* **30A**:792-797.

Hoffmann, W. (1988). A new repetitive protein from Xenopus laevis skin highly homologous to pancreatic spasmolytic polypeptide. *J Biol Chem* **263**:7686-7690.

Hoffmann, W. and Hauser, F. (1993). The P-domain or trefoil motif: a role in renewal and pathology of mucous epithelia? *Trends Biochem Sci* **18**:239-243.

Iannaccone, P. M., Evans, E. P., and Burtenshaw, M. D. (1985). Chromosomal sex and distribution of functional germ cells in a series of chimeric mice. *Exp Cell Res* **156**:471-477.

Itoh, H., Tomita, M., Uchino, H., Kobayashi, T., Kataoka, H., Sekiya, R., and Nawa, Y. (1996). cDNA cloning of rat pS2 peptide and expression of trefoil peptides in acetic acid-induced colitis. *Biochem J* **318 (Pt 3)**:939-944.

Jakowlew, S. B., Breathnach, R., Jeltsch, J. M., Masiakowski, P., and Chambon, P. (1984). Sequence of the pS2 mRNA induced by estrogen in the human breast cancer cell line MCF-7. *Nucleic Acids Res* **12**:2861-2878.

Jeffrey, G. P., Oates, P. S., Wang, T. C., Babyatsky, M. W., and Brand, S. J. (1994). Spasmolytic polypeptide: a trefoil peptide secreted by rat gastric mucous cells. *Gastroenterology* **106**:336-345.

Jørgensen , K. D., Diamant, B., Jørgensen , K. H., and Thim, L. (1982b). Pancreatic spasmolytic polypeptide (PSP): III. Pharmacology of a new porcine pancreatic polypeptide with spasmolytic and gastric acid secretion inhibitory effects. *Regul Pept* **3**:231-243.

Jørgensen , K. H., Thim, L., and Jacobsen, H. E. (1982a). Pancreatic spasmolytic polypeptide (PSP): I. preparation and initial chemical characterization of a new polypeptide from porcine pancreas. *Reg Peptides* **3**:207-219.

Joyner, A. L. (2000). "Gene Targeting", Oxford University Press Inc., New York, NY, USA.

Joyner, A. L., Skarnes, W. C., and Rossant, J. (1989). Production of a mutation in mouse En-2 gene by homologous recombination in embryonic stem cells. *Nature* **338**:153-156.

Kanai, M., Mullen, C., and Podolsky, D. K. (1998). Intestinal trefoil factor induces inactivation of extracellular signal- regulated protein kinase in intestinal epithelial cells. *Proc Natl Acad Sci USA* **95**:178-182.

Kayademir, T., Rosewell, I., Blin, N., and Gött, P. (2000). Genomic structure of the mouse trefoil factor (Tff) gene cluster in 17q. *GeneScreen* **1**:15-19.

Kayademir, T., Silva, E. d., Pusch, C., Beck, S., Machado, J. C., and Gött, P. (1998). A novel 25 bp tandem repeat within the human trefoil peptide gene TFF2 in 21q22.3: polymorphism and mammalian evolution. *Eur J Hum Genet* **6**:121-128.

Khulusi, S., Hanby, A. M., Marrero, J. M., Patel, P., Mendall, M. A., Badve, S., Poulsom, R., Elia, G., Wright, N. A., and Northfield, T. C. (1995). Expression of trefoil peptides pS2 and human spasmolytic polypeptide in gastric metaplasia at the margin of duodenal ulcers. *Gut* **37**:205-209.

Kim, H. S. and Smithies, O. (1988). Recombinant fragment assay for gene targetting based on the polymerase chain reaction. *Nucleic Acids Res* **16**:8887-8903.

Kindon, H., Pothoulakis, C., Thim, L., Lynch-Devaney, K., and Podolsky, D. K. (1995). Trefoil peptide protection of intestinal epithelial barrier function: cooperative interaction with mucin glycoprotein. *Gastroenterology* **109**:516-523.

Kinoshita, K., Taupin, D. R., Itoh, H., and Podolsky, D. K. (2000). Distinct pathways of cell migration and antiapoptotic response to epithelial injury: structure-function analysis of human intestinal trefoil factor. *Mol Cell Biol* **20**:4680-4690.

Krishnan, B. R., Blakesley, R. W., and Berg, D. E. (1991). Linear amplification DNA sequencing directly from single phage plaques and bacterial colonies. *Nucleic Acids Res* **19**:1153.

Labouvie, C., Machado, J.-C., Carneiro, F., Seitz, G., and Blin, N. (1997). Expression pattern of gastrointestinal markers in native colorectal epithelium, lesions and carcinoma. *Oncology Reports* **4**:1367-1371.

Labouvie, C., Machado, J. C., Carneiro, F., Sarbia, M., Vieth, M., Porschen, R., Seitz, G., and Blin, N. (1999). Differential expression of mucins and trefoil peptides in native epithelium, Barrett's metaplasia and squamous cell carcinoma of the oesophagus. *J Cancer Res Clin Oncol* **125**:71-76.

Lalani, E. N., Williams, R., Jayaram, Y., Gilbert, C., Chaudhary, K. S., Siu, L. S., Koumarianou, A., Playford, R., and Stamp, G. W. (1999). Trefoil factor-2, human spasmolytic polypeptide, promotes branching morphogenesis in MCF-7 cells. *Lab Invest* **79**:537-546.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.*, (2001). Initial sequencing and analysis of the human genome. International Human Genome Sequencing Consortium. *Nature* **409**:860-921.

Langer, G., Jagla, W., Behrens-Baumann, W., Walter, S., and Hoffmann, W. (1999). Secretory peptides TFF1 and TFF3 synthesized in human conjunctival goblet cells. *Invest Ophthalmol Vis Sci* **40**:2220-2224.

Lefebvre, O., Chenard, M. P., Masson, R., Linares, J., Dierich, A., LeMeur, M., Wendling, C., Tomasetto, C., Chambon, P., and Rio, M. C. (1996). Gastric mucosa abnormalities and tumorigenesis in mice lacking the pS2 trefoil protein. *Science* **274**:259-262.

Lefebvre, O., Wolf, C., Kedinger, M., Chenard, M. P., Tomasetto, C., Chambon, P., and Rio, M. C. (1993). The mouse one P-domain (pS2) and two P-domain (mSP) genes exhibit distinct patterns of expression. *J Cell Biol* **122**:191-198.

Lin, F. L., Sperle, K., and Sternberg, N. (1985). Recombination in mouse L cells between DNA introduced into cells and homologous chromosomal sequences. *Proc Natl Acad Sci USA* **82**:1391-1395.

Liu, D., el Hariry, I., Karayiannakis, A. J., Wilding, J., Chinery, R., Kmiot, W., McCrea, P. D., Gullick, W. J., and Pignatelli, M. (1997). Phosphorylation of beta-catenin and epidermal growth factor receptor by intestinal trefoil factor. *Lab Invest* **77**:557-563.

Longman, R. J., Douthwaite, J., Sylvester, P. A., Poulsom, R., Corfield, A. P., Thomas, M. G., and Wright, N. A. (2000). Coordinated localisation of mucins and trefoil peptides in the ulcer associated cell lineage and the gastrointestinal mucosa. *Gut* **47**:792-800.

Ludeking, A., Fegert, P., Blin, N., and Gött, P. (1998). Osmotic changes and ethanol modify TFF gene expression in gastrointestinal cell lines. *FEBS Lett* **439**:180-184.

Luqmani, Y., Bennett, C., Paterson, I., Corbishley, C. M., Rio, M. C., Chambon, P., and Ryall, G. (1989). Expression of the pS2 gene in normal, benign and neoplastic human stomach. *Int J Cancer* **44**:806-812.

Machado, J. C., Carneiro, F., Blin, N., and Sobrinho-Simoes, M. (1996a). Pattern of pS2 protein expression in premalignant and malignant lesions of gastric mucosa. *Eur J Cancer Prev* **5** :169-179.

Machado, J. C., Carneiro, F., Ribeiro, P., Blin, N., and Sobrinho-Simoes, M. (1996b). pS2 protein expression in gastric carcinoma. An immunohistochemical and immunoradiometric study. *Eur J Cancer* **32A**:1585-1590.

Machado, J. C., Nogueira, A. M., Carneiro, F., Reis, C. A., and Sobrinho-Simoes, M. (2000). Gastric carcinoma exhibits distinct types of cell differentiation: an immunohistochemical study of trefoil peptides (TFF1 and TFF2) and mucins (MUC1, MUC2, MUC5AC, and MUC6). *J Pathol* **190**:437-443.

Mansour, S. L. (1990). Gene targeting in murine embryonic stem cells: introduction of specific alterations into the mammalian genome. *Genet Anal Tech Appl* **7**:219-227.

Mansour, S. L., Thomas, K. R., and Capecchi, M. R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**:348-352.

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**:7634-7638.

Mashimo, H., Podolsky, D. K., and Fishman, M. C. (1995). Structure and expression of murine intestinal trefoil factor: high evolutionary conservation and postnatal expression. *Biochem Biophys Res Commun* **210**:31-37.

Mashimo, H., Wu, D. C., Podolsky, D. K., and Fishman, M. C. (1996). Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor. *Science* **274**:262-265.

Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A., and Chambon, P. (1982). Cloning of cDNA sequences of hormone-regulated genes from the MCF-7 human breast cancer cell line. *Nucleic Acids Res* **10**:7895-7903.

Maxam, A. M. and Gilbert, W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci USA* **74**:560-564.

May, F. E., Semple, J. I., Newton, J. L., and Westley, B. R. (2000). The human two domain trefoil protein, TFF2, is glycosylated in vivo in the stomach. *Gut* **46**:454-459.

May, F. E. and Westley, B. R. (1988). Identification and characterization of estrogen-regulated RNAs in human breast cancer cells. *J Biol Chem* **263**:12901-12908.

May, F. E. and Westley, B. R. (1997). Expression of human intestinal trefoil factor in malignant cells and its regulation by oestrogen in breast cancer cells. *J Pathol* **182**:404-413.

McKenzie, C., Marchbank, T., Playford, R. J., Otto, W., Thim, L., and Parsons, M. E. (1997). Pancreatic spasmolytic polypeptide protects the gastric mucosa but does not inhibit acid secretion or motility. *Am J Physiol* **273**:G112-G117.

Mead, D. A., Pey, N. K., Herrnstadt, C., Marcil, R. A., and Smith, L. M. (1991). A universal method for the direct cloning of PCR amplified nucleic acid. *Biotechnology (NY)* **9**:657-663.

Mori, K., Fujii, R., Kida, N., Takahashi, H., Ohkubo, S., Fujino, M., Ohta, M., and Hayashi, K. (1990). Complete primary structure of the human estrogen-responsive gene (pS2) product. *J Biochem (Tokyo)* **107**:73-76.

Mortensen, R (2000). Manipulating the Mouse Genome. *In* "Current Protocols in Molecular Biology" (Ausubel F.M., Ed.), pp. 23.5.1-23.5.11, John Wiley&Sons Inc., New York, NY, USA.

Muller, W. and Borchard, F. (1993). pS2 protein in gastric carcinoma and normal gastric mucosa: association with clincopathological parameters and patient survival. *J Pathol* **171**:263-269.

Newton, J. L., Allen, A., Westley, B. R., and May, F. E. (2000). The human trefoil peptide, TFF1, is present in different molecular forms that are intimately associated with mucus in normal stomach. *Gut* **46**:312-320.

Novacek, M. J. (1992). Mammalian phylogeny: shaking the tree. *Nature* **356**:121-125.

Nunez, A. M., Berry, M., Imler, J. L., and Chambon, P. (1989). The 5' flanking region of the pS2 gene contains a complex enhancer region responsive to oestrogens, epidermal growth factor, a tumour promoter (TPA), the c-Ha-ras oncoprotein and the c-jun protein. *EMBO J* **8**:823-829.

Olson, E. N., Arnold, H. H., Rigby, P. W., and Wold, B. J. (1996). Know your neighbors: three phenotypes in null mutants of the myogenic bHLH gene MRF4. *Cell* **85**:1-4.

Otto, W. R. and Patel, K. (1999). Trefoil factor family (TFF)-domain peptides in the mouse: embryonic gastrointestinal expression and wounding response. *Anat Embryol (Berl)* **199**:499-508.

Otto, W. R., Rao, J., Cox, H. M., Kotzian, E., Lee, C. Y., Goodlad, R. A., Lane, A., Gorman, M., Freemont, P. A., Hansen, H. F., Pappin, D., and Wright, N. A. (1996). Effects of pancreatic spasmolytic Polypeptide (PSP) on epithelial cell function. *Eur J Biochem* **235**:64-72.

Pallud, C., Le Doussal, V., Pichon, M. F., Prud'homme, J. F., Hacene, K., and Milgrom, E. (1993). Immunohistochemistry of pS2 in normal human breast and in various histological forms of breast tumours. *Histopathology* **23**:249-256.

Pham, C. T., MacIvor, D. M., Hug, B. A., Heusel, J. W., and Ley, T. J. (1996). Long-range disruption of gene expression by a selectable marker cassette. *Proc Natl Acad Sci USA* **93**:13090-13095.

Piggott, N. H., Henry, J. A., May, F. E., and Westley, B. R. (1991). Antipeptide antibodies against the pNR-2 oestrogen-regulated protein of human breast cancer cells and detection of pNR-2 expression in normal tissues by immunohistochemistry. *J Pathol* **163**:95-104.

Plaut, A. G. (1997). Trefoil peptides in the defense of the gastrointestinal tract. *N Engl J Med* **336**:506-507.

Playford, R. J., Marchbank, T., Chinery, R., Evison, R., Pignatelli, M., Boulton, R. A., Thim, L., and Hanby, A. M. (1995). Human spasmolytic polypeptide is a cytoprotective agent that stimulates cell migration. *Gastroenterology* **108**:108-116.

Playford, R. J., Marchbank, T., Goodlad, R. A., Chinery, R. A., Poulsom, R., and Hanby, A. M. (1996). Transgenic mice that overexpress the human trefoil peptide pS2 have an increased resistance to intestinal damage. *Proc Natl Acad Sci USA* **93**:2137-2142.

Podolsky, D. K. (1997). Healing the epithelium: solving the problem from two sides. *J Gastroenterol* **32**:122-126.

Podolsky, D. K., Lynch-Devaney, K., Stow, J. L., Oates, P., Murgue, B., deBeaumont, M., Sands, B. E., and Mahida, Y. R. (1993). Identification of human intestinal trefoil factor. Goblet cell-specific expression of a peptide targeted for apical secretion. *J Biol Chem* **268**:6694-6702.

Podolsky, D. K., Pleskow, D. K., and Jafari, H. (1988). Latent transformed growth-inhibiting factor in human malignant effusions. *Cancer Res* **48**:418-424.

Polshakov, V. I., Williams, M. A., Gargaro, A. R., Frenkiel, T. A., Westley, B. R., Chadwick, M. P., May, F. E., and Feeney, J. (1997). High-resolution solution structure of human pNR-2/pS2: a single trefoil motif protein. *J Mol Biol* **267**:418-432.

Poulsom, R. (1996). Trefoil peptides. *Baillieres Clin Gastroenterol* **10**:113-134.

Poulsom, R., Hanby, A. M., Lalani, E. N., Hauser, F., Hoffmann, W., and Stamp, G. W. (1997). Intestinal trefoil factor (TFF3) and pS2 (TFF1), but not spasmolytic polypeptide (TFF2) mRNAs are co-expressed in normal, hyperplastic, and neoplastic human breast epithelium. *J Pathol* **183**:30-38.

Probst, J. C., Skutella, T., Müller-Schmid, A., Jirikowski, G. F., and Hoffmann, W. (1995). Molecular and cellular analysis of rP1.B in the rat hypothalamus: *In situ* hybridization and immunohistochemistry of a new P-domain neuropeptide. *Mol Brain Res* **33**:269-276.

Probst, J. C., Zetzsche, T., Weber, M., Theilemann, P., Skutella, T., Landgraf, R., and Jirikowski, G. F. (1996). Human intestinal trefoil factor is expressed in human hypothalamus and pituitary: evidence for a novel neuropeptide. *FASEB J* **10**:1518-1523.

Prud'homme, J. F., Fridlansky, F., Le Cunff, M., Atger, M., Mercier-Bodart, C., Pichon, M. F., and Milgrom, E. (1985). Cloning of a gene expressed in human breast cancer and regulated by estrogen in MCF-7 cells. *DNA* **4**:11-21.

Ramirez-Solis, R., Davis, A. C., and Bradley, A. (1993). Gene targeting in embryonic stem cells. *Methods Enzymol* **225**:855-878.

Ramirez-Solis, R., Rivera-Perez, J., Wallace, J. D., Wims, M., Zheng, H., and Bradley, A. (1992). Genomic DNA microextraction: a method to screen numerous samples. *Anal Biochem* **201**:331-335.

Rasmussen, T. N., Raaberg, L., Poulsen, S. S., Thim, L., and Holst, J. J. (1992). Immunohistochemical localization of pancreatic spasmolytic polypeptide (PSP) in the pig. *Histochemistry* **98**:113-119.

Ribieras, S., Tomasetto, C., and Rio, M. C. (1998). The pS2/TFF1 trefoil factor, from basic research to clinical applications. *Biochim Biophys Acta* **1378**:F61-F77.

Rio, M.C., Lepage, P., Diemunsch, P, Roitsch, C., and Chambon, P. (1988). Structure primaire de la proteine humaine pS2. *Comptes Rendu des Academies de la Science III* **307**:825-831.

Rio, M. C., Bellocq, J. P., Daniel, J. Y., Tomasetto, C., Lathe, R., Chenard, M. P., Batzenschlager, A., and Chambon, P. (1988). Breast cancer-associated pS2 protein: synthesis and secretion by normal stomach mucosa. *Science* **241**:705-708.

Rio, M. C., Bellocq, J. P., Gairard, B., Rasmussen, U. B., Krust, A., Koehl, C., Calderoli, H., Schiff, V., Renaud, R., and Chambon, P. (1987). Specific expression of the pS2 gene in subclasses of breast cancers in comparison with expression of the estrogen and progesterone receptors and the oncogene ERBB2. *Proc Natl Acad Sci USA* **84**:9243-9247.

Rio, M. C. and Chambon, P. (1990). The pS2 gene, mRNA, and protein: a potential marker for human breast cancer. *Cancer Cells* **2**:269-274.

Rio, M. C., Chenard, M. P., Wolf, C., Marcellin, L., Tomasetto, C., Lathe, R., Bellocq, J. P., and Chambon, P. (1991). Induction of pS2 and hSP genes as markers of mucosal ulceration of the digestive tract. *Gastroenterology* **100**:375-379.

Robertson, E., Bradley, A., Kuehn, M., and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* **323**:445-448.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487-491.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**:1350-1354.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

Sands, B. E. and Podolsky, D. K. (1996). The trefoil peptide family. *Annu Rev Physiol* **58**:253-273.

Sanger, F., Nicklen, S., and Coulson, A. R. (1974). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **12**:5463-5467.

Sauer, B. (1993). Manipulation of transgenes by site-specific recombination: use of Cre recombinase. *Methods Enzymol* **225**:890-900.

Schmitt, H., Wundrack, I., Beck, S., Gött, P., Welter, C., Shizuya, H., Simon, M. I., and Blin, N. (1996). A third P-domain peptide gene (TFF3), human intestinal trefoil factor, maps to 21q22.3. *Cytogenet Cell Genet* **72**:299-302.

Schwartz, D. C. and Cantor, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**:67-75.

Schwartzberg, P. L., Goff, S. P., and Robertson, E. J. (1989). Germ-line transmission of a c-abl mutation produced by targeted gene disruption in ES cells. *Science* **246**:799-803.

Schwarzberg, H., Kalbacher, H., and Hoffmann, W. (1999). Differential behavioral effects of TFF peptides: Injections of synthetic TFF3 into the rat amygdala. *Pharmacology Biochemistry and Behavior* **62**:173-178.

Seib, T., Blin, N., Hilgert, K., Seifert, M., Theisinger, B., Engel, M., Dooley, S., Zang, K. D., and Welter, C. (1997). The three human trefoil genes TFF1, TFF2, and TFF3 are located within a region of 55 kb on chromosome 21q22.3. *Genomics* **40**:200-202.

Seib, T., Dooley, S., and Welter, C. (1995). Characterization of the genomic structure and the promoter region of the human intestinal trefoil factor. *Biochem Biophys Res Commun* **214**:195-199.

Seitz, G., Theisinger, B., Tomasetto, G., Rio, M. C., Chambon, P., Blin, N., and Welter, G. (1991). Breast cancer-associated protein pS2 expression in tumors of the biliary tract. *Am J Gastroenterol* **86**:1491-1494.

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc Natl Acad Sci USA* **89**:8794-8797.

Simpson, E. M., Linder, C. C., Sargent, E. E., Davisson, M. T., Mobraaten, L. E., and Sharp, J. J. (1997). Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nat Genet* **16**:19-27.

Singh, S., Poulsom, R., Hanby, A. M., Rogers, L. A., Wright, N. A., Sheppard, M. C., and Langman, M. J. (1998). Expression of oestrogen receptor and oestrogen-inducible genes pS2 and ERD5 in large bowel mucosa and cancer. *J Pathol* **184**:153-160.

Skilton, R. A., Luqmani, Y. A., McClelland, R. A., and Coombes, R. C. (1989). Characterisation of a messenger RNA selectively expressed in human breast cancer. *Br J Cancer* **60**:168-175.

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., and Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**:230-234.

Sommer, P., Blin, N., and Gött, P. (1999). Tracing the evolutionary origin of the TFF-domain, an ancient motif at mucous surfaces. *Gene* **236**:133-136.

Sonoda, M., Takano, M., Miyahara, J., and Kato, H. (1983). Computed radiography utilizing scanning laser stimulated luminescence. *Radiology* **148**:833-838.

Southern, E. M. (1975). Detection of specific sequences separated by gel electrophoresis. *J Mol Biol* **98**:503-517.

Stamp, G. W., Poulsom, R., Chung, L. P., Keshav, S., Jeffery, R. E., Longcroft, J. A., Pignatelli, M., and Wright, N. A. (1992). Lysozyme gene expression in inflammatory bowel disease. *Gastroenterology* **103**:532-538.

Sternberg, N., Sauer, B., Hoess, R., and Abremski, K. (1986). Bacteriophage P1 cre gene and its regulatory region. Evidence for multiple promoters and for regulation by DNA methylation. *J Mol Biol* **187**:197-212.

Suemori, S., Lynch-Devaney, K., and Podolsky, D. K. (1991). Identification and characterization of rat intestinal trefoil factor. *Proc Natl Acad Sci USA* **88**:11017-11021.

Szabo, S., Folkman, J., Vattay, P., Morales, R. E., Pinkus, G. S., and Kato, K. (1994). Accelerated healing of duodenal ulcers by oral administration of a mutein of basic fibroblast growth factor in rats. *Gastroenterology* **106**:1106-1111.

Taketo, A. (1988). DNA transfection of Escherichia coli by electroporation. *Biochim Biophys Acta* **949**:318-324.

Tan, X. D., Hsueh, W., Chang, H., Wei, K. R., and Gonzalez-Crussi, F. (1997). Characterization of a putative receptor for intestinal trefoil factor in rat small intestine: Identification by *in situ* binding and ligand blotting. *Biochem Biophys Res Commun* **237**:673-677.

Tanaka, S., Podolsky, D. K., Engel, E., Guth, P. H., and Kaunitz, J. D. (1997). Human spasmolytic polypeptide decreases proton permeation through gastric mucus in vivo and in vitro. *Am J Physiol* **272**:G1473-G1480.

Taupin, D., Ooi, K., Yeomans, N., and Giraud, A. (1996). Conserved expression of intestinal trefoil factor in the human colonic adenoma-carcinoma sequence. *Lab Invest* **75**:25-32.

Taupin, D., Wu, D. C., Jeon, W. K., Devaney, K., Wang, T. C., and Podolsky, D. K. (1999). The trefoil gene family are coordinately expressed immediate-early genes: EGF receptor- and MAP kinase-dependent interregulation. *J Clin Invest* **103**:R31-R38.

Taupin, D. R., Kinoshita, K., and Podolsky, D. K. (2000). Intestinal trefoil factor confers colonic epithelial resistance to apoptosis. *Proc Natl Acad Sci USA* **97**:799-804.

Taupin, D. R., Pang, K. C., Green, S. P., and Giraud, A. S. (1995). The trefoil peptides spasmolytic polypeptide and intestinal trefoil factor are major secretory products of the rat gut. *Peptides* **16**:1001-1005.

Te Riele H., Maandag, E. R., and Berns, A. (1992). Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc Natl Acad Sci USA* **89**:5128-5132.

Terada, T., Sakagami, R., Tabuchi, Y., and Maeda, M. (2001). Characterization of the mouse TFF1 (pS2) gene promoter region. *Biol Pharm Bull* **24**:135-139.

Theisinger, B., Seitz, G., Dooley, S., and Welter, C. (1996). A second trefoil protein, ITF/hP1.B, is transcribed in human breast cancer. *Breast Cancer Res Treat* **38**:145-151.

Theisinger, B., Welter, C., Grzeschik, K. H., and Blin, N. (1992). Assignment of the gene for human spasmolytic protein (hSP/SML1) to chromosome 21. *Hum Genet* **89**:681-682.

Thiebaud, C. H. and Fischberg, M. (1977). DNA content in the genus Xenopus. *Chromosoma* **59**:253-257.

Thim, L. (1988). A surprising sequence homology. *Biochem J* **253**:309.

Thim, L. (1989). A new family of growth factor-like peptides. 'Trefoil' disulphide loop structures as a common feature in breast cancer associated peptide (pS2), pancreatic spasmolytic polypeptide (PSP), and frog skin peptides (spasmolysins). *FEBS Lett* **250**:85-90.

Thim, L. (1994). Trefoil peptides: a new family of gastrointestinal molecules. *Digestion* **55**:353-360.

Thim, L. (1997). Trefoil peptides: from structure to function. *Cell Mol Life Sci* **53**:888-903.

Thim, L., Jørgensen, K. H., and Jørgensen, K. D. (1982). Pancreatic spasmolytic polypeptide (PSP): II. Radioimmunological determination of PSP in porcine tissues, plasma and pancreatic juice. *Regul Pept* **3**:221-230.

Thim, L. and Mortz, E. (2000). Isolation and characterization of putative trefoil peptide receptors. *Regul Pept* **90**:61-68.

Thim, L., Woldike, H. F., Nielsen, P. F., Christensen, M., Lynch-Devaney, K., and Podolsky, D. K. (1995). Characterization of human and rat intestinal trefoil factor produced in yeast. *Biochemistry* **34**:4757-4764.

Thomas, K. R. and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**:503-512.

Thomas, K. R., Deng, C., and Capecchi, M. R. (1992). High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Mol Cell Biol* **12**:2919-2923.

Tomasetto, C., Masson, R., Linares, J. L., Wendling, C., Lefebvre, O., Chenard, M. P., and Rio, M. C. (2000). pS2/TFF1 interacts directly with the VWFC cysteine-rich domains of mucins. *Gastroenterology* **118**:70-80.

Tomasetto, C., Rio, M. C., Gautier, C., Wolf, C., Hareuveni, M., Chambon, P., and Lathe, R. (1990). hSP, the domain-duplicated homolog of pS2 protein, is co-expressed with pS2 in stomach but not in breast carcinoma. *EMBO J* **9**:407-414.

Tomasetto, C., Rockel, N., Mattei, M. G., Fujita, R., and Rio, M. C. (1992). The gene encoding the human spasmolytic protein (SML1/hSP) is in 21q 22.3, physically linked to the homologous breast cancer marker gene BCEI/pS2. *Genomics* **13**:1328-1330.

Tomasetto, C., Wolf, C., Rio, M. C., Mehtali, M., LeMeur, M., Gerlinger, P., Chambon, P., and Lathe, R. (1989). Breast cancer protein PS2 synthesis in mammary gland of transgenic mice and secretion into milk. *Mol Endocrinol* **3**:1579-1584.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., *et al.*, (2001). The Sequence of the Human Genome. *Science* **291**:1304-1351.

Welter, C., Theisinger, B., Rio, M. C., Seitz, G., Schuder, G., and Blin, N. (1994). Expression pattern of breast-cancer-associated protein pS2/BCEI in colorectal tumors. *Int J Cancer* **56**:52-55.

Welter, C., Theisinger, B., Seitz, G., Tomasetto, C., Rio, M. C., Chambon, P., and Blin, N. (1992). Association of the human spasmolytic polypeptide and an estrogen- induced breast cancer protein (pS2) with human pancreatic carcinoma. *Lab Invest* **66**:187-192.

Wiede, A., Jagla, W., Welte, T., Kohnlein, T., Busk, H., and Hoffmann, W. (1999). Localization of TFF3, a new mucus-associated peptide of the human respiratory tract. *Am J Respir Crit Care Med* **159**:1330-1335.

Williams, R., Stamp, G. W., Gilbert, C., Pignatelli, M., and Lalani, E. N. (1996). pS2 transfection of murine adenocarcinoma cell line 410.4 enhances dispersed growth pattern in a 3-D collagen gel. *J Cell Sci* **109**:63-71.

Wong, W. M., Poulsom, R., and Wright, N. A. (1999). Trefoil peptides. *Gut* **44**:890-895.

Wright, N. A. (2001). Interaction of trefoil family factors with mucins: clues to their mechanism of action? *Gut* **48**:293-294.

Wright, N. A., Hoffmann, W., Otto, W. R., Rio, M. C., and Thim, L. (1997). Rolling in the clover: trefoil factor family (TFF)-domain peptides, cell migration and cancer. *FEBS Lett* **408**:121-123.

Wright, N. A., Pike, C., and Elia, G. (1990b). Induction of a novel epidermal growth factor secreting cell lineage by mucosal ulceration in human gastrointestinal stem cells. *Nature* **343**:82-85.

Wright, N. A., Poulsom, R., Stamp, G., Van Noorden, S., Sarraf, C., Elia, G., Ahnen, D., Jeffery, R., Longcroft, J., and Pike, C. (1993). Trefoil peptide gene expression in gastrointestinal epithelial cells in inflammatory bowel disease. *Gastroenterology* **104**:12-20.

Wright, N. A., Poulsom, R., Stamp, G. W., Hall, P. A., Jeffery, R. E., Longcroft, J. M., Rio, M. C., Tomasetto, C., and Chambon, P. (1990a). Epidermal growth factor (EGF/URO) induces expression of regulatory peptides in damaged human gastrointestinal tissues. *J Pathol* **162**:279-284.

Wurst, W., Auerbach, A. B., and 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**:2065-2075.

Zimmer, A. (1992). Manipulating the genome by homologous recombination in embryonic stem cells. *Annu Rev Neurosci* **15**:115-137.

Diese Arbeit wurde am Institut für Anthropologie und Humangenetik der Universität Tübingen angefertigt. Ein wichtiges Teilergebnis wurde veröffentlicht in:

Kayademir, T., Rosewell, I., Blin, N., and Gött, P. (2000). Genomic structure of the mouse trefoil factor (Tff) gene cluster in 17q. *GeneScreen* 1: 15-19.

Posterbeiträge:

1998 22. DGZ-Tagung (Deutsche Gesellschaft für Zellbiologie) in Saarbrücken
2000 24. DGZ-Tagung in Karlsruhe

Weitere wissenschaftliche Veröffentlichungen:

---

dos Santos, S. E., Kayademir, T., Regateiro, F., Machado, J. C., Savas, S., Dobosz, T., Blin, N., and Gott, P. (1999). Variable distribution of TFF2 (Spasmolysin) alleles in Europeans does not indicate predisposition to gastric cancer. *Hum Hered* **49**: 45-47.

Gött, P., Dos Santos Silva, E., Kayademir, T., Al-azzeh, E., Regateiro, F., Dobosz, T., and Blin, N. (1999). Tumor-associated trefoil peptides exert protective function in premalignant conditions. *Nowotwory* **49**: 26-29.

Kayademir, T., Silva, E. d., Pusch, C., Beck, S., Machado, J. C., and Gött, P. (1998). A novel 25 bp tandem repeat within the human trefoil peptide gene TFF2 in 21q22.3: polymorphism and mammalian evolution. *Eur J Hum Genet* **6**: 121-128.

Blin, N., Beck, S., Machado, J. C., Labouvie, C., Kayademir, T., Schmitt, H., and Gött, P. (1997). Trefoil peptides and their genes: oncogenetic markers and means for gene therapy. *Gin Pol* **68**: 191-198.

Und weitere Beitäge zum Thema:

1997 PCR In Clinical Diagnosis Symposium in Heidelberg
1997 Vortrag bei der Jahrestagung der Genetik in Giessen
2001 11. AEK Kongress der Deutschen Krebsgesellschaft in Heidelberg

Lebenslauf

Persönliche Daten

Name:	Tuncay Kayademir
geboren:	09.12.1970 in Herrenberg, Baden-Württemberg
Staatsangehörigkeit:	türkisch
Familienstand:	ledig
Anschrift:	Keplerstrasse 12, 72074 Tübingen Tel: 07071-22532

Schulbildung

1977-1981	Grundschule, Bondorf (Ba-Wü)
1981-1990	Schickhard-Gymnasium, Herrenberg (Ba-Wü)

Studium

10/1990-7/1991	Studium der Wirtschaftswissenschaften an der Universität Stuttgart-Hohenheim
10/1991-4/1996	Studium der Biologie an der Eberhard-Karls-Universität Tübingen
7/1996-4/1997	Diplomarbeit am Institut für Anthropologie und Humangenetik, Abteilung Molekulare Genetik an der Eberhard-Karls-Universität Tübingen
7/1997-dato	Promotion am Institut für Anthropologie und Humangenetik, Abteilung Molekulare Genetik an der Eberhard-Karls-Universität Tübingen unter Betreuung von PD Dr. Peter Gött und Prof. Dr. N. Blin

Auslandsaufenthalte

8/1995-9/1995	Aufenthalt in der Bosphorus University, Department of Molecular Biology and Genetics, Istanbul, Türkei
12/1998	Aufenthalt am ICRF und ICSM, London, England
11/2000	Aufenthalt am IPATIMUP Porto, Portugal

Im Selbstverlag herausgegeben von:

Tuncay Kayademir
Keplerstr. 12
72074 Tübingen