Chapter 2

Materials and Methods

Materials

2.1 Chemical reagents

All common chemicals were purchased from Sigma Chemical Co., BDH Chemical Ltd., and Difco Laboratories unless specified below or in the text.

Bio-Rad Laboratories	β-mercaptoethanol
Gibco BRL Life Technologies	Foetal Bovine Serum
	Phosphate buffered saline, PBS (pH7.2)
	ultraPURE™ agarose
	Ham's F12 media
Novagen	GeneJuice
	KOD polymerase
Novagiochem	X-gal
	(5-bromo-4-chloro-3-indolyl-β-D-galactoside)
Fluka	Formamide
Sigma	Anisomycin
	Cycloheximide
	Doxycycline
	Puromycin
	Streptomycin
	L-glutamine
	Dulbecco's Modified Eagle Medium (DMEM)
	4', 6-diamidino-2-phenylindole (DAPI) stain
Stratagene	Perfect Match™
	Taq Extender™

2.2 Enzymes and commercially prepared kits

All restriction endonucleases were purchased from New England Biolabs.

Ambion	DNA-free DNase treatment kit
Amersham Biosciences	Megaprime DNA labelling systems
	Sephadex G-50 Nick Columns
	Redivue™[α-^{32}P]-dCTP (AA 005) aqueous solution (370 MB/ml, 10 mCi.ml)
	2'-deoxynucleoside 5' triphosphates (dATP, dTTP, dGTP, dCTP)
BD Biosciences	Advantage-2 PCR Enzyme Mix

Bio-Rad	Bradford Protein Quantification
	Acrylamide Gels
Clontech	Luciferase Reporter Assay
	Human MTN Blot I
	Human MTN Blot II
Fuji	RX 1100 medical X-ray film
Invitrogen	Superscript II cDNA synthesis kit
	DNase I
New England Biolabs	T4 DNA ligase (1 U/μl)
PE Applied Biosystems	Amplitaq™
	AmplitaqFS
	SYBR Green Master Mix
Qiagen	Genomic DNA and DNA gel purification
	Midi- and Maxi- Prep Kits
	RNeasy RNA purification
Sigma Chemical Company	Ribonuclease A
	Deoxyribonuclease I
	DNA polymerase I (10 U/μl)

2.3 Hybridisation membranes, and X-ray and photographic film

Amersham Biosciences	Hybond-C™ Nylon (20 cm x 1 m)
	(used for western blotting)
Fuji	RX 100 Medical X-ray film
Polaroid	Polaroid 667 Professional film

2.4 Solutions and buffers

Solutions used in this thesis are listed below, alphabetically within each section. Final concentrations of reagents are given for most solutions. Amounts and/or volumes used in preparing solutions are given in some cases. Unless otherwise specified, solutions were made in nanopure water.

2.4.1 *Buffers*

10x DNase I Buffer
200 mM Tris-HCl (pH 8.4)
20 mM $MgCl_2$
500 mM KCl

10x Ligase Buffer
500 mM Tris-HCl (pH 7.5)
100 mM Dithiothreitol
100 mM $MgCl_2$

10x PCR buffer (Advantage)
400 mM Tricine – KOH (pH 8.7)
150 mM KOAc
35 mM $Mg(OAc)_2$
37.5 μg/ml BSA
0.05% Tween-20
0.05% Nonidet-P40

10x PCR buffer I
100 mM Tris-Hcl (pH 8.3)
500 mM KCl
15 mM $MgCl_2$

10x PBS pH 7.4
10.6 mM KH_2PO_4
1.5 M NaCl
30 mM Na_2PO_4-$7H_20$
5% v/v β-mercaptoethanol

PBS-T
0.1% v/v Tween 2
1 x PBS

10x TBE
890 mM Tris Base
890 mM Borate
20 mM EDTA (pH 8.0)

1x TE
10 mM Tris-HCl (pH 7.4)
1 mM EDTA

1x $T_{0.1}E$
10 mM Tris-HCl (pH 8.0)
0.1 mM EDTA

TFB I
30 mM KOAc
100 mM $RbCl_2$
10 mM $CaCl_2$
50 mM MnCl
15% v/v Glycerol
pH 5.8

TFB II
10 mM MOPS
75 mM $CaCl_2$
100 mM $RbCl_2$
15% v/v Glycerol
pH 6.5

2.4.2 *Northern blotting solutions*

20 x SSC
3 M NaCl
300 mM Trisodium Citrate

Hybridisation buffer
6 x SSC
1% w/v N-lauroyl-sarcosine
10 x Denhardt's
50 mM Tris-HCl (pH 7.4)
10% w/v Dextran sulphate

100 x Denhardt's Solution
20 mg/ml Ficoll 400-DL
20 mg/ml polyvinylpyrolidine 40
20 mg/µl BSA (pentax fraction V)

2.4.3 *Electrophoresis Solutions and Western Blotting Solution*

Blocking Solution
10% w/v Milk Powder
0.1% v/v Tween 20
PBS

6x Glycerol Dye
30% v/v Glycerol
0.1% w/v Bromophenol Blue
0.1% w/v Xylene Cyanol
5 mM EDTA (pH7.5)

1x Protein Sample Buffer
2% w/v SDS
10% v/v Glycerol
60 mM Tris pH6.8
0.01% w/v Bromophenol Blue

10x Running Buffer
0.25 M Tris
1.92 M Glycine
1% w/v SDS

1x Transfer Buffer
0.025 M Tris
0.192 M Glycine
0.1% w/v SDS
25% v/v Ethanol

2.4.4 *Immunofluorescence Solutions*

Blocking Solution
0.2% w/v Gelatine
0.05% w/v Saponin
PBS

Washing Solution
0.05% w/v Saponin
PBS

Quenching Solution
50 mM NH_4Cl

2.4.5 *Media*

All media were prepared in nanopure water and either autoclaved or filter-sterilised prior to use. When used for bacterial growth, 15 mg/ml bacto-agar was added to the appropriate media. Where appropriate Ampicillin (dissolved in 1 M sodium bicarbonate, stored at -20°C) was added to media at a final concentration of 75 μg/ml.

LB
10 mg/ml Bacto-tryptone
5 mg/ml Yeast extract
10 mg/ml NaCl
pH 7.4

2 X TY
15 mg/ml Bacto-tryptone
20 mg/ml Bacto-peptone
2% w/v dextrose
pH 5.8

LB plates
LB media
15 g/l agar
75 ug/ml Amplicillin

X-gal plates
As for LB plates plus
100 μg/ml Xgal
200 μg/ml IPTG

2.4.6 *General DNA preparation solutions*

GTE
50 mM Glucose
1 mM EDTA
25 mM Tris-HCl (pH 8.0)

3 M K^+/5 M Ac^-
60 ml 5M potassium acetate (pH 4.8)
11.5 ml glacial acetic acid
28.5 ml H_20

2.5 Size Markers

1 kb ladder (1 mg/ml) (Gibco BRL Life Technologies)
Contains 1 to 12 repeats of a 1,018 bp fragment and vector fragments from 75 bp to 1,636 bp to produce the following sized fragments in bp: 75, 142, 154, 200, 220, 298, 344, 394, 516/506, 1,018, 2,036, 3,054, 4,072, 5,090, 6,108, 7,125, 8,144, 9,162, 10,180, 11,198, 12,216.

100 bp ladder (Invitrogen Life Technologies)
The 100 bp ladder consists of 15 blunt ended fragments between 100 and 1500 bp in multiples of 100 with an additional fragment of 2072 bp.

SeeBlue Protein Standard (Invitrogen Life Technologies)
Consists of 10 pre-stained protein bands in the range of 4 – 250 kDa. The proteins and their approximate molecular weights (kDa) are: Myosin – 250, Phosphorylase – 148, BSA – 98, Glutamic Dehydrogenase – 64, Alcohol dehydrogenase – 50, carbonic Anhydrase – 36, Myoglobin Red – 22, Lysozyme – 16, Aprotinin – 6, Insulin, B chain – 4.

2.6 *E.coli* strains

The bacterial strains used in this study are listed in Table 2.1.

Table 2.1 Strains of *E. coli* used in this study

Strain	Source	Genotype
JM109	Clontech	*e14-(McrA-) recA1 endA1 gyrA96 thi-1 hsdR17*(rK- mK+) *supE44 relA1 Δ(lac-proAB)* [F' *traD36 proAB lacI*^{q}*ZΔM15]*
DH5α	Invitrogen	FΦ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 *ton*A
DH10b	Invitrogen	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ - *rps*L *nup*G *ton*A

2.7 Mammalian Cell Lines

The mammalian cell lines used in this study, together with their associated experiments and growth conditions are listed in Table 2.2.

Table 2.2 Mammalian cell lines used in this study

Cell Line	Cell Line Description	Use	Media *	Source
Cos-7	African Rhesus Monkey (kidney)	Intracellular Localisation	DMEM (Sigma)	Dr J Collins
CHO-K1	Chinese Hamster Ovary	Intracellular Localisation	Ham's F12 (Gibco)	Dr P Couttet
CHO-AA8-Luc-Off	Chinese Hamster Ovary transfected with tetracycline regulatory element	Transcript Stability	DMEM (Gibco)**	Dr P Couttet
Hek293FT	Human embryo kidney	Translation Inhibition	DMEM (Sigma)	Dr J Collins

* Media was supplemented with 10% v/v foetal bovine serum (Gibco, BRL), 100 U/ml Penicillin (Sigma), 100 μg/ml Streptomycin (Sigma), and 2 mM L-glutamine (Sigma).
** Media was supplemented with 10% v/v tetracycline approved foetal bovine serum (Clontech), 100 U/ml Penicillin (Sigma), 100 μg/ml Streptomycin (Sigma), and 2 mM L-glutamine (Sigma).

2.8 RNA samples

2.8.1 *Sources of human total RNA*

Total RNA was obtained from Ambion, Clontech and Stratagene. Tissue origins are given overleaf in Table 2.3. All commercial samples were extracted from single tissues and are not pools of samples.

Table 2.3 Total RNA samples used in this study

Supplier	Tissue panel number	Tissue	Supplier	Tissue panel number	Tissue
Clontech	1	Adrenal gland	Ambion	26	Cervix
Clontech	2	Bone marrow	Ambion	27	Colon
Clontech	3	Brain (cerebellum)	Ambion	28	Heart
Clontech	4	Brain (whole)	Ambion	29	Kidney
Clontech	5	Foetal brain	Ambion	30	Liver
Clontech	6	Foetal liver	Ambion	31	Lung
Clontech	7	Heart	Ambion	32	Ovary
Clontech	8	Kidney	Ambion	33	Pancreas
Clontech	9	Liver	Ambion	34	Placenta
Clontech	10	Lung	Ambion	35	Prostate
Clontech	11	Placenta	Ambion	36	Skeletal muscle
Clontech	12	Prostate	Ambion	37	Small intestine
Clontech	13	Salivary gland	Ambion	38	Spleen
Clontech	14	Skeletal muscle	Ambion	39	Stomach
Clontech	15	Spleen	Ambion	40	Testis
Clontech	16	Testis	Clontech	41	Thymus
Clontech	17	Thymus	Clontech	42	Bone marrow
Clontech	18	Thyroid gland	Stratagene	43	Foetal stomach
Clontech	19	Trachea	Stratagene	44	Foetal lung
Clontech	20	Uterus	Stratagene	45	Foetal heart
Clontech	21	Foetal brain	Stratagene	46	Foetal kidney
Clontech	22	Foetal liver	Stratagene	47	Foetal skeletal muscle
Ambion	23	Adrenal gland	Stratagene	48	Foetal colon
Ambion	24	Bladder	Clontech	49	Uterus
Ambion	25	Brain	Clontech	50	Thymus

2.8.2 *Additional sources of total RNA*

Total RNA from *S. pombe* was a gift from Dr. J. Bähler (WTSI).

2.9 cDNA libraries

Nineteen different cDNA libraries were used in the study (see Table 2.4). cDNA libraries were imported and maintained by Jacqueline Bye. Each library contains 500,000 cDNA clones, divided into 25 pools of 20,000 clones. Five pools were combined to form a superpool that contained 100,000 clones. Prior to their use in PCR, each superpool was diluted 1:100 in $T_{0.1}E$.

Table 2.4 cDNA libraries used in this study

cDNA library code	cDNA library description	Vector	Source/ Reference
1. U	Monocyte NOT activated-from a patient with promonocytic leukaemia (U937+)	pCDM8	Simmons (1993)
2. H*	Placental, full term normal pregnancy (H9)	pH3M	Simmons (1993)
3. P	Adult brain	pCDNA1	Pfizer
4. DAU	B lymphoma (Daudi)	pH3M	Simmons (1993)
5. FB	Fetal brain	pCDNA1	Invitrogen
6. FL	Fetal liver	pcDNA1	Invitrogen
7. HL	Peripheral blood (HL60)	pCDNA1	Invitrogen
8. SK	Neuroblastoma cells	pCDNA1	Invitrogen
9. T	Testis	pCDM8	Clontech
10. FLU	Fetal lung	pCDNA1	Invitrogen
11. AL	Adult lung	pCDNA1	Clontech
12. UACT*	(Monocyte PMA activated - from a patient with promonocytic leukaemia) (U937act)	pCDM8	Simmons (1993)
13. YT*	HTLV-1+ve adult leukaemia T cell	pH3M	Simmons (1993)
14. NK*	Natural killer cell	pH3M	Simmons (1993)
15. HPB*	T cell from a patient with acute lymphocytic leukaemia (HPBALL)	pH3M	Simmons (1993)
16. BM*	Bone Marrow	pH3M	Simmons (1993)
17. DX3*	Melanoma	pH3M	Simmons (1993)
18. AH	Adult Heart	pcDNA3-Uni	Invitrogen
19. SI **	Small Intestine	pcDNA3	Stammers

* Generously provided by Dr Simmons, Oxford (Simmons *et al.*, 1993)
** Generously provided by Dr Stammers (Sanger Institute)

2.10 Primer sequences

Appendices I to III list the STSs (sequence tag sites) used in this thesis and give the sequence of each primer. Where appropriate, the clones, or genes from which the STSs were derived are also listed.

Primers were synthesised in house by Dave Fraser or externally by Sigma.

2.11 Key World Wide Web addresses

Table 2.5 Key world wide web addresses used in this study

Website	Address
Baylor College of Medicine Search Launcher	http://searchlauncher.bcm.tmc.edu/
CCDS	http://www.ncbi.nlm.nih.gov/CCDS/
Dotter	http://www.cgr.ki.se/cgr/groups/sonhammer/ Dotter.html
EBI	http://www.ebi.ac.uk/
EBI- ClustalW	http://www.ebi.ac.uk/clustalw/index.html
EMBOSS	http://www.hgmp.mrc.ac.uk/Software/EMBOSS/
Ensembl	http://www.ensembl.org
Eponine	http://servlet.sanger.ac.uk:8080/eponine/
First Exon Finder	http://genome.ucsc.edu/cgi- bin/hgTrackUi?hgsid=42528583&c=chrX&g=firstEF
Gap4	http://staden.sourceforge.ent/overview.html
Gene Expression Atlas	http://expression.gnf.org/cgi-bin/index.cgi#Q
NCBI	http://www.ncbi.nlm.nih.gov/
NCBI - BLAST server	http://www.ncbi.nlm.nih.gov/BLAST/
NCBI - Entrez	http://www.ncbi.nih.gov/Entrez/
NCBI - Locus Link	http://www.ncbi.nlm.nih.gov/projects/LocusLink/
NCBI - Spidey	http://www.ncbi.nlm.nih.gov/spidey/
NCBI - UniGene	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=un igene
OMIM	http://www3.ncbi.nlm.nih.gov/Omim/
Primer3	http://frodo.wi.mit.edu/cgi- bin/primer3/primer3_www.cgi
Prosite	http://au.expasy.org/prosite/
PSORT	http://psort.nibb.ac.jp/
RepeatMasker	http://ftp.genome.washington.edu/RM/ webrepeatmaskerhelp.html
SpliceSiteFinder	http://www.genet.sickkids.on.ca/~ali/splicesitefinder. html
The HGMP Resource Centre	http://www.hgmp.mrc.ac.uk/
The Wellcome Trust Sanger Institute	http://www.sanger.ac.uk/
UCSC Genome Browser	http://genome.ucsc.edu/
Vega	http://vega.sanger.ac.uk/
zPicture	http://zpicture.dcode.org/

Methods

2.12 Mammalian cell culture

The mammalian cells used in this study are listed in Table 2.2.

2.12.1 *Growing and harvesting cells*

All cell lines were adherent and were grown at 37°C, 5% CO_2 in either 75 cm^2 or 150 cm^2 flasks, 8-well cell culture slides, or 6-well dishes.

When a cell density of greater than 80% confluency was reached, cells were sub-cultured. Briefly, the media was aspirated and the cells were washed twice in PBS. To detach the cells from the flask/wells the cells were then incubated with approximately 0.07 volumes of trypsin (Sigma) at 37°C for 2 to 5 minutes. Approximately 0.7 volumes of pre-warmed media were then added to inactivate the typsin. Cell clumps were disrupted with gentle pipetting. From here, the cells were distributed into fresh flasks for sub-culturing or were counted and aliquoted for future experimental analysis.

Cells were counted with a haemocytometer (Sigma) with a 0.1 mm sample depth and light microscope (Olympus).

For frozen storage, cells were harvested at 200 x g following trypsinisation. Cell pellets were resuspended at a density of approximately 3 x 10^6 cells per ml in 10% DMSO in foetal bovine serum in polypropylene vials. Vials were frozen at 1°C/per minute for 12 to 24 hours before permanent storage in the gas phase of a liquid nitrogen vessel (at -180°C). Cells were recovered by rapid thawing at 37°C. They were washed and then mixed directly with 10 to 15 ml of complete medium.

2.12.2 *Transfection*

On the day prior to transfection, cells were trypsinized and plated to the required density with the media, FCS and antibiotics. The next day, the appropriate media was mixed with GeneJuice and incubated at room temperature for between 5 and 15 minutes (volumes and amount of DNA used are listed in Table 2.6). DNA was added to the GeneJuice/media mix and incubated for a further 5 to 15 minutes at room temperature. This was then added to the cells in a dropwise manner, and was gently swirled to mix.

Cells were then grown at 37°C, 5% CO_2 with humidity for a further 24 to 48 hours.

Table 2.6 Amounts of DNA and reagents used in mammalian cell transfections.

Dish	Media (µl)	GeneJuice (µl)	DNA (µg)
8-well culture slide	20	0.75	0.25
24-well cell culture dish	20	0.75	0.25
6-well cell culture dish	100	4	2

2.13 RNA Manipulation

All reagents for RNA work were prepared with Diethylene Pyrocarbonate (DEPC) treated water. Bench surfaces, tubes and labware were cleaned before use with RNaseZap (Ambion).

2.13.1 *Preparation of RNA from cellular extracts*

Total RNA was prepared from mammalian cell pellets using the RNeasy RNA extraction kit (Qiagen) in accordance with the manufacturer's protocol. RNA was eluted with 50 µl of DEPC-treated water.

Following elution, the integrity of the RNA was confirmed by visualisation on 1% agarose gels using ethidium bromide staining. The concentration of RNA was determined by spectroscopy (GENE QUANT) where an absorbance of 1 at 260 nm equates to a concentration of 40 ng/µl. A_{260}:A_{280} ratios were also calculated for each sample. Samples with ratios smaller than 1.7 or greater than 1.9 were discarded.

2.13.2 *DNase treatment of total RNA*

Residual plasmid or genomic DNA was removed from RNA samples using DNase I (Invitrogen). Briefly, samples (up to 2 µg) of total RNA were equilibrated with 10 x DNase buffer to which 1 U of DNase was added. Samples were then incubated for 15 minutes at room temperature and the enzyme was denatured by incubation at 65°C for 10 minutes.

2.13.3 *cDNA synthesis*

cDNA was synthesized from (DNase treated) total RNA (0.5 to 5 µg) using Superscript II (Invitrogen). Where appropriate, and as outlined in the text, either random hexamers (50 ng) or Oligo(dT)$_{12-18}$ primers (500 ng) were used to prime the cDNA synthesis. The synthesis was completed in accordance with the

manufacturer's instructions, with an incubation temperature of 25ºC (random hexamers) or 42ºC (Oligo(dT)$_{12-18}$).
The resulting cDNA was diluted to 400 μg/μl and was stored at -20ºC.

2.13.4 *Northern blotting*

2.13.4.1 Probe Preparation

Approximately 25 ng of purified, nested PCR product were randomly labelled with [α-^{32}P] dATP using the Megaprime DNA labelling Kit (Amersham Biosciences) in accordance with manufacturers instructions.

Incorporation of the label was confirmed by spotting 2 μl of radiolabelled probe onto a Polygram 300 PEI/UV Thin Layer Chromatography (TLC) plate in 1M KH_2PO_4 (pH3). TLC plates were exposed to autoradiography film for approximately 1 hour.

Unincorporated label was then removed by elution on a Sephadex G50 column (Pharmacia Biotech). The probe was diluted to 400 μl in $T_{0.1}E$ and applied to the column. Five 400 μl fractions were collected. The amount of label in each fraction was monitored in a scintillation counter, with fraction 2 containing the labelled probe. All probes were denatured at 95 °C for 5 minutes prior to hybridisation.

2.13.4.2 Hybridisation

Northern blots were pre-hybridised in 50 ml of hybridisation buffer for at least 2 hours at 65 °C in Hybaid tubes with gentle rotation. Twenty-five ml of hybridisation buffer were removed and the radiolabelled probe was added. The blots were hybridised at 65 °C overnight with gentle rotation (approximately 16 hours).

Following overnight hybridisation, the hybridisation solution was then discarded and the blots were washed under the following conditions:

2 x SSC	room temperature	2 x 5 minutes
2 x SSC, 1 % sarcosyl	65 °C	2 x 30 minutes
1 x SSC, 1 % sarcosyl	65 °C	1 x 30 minutes
0.5 X SSC, 1 % sarcosyl	65 °C	1 x 20 minutes
0.2 x SSC	room temperature	2 x 5 minutes.

All washes were carried out with gentle shaking.

Blots were washed until the signal from a Geiger counter dropped below ~ 5cps. Blots were wrapped in Saran Wrap (Dow Chemical Co.) and were then exposed to pre-flashed autoradiography film at –70 °C.

2.13.4.3 Removal of radiolabelled probe from Northern blots.

Blots were stripped of any remaining hybridised probe by incubation in 5 mM Tris-HCl pH7.4 at 75 °C for 1 hour. Removal of the probe was confirmed by exposing the blots to pre-flashed autoradiography film for 1 week.

2.14 DNA manipulation

2.14.1 *Purification of DNA*

Ethanol precipitation

DNA was precipitated from solution by the addition of 0.1 volumes of 3M sodium acetate and either 2 volumes of 100% ethanol, or 0.7 volumes of isopropanol. Samples were centrifuged for 20 minutes at 13,000 x g. The pellet was washed with 1 volume of 70% ethanol, followed by centrifugation for 5 minutes at 13,000 x *g*. DNA pellets were air-dried and resuspended in $T_{0.1}E$.

Direct PCR product purification

PCR products (10 to 50 µl) were purified for subsequent analysis using the PCR purification kit (Qiagen) in accordance with the manufacturer's protocol.

ExoSAP purification of PCR products

PCR reactions were equilibrated with 0.1 volumes of 10 x EXOSAP buffer, 1 U/µl Shrimp Alkaline Phosphatase and 1 U/µl exonuclease I.
Reactions were incubated at 37°C for 30 minutes, followed by 70°C for 15 minutes.

Agarose gel purification

Following visualisation by agarose gel electrophoresis, the appropriately sized DNA fragment was excised from the gel using a clean scalpel. Purification proceeded using the gel purification kit (Qiagen) in accordance with the manufacturer's instructions.

2.14.2 *Alkaline phosphatase treatment of DNA*

The removal of the 5' phosphate group from up to 10 µg of digested DNA was completed using 50 units of Shrimp Alkaline Phosphatase (SAP) in SAP buffer. Samples were incubated at 37°C for 1 hour.

2.14.3 *Restriction digests*

Restriction digests of plasmid DNA (up to 10 µg) were completed using the appropriate buffer, and 20-50 units of enzyme. Samples were incubated at 37°C for 2 hours, and the resulting digest was confirmed by agarose electrophoresis.

2.14.4 *Mini-preps of plasmid DNA*

A single colony was inoculated in up to 10 ml of LB broth containing the appropriate antibiotic and grown overnight at 37°C, shaking at 250-300 rpm. On the following day the cells were pelleted at 1500 x *g*, and resuspended in 200 µl GTE on ice. To this, 400 µl of fresh 0.2M NaOH and 1% SDS were added with gentle inversion of the tube. The tube was left on ice for 5 minutes. Three hundred microlitres of 5M Acetate, 3M K^+ were added. The tube was inverted to mix, and the samples were incubated on ice for at least 10 minutes. The precipitate was pelleted by centrifugation at 13,000 x *g*. This procedure was repeated if the resulting supernatant was still cloudy. The DNA was pelleted by the addition of 0.7 volumes of isopropanol and subsequent centrifugation at 13,000 x *g*.

The DNA was then ethanol precipitated (see section 2.14.1), and washed with 70% ethanol before being resuspended in 30 µl $T_{0.1}E$.

2.14.5 *Midi and maxi preps of plasmid DNA*

A single colony was inoculated into 5 ml LB broth containing the appropriate antibiotic and grown for 8 hours at 37°C, shaking at 20-300 rpm. This inoculation was diluted 1:500 into 100-500 ml LB broth containing the same antibiotic and grown overnight at 37°C, shaking at 200 to 300 rpm.

Purification of plasmid DNA was completed using the appropriate kit (Qiagen Midi- and Maxi Prep Kits) in accordance with the manufacturer's protocol.

2.14.6 *Quantification*

Absorbance Spectroscopy

DNA was quantified by applying the Beer-Lambert equation relating absorbance, and extinction co-efficient to DNA concentration. Absorbance readings were measured at 260 nm and the extinction coefficients used were 50 for dsDNA and 33 for ssDNA. Absorbance readings were taken on either Gene Quant (Biochrom Ltd.) or BioQuant (Eppendorf).

Pico Green Quantification

Double stranded DNA was quantified using PicoGreen® dsDNA Quantitation Reagents (Molecular Probes). The procedure was completed in accordance with the manufacturer's protocol. Lambda DNA (Molecular Probes) was used to construct the standard curve in the range of 0-200 ng/µl from which the concentration of unknown samples was extrapolated. Absorbance readings were performed using a luminescence spectrofluorimeter (BioQuant) measuring at a wavelength of 480 nm for excitation and 520 nm for emission.

2.15 Polymerase Chain Reaction

Primers were designed using Primer3 (section 2.11) or Primer Express software packages (Applied Biosystems). Where possible primers were selected using the following parameters:

- Melting temperature between 57°C and 63°C.
- G/C content between 30-80%
- Length between 17 – 22 bp
- Less than 2°C difference in melting temperature between the two primers

ePCR or BLAST searches were performed to test the specificity of the primer pair.

PCR was performed in either a 96-well micro-titer plate (Costar Thermowell™ M-type plate or 0.2 µl tubes (Falcon) in a PTC-225 (MJ Research) thermocycler. Unless stated otherwise template DNA was amplified in a reaction volume of 15 µl. Reactions contained approximately 1.3 µM of each oligonucleotide primer, 67 mM Tris-HCL (pH 8,8), 16.6 mM $(NH_2)SO_4$, 6.7 mM $MgCl_2$, 0.5 mM of each deoxyribonucleoside triphosphate (dATP, dCTP, dGTP, dTTP), 1.5 U of Amplitaq™ (Cetus Inc.). 10 mM β-mercaptoethanol and 170 µg/ml of BSA (Sigma Chemical Co., A-4628) were also added to the reactions.

Unless specified otherwise, cycling conditions were as follows: an initial denaturing step of 5 minutes at 94 °C, followed by thermal cycling at 94 °C for 30 seconds, [primer-specific annealing temperature] for 30 seconds, and 72 °C for 30 seconds. A final extension step of 5 minutes at 72 °C completed the amplification reaction. PCR products were then separated on 2.5% agarose gels and were visualised by ethidium bromide staining.

2.15.1 *STS pre-screen*

STS pre-screens were performed using 5 ng of the following templates: human genomic DNA, Clone 2D (a hybrid containing only X chromosome DNA), hamster DNA and a negative control. Pre-screens were performed using three different annealing temperatures (55 °C, 60 °C, 65 °C) to determine the cycling parameters that will give a visible and specific DNA product. Buffer and PCR conditions are described above (section 2.15) and section 2.4.1.

2.15.2 *cDNA library screening*

Nineteen different cDNA libraries were subdivided into 25 subpools of 20,000 clones which were then combined to produce 5 superpools of 100,000 clones (J.Bye). The cDNA libraries are listed in Table 2.4.

To screen for expression, aliquots of the superpools of each library were arranged in a micro-titre plate to facilitate subsequent manipulation and gel-loading post PCR with a multichannel pipetting device. Five microlitres of each superpool were used as template in a 15 μl final reaction volume in the primary screens. Buffer and PCR conditions are described in section 2.15 and section 2.4.1.

2.15.3 *Vectorette PCR*

Vectorette PCR on cDNA libraries.

Vectorette PCR was performed on the cDNA library superools (listed in Table 2.4). PCR was performed using 5 μl of the diluted superpools (1:100 dilution in $T_{0.1}E$) as the template in a 15 μl final reactions volume using buffer conditions as described in section 2.15.1) Primer combinations were as follows: universal primer 224 and specific primer A, universal primer 224 and specific primer B for each STS.

The PCR was performed in a DNA thermocycler (Omnigene) using a hot start. Cycling conditions were as follows: an initial denaturation step of 5 minutes at 95 °C, followed by 17 cycles of: 94 °C for 5 seconds, 65 °C for 30 seconds and 72 °C for 3 minutes; followed by cycles of 94 °C for 5 seconds, 60 °C for 30 seconds and 72 °C for 3 minutes. An incubation of 72 °C for 5 minutes completed the reaction. The PCR was paused after 4 minutes of the initial denaturation and 2 μl of Taq premix (containing 0.12 units Amplitaq, 0.12 units Taq Extender and 0.l2 units of Perfect Match, 10 % sucrose + cresol red, and $T_{0.1}E$) were added to each reaction.

Re-amplification of vectorette PCR products.

In reactions where multiple bands or weaker bands were observed, bands were excised and placed in 100 μl of $T_{0.1}E$ for at least 10 hours. Re-amplification of each band was carried out by PCR using 5 μl of $T_{0.1}E$ taken from the 100 μl containing the excised band followed by the addition of PCR reagents as described in section 2.15.1). Cycling conditions were as follows: an initial denaturation step of 5 minutes at 95 °C followed by 25 cycles of 94 °C for 5 seconds, 60 °C for 30 seconds and 72 °C for 3 minutes, followed by 72 °C for 5 minutes.

The vectorette PCR products were separated by electrophoresis in 2.5% agarose gels and were visualised by ethidium bromide staining. Products were gel purified using the Qiagen gel extraction kit prior to sequencing.

2.15.4 *Colony PCR*

Following bacterial transformation (section 2.17.3) individual white colonies were picked using a sterile toothpick and resuspended in 100 µl of sterile water. PCR was performed using 5 µl of the resuspended colony as the template DNA. Different primer combinations were used to either confirm the presence and size of an insert, or to confirm the presence and orientation of the insert. These primer combinations are listed in Appendices II and III.

Reaction products were visualised by agarose gel electrophoresis and staining with ethidium bromide (section 2.16.1).

2.15.5 *RT-PCR*

PCR was performed using up to 5 µl of single stranded cDNA synthesised from total RNA (section 2.13.3) with the final reaction volume ranging between 15-25 µl. Included in the reaction were 0.01 volumes of 50 x Advantage 2 PCR Enzyme System (BD Biosciences) which was used with 0.1 volumes of 10 x Advantage 2 buffer, 1.5 U of Amplitaq™ (Cetus Inc.), 0.4 mM of each deoxyribonucleoside triphosphate (dATP, dCTP, dGTP, dTTP), and 0.5 µM of each oligonucleotide primer.

Unless otherwise noted in the text, 35 cycles of PCR were performed. Cycling conditions were as follows: an initial denaturation step of 5 minutes at 95 °C, followed by 35 cycles of: 94 °C for 30 seconds, 65 °C for 30 seconds and 72 °C for

30 seconds to 3 minutes depending on the expected size of the PCR product. The extension time was increased by 1 min per kb when longer PCR products were expected.

2.15.6 *Quantitative PCR*

Quantitative PCR reactions were performed on an ABI7000, using SYBR-Green Master Mix (Applied Biosystems), ABI PRISM® 96 well optical reaction plates and ABI PRISM™ optical adhesive plate sealers. All reactions were completed in triplicate, with minus RT controls. Each 25 µl PCR reaction contained 0.02 to 0.1 µg cDNA (section 2.13.3), 2 x SYBR-Green Master Mix, and primers diluted to a final concentration of 0.5 µM.

The following cycling parameters were used: 50ºC for 10 minutes, then 95ºC for 10 minutes. This was followed by 40 cycles of 95ºC for 10 seconds and a combined annealing/extension temperature of 60ºC for 2 minutes. During each cycle of the PCR the fluorescence emitted by the binding of SYBR-Green to the dsDNA produced in the reaction was measured. To confirm the specificity of the reactions dissociation curves were constructed for each primer pair at 0.1ºC intervals between the temperatures of 60 ºC and 95ºC.

Analysis of Quantitative PCR

The SYBR-Green fluorescent spectra collected during the PCR were analysed using the Sequence Detection System Software (ABI). Firstly, background threshold levels were set at the number of cycles before any SYBR-Green fluorescence was detected. The detection threshold was set at the point where the increase in SYBR-Green fluorescence became exponential. Assuming specific amplification, the cycle number at which the sample's fluorescence intersected with the detection threshold, was directly proportional to the amount of DNA in the sample, and was expressed at C_T values. Two different methods were employed to quantify PCR products, absolute and relative concentration.

For absolution concentration analysis, standard curves were generated using a known amount of template (purified plasmid DNA). All unknown samples were then plotted on the standard curve and from which the concentration of DNA in the samples was extrapolated. This method was employed to determine the relative abundance of *PQBP1* alternative transcripts.

Determination of the relative abundance was achieved using a ubiquitously expressed gene as a calibrator. Calibrators used in this thesis, and their primer sequences are listed in Table 2.7. This approach requires the calibrator/sample reactions to have the same amplification efficiency which was determined by titrating the calibrator and sample 1,000 fold, where the gradient of the titration series equates to the amplification efficiency of the reaction. Calibrator/sample primer pairs with similar amplification efficiencies (< 0.01) were used for further analysis.

Table 2.7 Real time PCR control primers

Calibrator	Species	Forward Primer ($5' \rightarrow 3'$)	Reverse Primer ($5' \rightarrow 3'$)
GAPDH	*H. Sapiens*	GAAGGTGAAGGTCGGAGTC	GAAGATGGTGATGGGATTTC
β-actin	*C. griseus*	ACCAACTGGGACGACATGGAGAAGA	TACGACCAGAGGCATACAGGGACAA

The calculation for quantitation first determined the difference (ΔC_T) between the C_T values of the target and the calibrator:

$$\Delta C_T = C_T \text{ (target)} - C_T \text{ (calibrator)}$$

This value was calculated for each sample after which one sample (either time = 0 for time course experiments or brain cDNA for expression profiles) was designated as the reference sample. The comparative ($\Delta\Delta C_T$) calculation was then used to determine the difference between each sample's ΔC_T and the reference's ΔC_T.

$$\text{Comparative expression level} = \Delta C_T \text{ target} - \Delta C_T \text{ reference}$$

Finally, these values were transformed to absolute values using the formula:

$$\text{Absolute comparative expression level} = 2^{-\Delta\Delta CT}$$

2.16 Electrophoretic analysis of DNA, RNA and proteins

2.16.1 *Agarose electrophoresis*

Electrophoresis was carried out using gels containing agarose (BioRad, UK) melted in 1x TBE (section 2.4.3). The concentration of agarose ranged between 0.8% and 2.5% w/v depending on the resolution of separation required. Electrophoresis was

performed in 1x TBE with a voltage ranging between 25-150 V, for between 15-120 minutes depending on the resolution required.

2.16.2 *SDS-PAGE*

Denaturing polyacrylamide gels (15% w/v) were purchased from BioRad, UK. Gels were run in a mini-gel apparatus (BioRad, UK) in 1x Running Buffer (section 2.4.3) at 150 V for 45 minutes.

2.17 Bacterial cloning

2.17.1 *Preparation of chemically competent E.coli*

On the day before preparation an inoculation of a single colony representing the appropriate bacterial strain was established in LB broth and grown at 37°C for at least 12 hours. Cells were diluted 1:500 in LB broth and were grown until the A_{550} was between 0.4 and 0.7. Cells were then pelleted by centrifugation at 4000 x g for 20 minutes.

All subsequent procedures following collection of the bacterial cells were completed at 4°C with minimal agitation to the cells to preserve their viability.

Cells were then gently resuspended in 0.2 volumes TFB I and were again pelleted by centrifugation at 4,000 x g for 20 minutes. The cell pellets were then resuspended in 0.02 volumes TFB II. Aliquots of the prepared cells were stored at –70°C for use in bacterial transformations. To ensure that the cells were chemically competent, test transformations were completed used 1 µg of pUC control plasmid (Clontech).

2.17.2 *Subcloning*

The vector and insert to be used were digested with appropriate restriction enzymes (section 2.14.3). The products were gel purified (section 2.14.1) and the concentration of each product was estimated by visualisation on agarose gels.

A 10 µl ligation reaction was prepared using an approximate 3:1 molar ratio of the insert and vector (roughly 150 ng insert and 50 ng vector) together with 0.1 volumes of 10 x ligation buffer, and 1 unit of T4 DNA ligase. The reaction was incubated at 4°C for at least 12 hours.

2.17.3 *Transformations*

Between 20-50 µl of competent cells were thawed on ice, prior to incubation with 2 to 6 µl of ligation mix (section 2.17.2) for 20 minutes (also on ice).The cells were heat-shocked at 42°C for 45 seconds followed by a brief incubation on ice for 5 minutes. Following this, 950 µl of LB were added to each transformation, which was then incubated at 37°C degrees for 1 to 1.5 hours. The transformation mix was then plated out in varying quantities onto agar plates containing amplicillin and where necessary IPTG and Xgal. The plates were incubated overnight at 37ºC. Individual white colonies were resuspended in either 100 µl of sterile water for colony PCR (section 2.15.4), or 1 to 5 ml of media to harvest the plasmids (section 2.14.4 and 2.14.5).

2.18 Identification of transcript variants

PCR products were purified by direct PCR purification (section 2.14.1), and were ligated to pGEM®T- Easy vector (50 ng) (section 2.17.2) before being chemically transformed into JM109 cells (section 2.17.3). Colony PCR (section 2.15.4) followed by agarose electrophoresis with ethdium bromide stainin was completed to confirm the presence and the size of the inserts.

2.19 *PQBP1* open reading frame cloning

Nested primers were designed to amplify the open reading frame of *PQBP1*, and 16 bp upstream from the translation start site and 20 bp downstream from the stop codon. KOD, a proof reading DNA polymerase (Novagen), was used for all amplification procedures.

The first round of PCR was completed on the panel of cDNA from 20 different human tissues (section 2.13.3) using the following cycling profile: 94°C for 2 minutes, followed by 35 cycles at 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 2 minutes, finally followed by 1 cycle at 72°C for 5 minutes. The reaction mix was diluted 1:200 with $T_{0.1}E$. Five microlitres of the diluted reaction mix were used as a template for the second round of PCR.

An additional round of PCR was completed using the internal set of nested primers and the number of amplification cycles was decreased from 35 to 25. Reaction products were purified directly as described in section 2.14.1 and visualised by agarose electrophoresis and staining with ethidium bromide (section 2.16.1).

2.19.1 *A-tailing of purified PCR products.*

The 3' end of PCR products generated using KOD polymerase were adenylated prior to ligation with pGEM®T-Easy. Between 2 and 6 µl of purified PCR product were incubated with 0.1 mM dATP and 0.12 units of *Taq* polymerase for 30 minutes at 37°C. The reaction was terminated by incubation at 70°C for 15 minutes.

2.19.2 *Ligation into pGEM®T-Easy.*

A-tailed PCR products were cloned into pGEM®T-Easy as outlined in section 2.17.2 and transformed into DH10B cells (section 2.17.3). Colony PCR (section 2.15.4) was performed to confirm the presence and the size of the inserts.

2.20 Preparation of constructs for immunofluorescence

The vectors pCDNA3.NT7 and pCDNA3.CT7 containing a T7 epitope sequence were a kind gift from Dr J. Collins. These vectors express a fusion protein consisting of the desired open reading frame and the T7 epitope (at either the amino- or carboxy- terminal). Here, the insert was ligated after digestion with restriction enzymes restriction enzyme sites located either side of the T7 tag.

2.20.1 *N-terminal T7 tag*

PCR using a proofreading enzyme (KOD) was used to introduce *Not*I and *Xba*I restriction sites flanking the open reading frame of the various *PQBP1* clones made in section 2.19. Following amplification, the PCR products were restriction digested with *Xba*I and *Not*I (2.14.3) and agarose gel purified (section 2.14.1). The translation termination codon was retained.

In preparation for ligation with pCDNA.3-Ntag, the vector was also digested with *NotI* and *XbaI* (section 2.14.3) and purified (section 2.14.1). Ligation was performed using 3:1 molar ration of insert and vector (300 ng of insert and 100 ng of vector, section 2.17.2). The ligation was transformed into DH10B and colony PCR (section 2.15.4) confirmed the presence of the inserts.

Individual white colonies were screened for the pCDNA-3-Ntag: PQBP1 plasmids by PCR. Colonies that contained the insert in the correct orientation were harvested and the plasmid purified by midi-prep purification (section 2.14.5). All plasmids were sequenced in order to verify the fidelity of the insert.

2.20.2 *C-terminal T7 tag*

PCR using a proofreading enzyme (KOD) was used to introduce *Hind* III and *Nhe*I restriction sites flanking the open reading frame of the various PQBP1 clones made in section 2.19. Inserts were prepared as outlined previously (section 2.20.1). The stop codon was removed from the PQBP1 ORF to ensure translation of the C-terminal T7 tag.

In preparation for ligation the vector was digested with *Hind*III and *Nhe*I (section 2.14.3) and purified (section 2.14.1). Ligation, and transformation and subsequent analysis of the construct were performed in the same way as described in section 2.20.1.

2.20.3 *Non-directional cloning*

Transcripts containing either a *Not*I, *Xba*I, *Hind*III or *Nhe*I restriction enzyme sites could not use the protocol outlined in sections 2.20.1 and 2.20.2. In such cases, vectors and inserts where digested with XmnI (section 2.14.3). Digested vector was prepared for ligation with the addition of a thymidine residues to generate a 5' overhang. The prepared vector was a kind gift from Dr J. Collins. To complement the thymidine overhang, the 3' end of the insert was adenylated as outlined in section 2.19.1. Ligation, transformation, propagation and sequence verification were completed as before (section 2.20.1).

2.21 Transcript stability assays

2.21.1 *Preparation of inserts*

The *PQBP1* transcripts were excised from the pGEM®T-Easy by restriction enzyme digest with *EcoR*I (section 2.14.3). The excised DNA containing the PQBP1 open reading frame with 24 bp upstream and 17 bp downstream was then gel purified (section 2.14.1).

2.21.2 *Preparation of constructs*

The vector pTRE-TIGHT (Clontech) was prepared for ligation by digestion with *EcoR*I (section 2.14.3) to generate ends complementary to the DNA insert. The digested vector was then treated with alkaline phosphatase (section 2.14.2) to remove the 5'-phosphate group from the exposed ends.

2.21.3 *Ligation of Vector and DNA*

PQBP1 transcripts were ligated with the pTRE-TIGHT (tet-off) vector as outlined in section 2.17.2 and transformed into DH10B cells (section 2.17.3). Colony PCR (section 2.15.4) was used to confirm the presence and orientation of the inserts.

Individual colonies were screened for the incorporation of the pTRE-TIGHT:*PQBP1* DNA construct by restriction digestion with *EcoR*I (section 2.14.3). Colonies that contained that appropriate plasmid were purified by bacterial mini-preps (2.14.4) prior to DNA sequencing to confirm the integrity of the plasmids.

2.21.4 *Transfection of mammalian cells*

CHO-AA8 Tet-off (Luc) cells were plated at a density of 30,000 cells per well on an 6-well culture dish (Falcon). Individual wells were used for each time point as well as each *PQBP1* transcript variant. Cells were incubated for 30 minutes at room temperature, before growing overnight at 37°C, 5% $C0_2$ with humidity. Transfection of the cells proceeded as described in section 2.12.2 using 1 µg of each the pTRE-TIGHT-*PQBP1* construct, and pCMVß plasmid. Transfected cells were grown overnight prior to the addition of doxycycline.

2.21.5 *Time course experiment*

Doxycycline was added to the media of the transfected CHO-AA8 tet-off cells to final concentration of 50 ng/µl and was mixed by gentle swirling. Cells were then incubated at 37°C, 5% $C0_2$ with humidity for the appropriate length of time. Cells were harvested (section 2.12.1) after 0, 1, 2, 3, 4, 6 and 8 hours. Total RNA was harvested from the cells as described in section 2.13.1 and was treated with DNase (section 2.13.2). cDNA was synthesised from 2 µg of total RNA (section 2.13.3).

2.21.6 *Real-time PCR analysis*

Real-time PCR was established using the reagents and conditions described in section 2.15.6. The amount of cDNA used in each reaction varied with the primer combination that was used. The primer combination and amount of cDNA used (given in the amount of total RNA from which the cDNA synthesised) per reaction were: *ActB* (10 ng), PQBP1-Q10 (25 ng), *LacZ* (25 ng).

All real-time PCRs were performed in triplicate using cDNAs prepared in the presence and absence of RT. Some of the experiments were duplicated.

2.21.7 *Analysis of results*

The relative abundance of *PQBP1* transcript variants and *LacZ* cDNA were normalised to the abundance of *ActB* as outlined in section 2.15.6. To correct for varying transfection efficiencies the relative abundance of the *PQBP1* transcripts were normalised to the amount of LacZ.

Changes in the abundance of the *PQBP1* transcripts were then expressed in relation to t=0. mRNA decay plots were determined by plotting the logarithmic value of the transcript abundance against the incubation period with doxycycline. A linear relationship between the log(relative abundance) and the incubation time was observed between 1 and 4 hours and these time points were used to determine the mRNA half lives of the *PQBP1* transcript variants. The first-order rate constants for *PQBP1* degradation (k= mRNA half-life = gradient), correlation coefficient (r^2) and standard error were calculated from linear regression analysis of the mRNA decay plots.

2.22 Translation Inhibition Time Course Analysis

On the day prior to treatment with translation inhibitors, 2 x 10^5 HEK293FT cells were plated into 6 well culture dishes. The cells were grown to 50-80% confluency. Translation inhibitors cycloheximide (CHX, Sigma), anisomycin (ANS, Sigma) and puromycin (PUR, Sigma) were added to the cells at a final concentrations of 100, 10 and 20 µg/ml respectively. Cells were harvested after 0, 1, 2, 4 and 6 hours.
The samples were prepared for analysis by extracting the RNA (section 2.13.1), DNase I treatment (section 2.13.2) and cDNA synthesis (section 2.13.3).

2.22.1 *Real-time PCR analysis*

Real-time PCR was established using the reagents and conditions described in section 2.15.6. The amount of cDNA used in each reaction varied with the primer combination that was used. The primer combination and amount of cDNA used (given in the amount of total RNA from which the cDNA synthesised) per reaction were: *GAPDH* (10 ng), PQBP1-Q10 (25 ng), *PQBP1*- Q2b, Q3, Q4 and Q6 (50 ng).

All real-time PCRs were performed in triplicate using cDNAs prepared in the presence and absence of RT. Each experiment was duplicated.

2.23 Detection of luciferase activity

2.23.1 *Preparation of cell lysates*

On the day prior to treatment with doxycycline, 2 x 10^5 CHO-AA8 Luc Off cells were plated into 6-well culture dishes. The cells were grown to 50 to 80% confluency. Doxycycline was added to the cells in concentrations ranging between 0 to 1,000 ng/µl which were incubated at 37°C, 5% humidity for 4 hours. Following incubation, the media was removed from the cells which were washed twice in PBS prior to lysing. Cells were lysed in 1x lysis buffer (Clontech) at room temperature for 15 to 20 minutes and were centrifuged at 13 000 x *g* to remove the cellular debris. All samples were assayed immediately after lysis.

2.23.2 *Luciferase activity assay*

The luciferase assay was performed in a white opaque 96-well flat-bottomed plate (Costar) in accordance with the manufacturer's protocol (Clontech). Luciferase activity was measured on a luminometer (BioRad, UK). All samples were assayed in duplicate.

Luciferase activity was normalised against the sample's total protein concentration of the sample which was determined using the Bradford assay.

2.23.3 *Bradford Assay*

Total protein concentrations were measured from the cell lysates prepared in section 2.23.1). The Bradford protein quantification assay was performed in accordance with the manufacturer's instructions (BioRad, UK). Here, a standard curve was constructed using bovine serum albumin (BSA, Sigma) between the concentrations 0 to 1 mg/ml. Absorbance readings were taken at 595 nm, and all samples were measured in duplicate.

2.24 Western Blotting

2.24.1 *Preparation of samples for Western blotting.*

On the day prior to transfection, cells were trypsinized and diluted to a concentration of 8 x 10^4 cells/ml. Transfections were carried out as described in section 2.12.2 in a 24-well culture plate.

Cells were grown for an additional 48 hours and were harvested using 1 x protein sample buffer. All samples were denatured by boiling for 5 minutes, prior to loading on a 15% denaturing polyacrylamide gel (Biorad).

2.24.2 *Electrophoresis of proteins using SDS-PAGE*

SDS-PAGE was carried out using a Mini-PROTEAN® Electrophoresis cell (Biorad) using 1 x running buffer. Proteins were resolved using a 15% separating gel, with a 4% stacking gel (Biorad), and SeeBlue Protein standards (section 2.5). Gels were run at 150 V for approximately 70 minutes.

Electrophoretic transfer

Proteins were transferred to a nitrocellulose membrane using the Mini Trans-Blot® Electrophoresis Transfer cell (Biorad) in 1 x transfer buffer. The electrophoretic transfer was performed at 100 V for 1½ hours at 4ºC.

Detection of proteins

Once proteins were transferred onto the nitrocellulose membrane, the membrane was blocked in blocking buffer for 45 minutes. The blot was then incubated with a mouse anti-T7 monoclonal antibody (stock at 1 mg/ml) (Novagen #69522-4) at a dilution of 1/7,500 in blocking buffer for between 1-12 hours. The blot was washed 3 x10 minutes in PBS-T. The secondary antibody, a sheep-anti-mouse-IgG HRP-conjugate (stock at 0.32 mg/ml) (Sigma #67782) was used at a dilution of 1/7,500 in blocking buffer. Again, the blot was washed for 3 x 10 minutes in PBS-T and then 1 x 5 minutes in PBS. Visualisation of immunoreactivity was completed using ECL (NEN) according to the manufacturers instructions. Membranes were wrapped in Saran Wrap and were then exposed to autoradiography film for between 1 and 15 minutes.

2.25 Intracellular localisation

2.25.1 *Transfection of mammalian cells for immunofluorescence*

Cells were plated at a density of 15,000 cells per well on an 8-well culture slide (Falcon). Cells were incubated for 30 minutes at room temperature, before growing overnight at 37°C, 5% $C0_2$ with humidity. Transfection of the cells proceeded as described in section 2.12.2.

2.25.2 *Fixation of cells*

Twenty-four hours after transfection the cells were washed 3 times in 0.5 ml of 250 mM Hepes, and were then fixed in HISTOCHOICE™ MB® (Amresco), for 20 minutes. The fixation reagent was removed by washing the cells four times with PBS. Cells were then quenched in quenching buffer for at least 15 minutes.

2.25.3 *Antibody staining and visualisation*

The cells were rinsed 3 times with PBS before incubating with blocking buffer for 15 minutes. The well dividers were then removed from the slides.

Primary antibody mouse anti-T7 monoclonal antibody (Novagen) was diluted 1:500 in blocking buffer and cells were incubated with the antibody for 45-60 minutes at room temperature. To remove the antibody cells were rinsed twice with washing buffer, and were then washed for 3 x 10 minutes in washing buffer.

Prior to incubation with the secondary antibody, the cells were incubated in blocking buffer for 15 minutes. The secondary antibody, a goat anti-mouse FITC (stock at 1.1 mg/ml) (Sigma #F2012), was diluted 1:100 in blocking buffer and was incubated with the cells for 45-60 minutes.

Cells were again rinsed twice with washing buffer, and were then washed twice for at least 1 hour. Cells were then stained with DAPI (diluted 1:10,000 in PBS). Finally, cells were washed for 2 x 10 minutes in PBS. Coverslips were mounted using Vectashield (vectorlabs). The FITC fluorescein was typically excited by a 488 nm line of an argon laser, and emission was collected at 530 nm using Nikon Eclipse E800 Microscope confocal microscope (Nikon) and images were captured with a Laser Scanning System ' Radiance 2100fs' (Bio-Rad).

2.26 Computational Analysis

Additional details of the computer programmes used within this thesis can be found at the Wellcome Trust Sanger Institute web site (http://www.sanger.ac.uk/Software/). Frequently used programmes are discussed below.

2.26.1 *Xace*

Human X chromosome data and annotation generated in this thesis were stored in Xace, a chromosome-specific implementation of ACeDB. Other ACeDB implementations were used to store mouse annotation data, as described in Chapter 5. ACeDB was originally developed for the *C. elegans* genome project (Durbin and Thierry-Mieg, 1996). Documentation code and data available from http://www.acedb.org.

ACeDB works using a system of windows and presents data in different types of windows according to the type of data. All windows are linked in a hypertext fashion, so that clicking on an object will display further information about that object. For example, clicking on a region of a chromosome map will highlight landmarks mapping to that part of the chromosome; clicking on a landmark will display information about that landmark including landmark-clone associations .

In addition to the data generated by the X chromosome mapping group, Xace also contains displays of published X chromosome maps. Genomic sequence data is also displayed in ACeDB along with the collated results from the computational sequence analysis performed by the Sanger Institute Human Sequence Analysis Group.
Xace can be accessed by following the instructions at: http://www.sanger.ac.uk/HGP/ChrX.

2.26.2 *Blixem*

Individual matches identified as a result of similarity searches using the BLAST algorithm, or matches between sequences of cDNA clones or PCR products amplified from genomic DNA generated as part of the project, were viewed in more detail using BLIXEM. BLIXEM, (Blast matches In an X-windows Embedded Multiple alignment) is an interactive browser of pairwise Blast matches displayed as a multiple alignment. Either protein or DNA matches can be viewed in this way at either the amino acid or nucleotide level respectively. BLIXEM contains two main displays: the bottom display panel shows the actual alignment of the matches to the genomic DNA sequence, and the top display shows the relative position of the sequence being viewed within the context of the larger region of genomic DNA. A program “pfetch” retrieves the record from an external database (e.g. EMBL, SWISSPROT).

2.26.3 *RepeatMasker*

Human repeat sequences were masked using RepeatMasker, a program that screens DNA sequence for interspersed repeats and low complexity DNA sequence (Smit & Green RepeatMasker at http://www.repeatmasker.org). The output of the program is a detailed annotation of the repeats that are present in the query sequence and a version of the sequence with repeats masked by "N" characters. Sequence comparisons are performed by the program cross_match, an implementation of the Smith-Waterman-Gotoh algorithm developed by P. Green. The interspersed repeat databases screened by RepeatMasker are based on the repeat databases (Repbase Update Jurka 2000) copyrighted by the Genetic Information Research Institute.

2.26.4 *GAP4*

The quality of DNA sequences generated in this thesis was assessed using Gap4. This sequence analysis software was written to aid the finishing process during genome sequence acquisition. However, several functions of this programme have also been utilised to for smaller scale sequence analysis. The contig editor used phred confidence values to calculate the confidence of the consensus sequence and identifies places requiring visual trace inspection or extra data. Traces were automatically checked for mutation assignments. Vector clip located and tagged vector segments of sequence reads. Additional information about Gap4 can be found at: http://staden.sourceforge.ent/overview.html.

2.26.5 *Perl Scripts*

Computational analysis of large datasets was also performed using customed perl scripts. The scripts, author and function of these scripts are appropriately noted throughout the text.

2.26.6 *Emboss*

EMBOSS is a free open source software analysis package developed by the molecular biology community. It has over 100 applications for sequence analysis which can be accessed via the url - http://www.hgmp.mrc.ac.uk/Software/EMBOSS/. Emboss applications can be run using a webserver, Jemboss, or over the command line when installed locally. Some of the individual programmes that have been used in this thesis are outline in Table 2.8.

2.26.7 *Excel analysis*

All mathematical analysis was carried out using Microsoft Excel.

Statistical analysis was determined using the TTEST function in Microsoft Excel which returns the probability associated with a Student's t-Test. A two-tailed distribution and two-sample unequal variance test was performed.

Table 2.8 Emboss applications used in this study

Programme	Source	Function
Cons	RFCGR	Creates a consensus from multiple sequences
Cutseq	RFCGR	Removes a specified section from a sequence
Diffseq	RFCGR	Finds differences between nearly identical sequences
eprimer3	RFCGR	Picks PCR primers and hybridisation probes
est2genome	Sanger	Aligns EST and genomic DNA sequences
Geecee	Sanger	Calculates the fractional GC content of nucleic acid sequences
Getorf	RFCGR	Finds and extracts open reading frames
Matcher	Sanger	Local alignment of two sequences
Merger	RFCGR	Merge two overlapping sequences
Newseq	RFCGR	Type in a new short sequence
Restrict	RFCGR	Finds restriction enzyme cleavage sites
Revseq	RFCGR	Reverse and complement a sequence
Seqret	Sanger	Reads and writes a sequence
Seqretsplit	RFCGR	Reads writes and returns sequences in individual files
Splitter	RFCGR	Split a sequence into (overlapping) smaller sequences
Transeq	RFCGR	Translates nucleic acid sequences

2.27 Sequence Analysis

The sequence of finished clones was analysed using a standard protocol at the Wellcome Trust Sanger Institute. Briefly, sequence properties of the clones were determined including GC content, and repeat content using RepeatMasker (2.26.3). Prior to repeat masking, the sequence was analysed for other features such as CpG Islands (G. Micklem unpublished), tandem repeats and tRNA genes. Masked sequences were then used for much of the subsequent prediction of potential coding regions. Firstly, a variety of similarity searches against the public domain DNA and protein databases using the BLAST suite of programmes were performed. In addition, *in silico* analysis was also completed using a number of gene and exon prediction programmes, including Genscan (Burge, 1997) and FGenesH (Solovyev *et al.*, 1995). The analysis was performed automatically by the Human Sequence Analysis Group at the Wellcome Trust Sanger Institute and was collated in Xace (section 2.26.1) for manual examination.

2.27.1 *Transcript annotation*

Transcripts were annotated in accordance with the following criteria:

Gene classification	Description
Known	Identical to known cDNAs or protein sequences
Novel CDS	Has an open reading frame and is identical or homologous to cDNAs from human or proteins from all species
Novel transcripts	Similar to above however it cannot be assigned an unambiguous ORF
Putative	Identical or homologous to human spliced ESTs but do not contain an ORF
Predicted	Based on *ab initio* prediction and for which at least one exon is supported by biological data (unspliced ESTs, protein sequence similarity with mouse or tetradon genomes
Processed pseudogenes	Pseudogenes that lack introns and are thought to arise from reverse transcription of mRNA followed by reinsertion of DNA into the genome.
Nonprocessed pseudogenes	Pseudogenes that contain introns and are produced by gene duplications.

Gene structures were annotated onto the genomic sequence using Xace .

2.27.2 *Alignment of nucleic acid and protein sequences*

Nucleic acid and protein sequences were aligned using the program ClustalW (Pearson, 1990; Peasron and Lipman, 1998) via a web-based server at the EBI (http://www.ebi.ac.uk/), or ClustalX installed locally on a PC unless otherwise noted in the text. User-defined parameters were left at their default settings unless directed otherwise in the text. Alignments were then manually edited and presented using the program GeneDoc (Nicholas *et al.*, 1997).

2.27.3 *Calculation of sequence identities and similarities*

Nucleic acid and protein sequences were aligned as described in section 2.24.1. The “statistics report” function of GeneDoc was then used to calculate and display sequence identities and similarities.

2.27.4 *Phylogenetic analysis of protein sequences*

Protein sequence alignments were subjected to various phylogenetic analyses to estimate their order of relationship. In each case, alignments produced as described in section 2.24.1 were manually edited as necessary to minimise the number of gaps, and the most reliably aligned region of the alignment was then

used for the respective phylogenetic analyses. Any columns within the alignment containing gaps were removed prior to phylogenetic analysis.

Phylogenetic analyses were performed using the program Phylo-win (Galtier *et al.,* 1996) installed locally on a PC. This package combines various phylogenetic analysis methodologies in a straightforward interface. In all analyses, 500 bootstrap replicates were selected to assess robustness of the tree produced.

2.28 Comparative sequence analysis

2.28.1 *zPicture*

PIP plots were generated using zPicture as per the authors instructions (Ovcharenko *et al.,* 2004). Text files were generated containing relevant sequences in fasta format, and an annotation file was generated as per the authors instructions. The annotation file was also used to generate an underlay file as per the authors instructions. The base sequence (human unless otherwise specified) was masked for repeats using RepeatMasker. Most program parameters were as default, except sequences were searched on both strands, and chaining was employed. Chaining reports only those matches occurring in the same order in the different species, and avoids build-up of matches due to repetitive sequence occurring throughout the sequences. Chaining assumes that the order of matches should be conserved. The “high sensitivity” setting was also employed.