EXPRESSION AND FUNCTION OF GENES IDENTIFYING PLURIPOTENT CELL SUB-POPULATIONS IN THE EARLY MOUSE EMBRYO

A Thesis Submitted To The University Of

Adelaide For The Degree Of Doctor Of Philosophy

by

Tricia Ann Pelton, B Sc. (Hons.)

Department Of Biochemistry

University Of Adelaide

Adelaide, South Australia

July, 2000

TABLE OF CONTENTS

CHAPTER 2: MATERIALS AND METHODS

SUMMARY

The coordinated regulation of pluripotent cell developmental progression during early mouse development is critical for formation of the three primary germ layers of the embryo. Accumulating evidence suggested the existence of transient pluripotent cell sub-populations in the early mouse embryo, which could not be recognised at the molecular level. Marker genes generated by dd-PCR analysis of the ES to EPL cell transition *in vitro*, a model of the ICM to primitive ectoderm transition *in vivo* (Rathjen *et al.*, 1999; Lake *et al.*, 2000), potentially provided a means to distinguish these transient pluripotent cell sub-populations.

Wholemount *in situ* hybridisation analysis of the reported marker genes, *Oct4*, *Rex1* and *Fgf5* (Haub and Goldfarb, 1991; Hébert *et al.*, 1991; Rogers *et al.*, 1991; Rosner *et al.*, 1990, Schöler *et al.*, 1990a and b) in conjunction with markers identified and isolated from the dd-PCR screen of ES and EPL cell RNA, *L17*, *Psc1* and *K7* (Schulz, 1996; S. Sharma, unpublished) provided a molecular description of pluripotent cell developmental progression during peri-implantation mouse development. Expression of *Rex1* and *L17* was observed in the ICM and early epiblast populations of the developing blastocyst, and downregulated at approximately 4.75 d.p.c.. *Psc1* expression was detected in the pluripotent cells of the developing blastocyst and epiblast bud, and downregulated following proamniotic cavitation at 5.25 d.p.c.. *K7* was not expressed in the early blastocyst, but was expressed by pluripotent cells between 4.75 and 5.25 d.p.c., following the formation of the late stage blastocyst to immediately prior to primitive ectoderm organisation into a pseudostratified epithelial sheet. *Fgf5* expression was not detected during blastocyst development, but was detected in the primitive ectoderm following proamniotic cavitation at 5.25 d.p.c.. Compilation of gene expression data indicated the existence of at least four distinct pluripotent cell sub-populations within the $Oct4^{+}$ pool. Comparison of pluripotent cell sub-populations with embryonic development indicated that the transient sub-populations correlated with embryonic events, including increased pluripotent cell proliferation, proamniotic cavitation and formation of a primitive ectoderm monolayer. Furthermore, gradual progression of pluripotent cells *in vivo* correlated closely with marker gene expression *in vitro* during the ES to EPL cell transition, validating the conversion of ES cells to EPL cells as an *in vitro* model of the ICM to primitive ectoderm transition and allowing more precise definition of ES and EPL cell embryonic equivalents.

K7 was selected for further analysis, based on its tight expression window during early development. cDNA clones spanning a 6.6 kb *K7* sequence were isolated and the complete *K7* cDNA was shown to be 79.3% identical to a human gene of unknown function, termed KIAA0165. This human gene was expressed in similar tissues to *K7*, including the testis, haematopoietic system and gastrointestinal tract, suggesting that these two genes were homologues. The C-terminal putative protein sequence of K7 and KIAA0165 demonstrated similarity to three fungal proteins, *cut1* (*S. pombe*), *ESP1* (*S. cerevisiae*) and *bimB* (*A. nidulans*). Sequence similarity of the conserved C-terminal region of K7 with Cut1 and Esp1 C-terminal protein sequence implicated K7 in the control of cell division in mammals. A mitotic role for K7 was consistent with the *in vivo* expression pattern of *K7* in the adult, which was confined to tissues containing rapidly proliferating cells for tissue and cellular regeneration. Functional investigation of K7 in mammalian cells using overexpression studies demonstrated that like KIAA0165, K7 interacted with a known cell cycle regulator encoded by the *pituitary transforming tumour gene (PTTG)*. Furthermore, overexpression of mPTTG in Cos-1 cells increased the proportion of cells in the G2/M phase of the cell cycle, consistent with a role for this complex in cell cycle progression through G2/M. K7 was also shown to localise to the centrosome, the microtubule organising centre from which the mitotic spindle emanates. Taken together these results suggested that the K7/mPTTG interaction shared similarities with the Cut1/Cut2 and Esp1/Pds1 complexes of budding yeast and fission yeast respectively, suggesting that aspects of the chromosome segregation mechanism may be conserved between yeast and mammals.

STATEMENT

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give my consent for this thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED ..... DATE 11.7.00

ACKNOWLEDGMENTS

I am grateful to Professor Peter Rathjen for allowing me to undertake a Ph.D. in the Department of Biochemistry, for his supervision, interest in this project and the much valued contribution of critically reading this thesis.

I would also like to thank all members of the Rathjen lab, past and present, for their friendship and support throughout my years here. Particular mention must go to -; Drs. Tom Schulz and Shiwani Sharma for discovering the marker genes and thus providing me with a project, thanks also to Tom for passing on his knowledge of embryonic dissection and *in situ* hybridisation; Julie-anne Lake and Jenny Washington for extra-long coffee breaks and the bum-bag, but most of all for their continued friendship; past and present inhabitants of room 102 for company and innumerable laughs, particularly Gavin Chapman for introducing me to the wonderful world of molecular biology and his willingness to answer the call of Satan, Steve Kavanagh and Steve Rodda for a gravy booby time during their honours years, and Roger "Ratzenberger" Voyle for his vocabular intricacies; thanks also to Dr. Stephen Wood and Joy Rathjen for their comments on the introduction section of this thesis.

The assistance of Drs. Graham Webb and Peter Kolesik from the WAITE Institute in conducting the chromosome localisation analysis and assistance with confocal microscopy respectively, was much appreciated. Thanks also to all members of the Biochemistry Department for their willingness to share their knowledge and assist whenever necessary, it makes the department a great place to work. In particular, Tom Schulz, Steve Rodda, Stephen Wood and Roger Voyle for the odd RNA sample, Dr. Steve Dalton for introducing me to yeast genetics, Elaine Stead and Jo White for all their help with FACS analysis, and Dr. Terry Mulhern for analysing the protein structure data base searches (and for sharing his home-brew). Thanks must also go to the departmental support staff, particularly Amanda and Sharon for knowing where to locate Pete, Jan and Serge for the swiftness of orders and Jackie, Sheila and Judy for providing solutions, clean glassware and packed tips. Special thanks must also go to Sharon Pursglove, Anita Merkel and Ines Atmosukarto from the Booker lab for their friendship and chats over coffee.

I am forever indebted to my family for their love, support and interest, thanks Mum, Dad, Grant and Nicole. Thanks also to Jan for adopting me as one of her own.

Last but certainly not least I must thank Bryan Haines for his love and support, for being a sounding board for ideas and his constructive comments on this work, but most of all for his fun-loving nature which makes him a wonderful person to be around.

CHAPTER 1:

GENERAL INTRODUCTION

1.1 MAMMALIAN DEVELOPMENT

The developmental mechanisms involved in the formation of a complex living organism from a single founder cell, the fertilised egg, have been scientifically investigated for centuries in a variety of model organisms. These studies have not only increased our understanding of basic human embryology, but knowledge of the molecular basis of developmental decisions has proven to be applicable to genetic and developmental disease states and provided a basis for the generation of therapeutics.

Analysis of metazoan development has greatly benefited from the study of model organisms such as *Drosophila melanogaster* and *Xenopus laevis*. However, although many fundamental concepts derived from such studies also apply to mammalian development, the mammalian developmental regime differs from that of lower organisms during early pre- and peri-implantation periods. Early developmental events, such as cell specification and body axis determination, are programmed by maternally loaded cytoplasmic components in lower organisms (reviewed in Melton, 1991). In contrast, there is little evidence of maternal direction in early mammalian embryogenesis, and instead a dependence on zygotic transcription (reviewed in Kidder, 1992).

1.2 PLURIPOTENT CELLS AND MAMMALIAN EMBRYOGENESIS

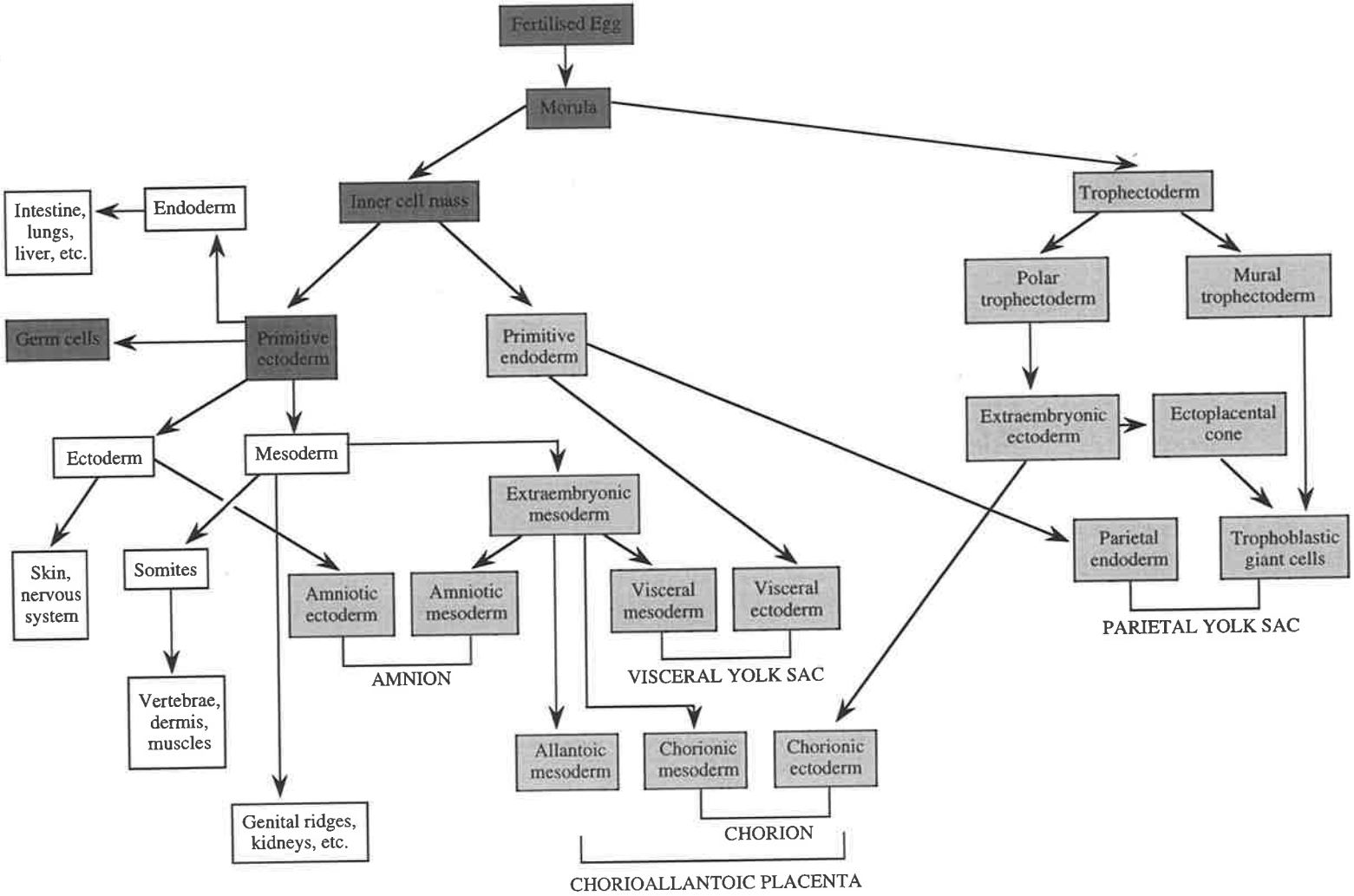
A distinctive and early feature of mammalian embryogenesis is the formation and maintenance of a population of pluripotent cells. They retain the ability to give rise to all three germ layers of the embryo proper during the process of gastrulation and thus can be defined as pluripotent (Martin, 1980). Pluripotent cells undergo progressive and coordinated differentiation in response to signals encountered during early development, and become developmentally restricted in their potential to form extra-embryonic lineages. A summary of the commitment of pluripotent cells to various embryonic and extra-embryonic lineages is depicted in Figure 1.1.

While appropriate regulation of pluripotent cells is a feature of normal mammalian development, the mammalian embryo appears to retain considerable developmental flexibility. Removal of up to 80% of the embryo prior to gastrulation by mitomycin C treatment (Snow and Tam, 1979; Tam, 1988), or aggregation of several pre-implantation embryos (Buehr and McLaren, 1974; Lewis and Rossant, 1982; Rands, 1986) can give rise to normal embryos, and

Figure 1.1: Summary of the cell lineages formed during mouse embryogenesis

Darkly shaded boxes represent totipotent and pluripotent cells which give rise to all extra-embryonic tissues and the embryo proper, and also allow the continuation of the species through the germ cell lineage. Lightly shaded boxes represent extra-embryonic lineages and unshaded boxes designate the three primary germ layers and their derivatives.

Adapted from Hogan *et al.* (1994)


Fertilised Egg
Morula
Inner cell mass
Trophectoderm
Intestine, lungs, liver, etc.
Endoderm
Polar trophectoderm
Mural trophectoderm
Germ cells
Primitive ectoderm
Primitive endoderm
Extraembryonic ectoderm
Ectoplacental cone
Ectoderm
Mesoderm
Extraembryonic mesoderm
Parietal endoderm
Trophoblastic giant cells
Skin, nervous system
Somites
Amniotic ectoderm
Amniotic mesoderm
Visceral mesoderm
Visceral ectoderm
PARIETAL YOLK SAC
AMNION
VISCERAL YOLK SAC
Vertebrae, dermis, muscles
Allantoic mesoderm
Chorionic mesoderm
Chorionic ectoderm
Genital ridges, kidneys, etc.
CHORION
CHORIOALLANTOIC PLACENTA

offspring of typical size. This developmental plasticity is thought to reside in the ability of pluripotent cells to reprogram their development in response to environmental signals (Smith, 1992). Early mammalian embryogenesis therefore appears dependent on a flexible signaling system which can modulate the behaviour of pluripotent cells in response to embryonic and environmental signals (Smith, 1992). Consequently, deciphering the signals which regulate pluripotent cellular decisions is essential for a comprehensive understanding of mammalian embryogenesis.

1.3 MOUSE DEVELOPMENT

The most effective model system for deciphering the molecular controls of mammalian development is the mouse (*Mus musculus*). The mouse is small, has a relatively short gestation time, produces relatively large litters, and is easily kept in breeding colonies. The capacity to inbreed mice has led to the generation of over 400 inbred strains for research (Festing, 1992). Detailed documentation of the morphological changes which occur during mouse development has been carried out using histological and microscopic techniques (Kaufman, 1992). The genome of the mouse is well characterised with over ten thousand genes having been mapped. Finally, the ability to isolate embryonic stem (ES) cells from mouse embryos and maintain them *in vitro* (Evans and Kaufman, 1981; Martin, 1981) has enabled precise genetic modifications to be made in ES cells. Since genetically modified ES cells have the capacity to contribute to all tissues of the embryo proper upon reintroduction into host embryos, analysis of gene function can be carried out in a developmental context.

1.3.1 Pre-implantation Development

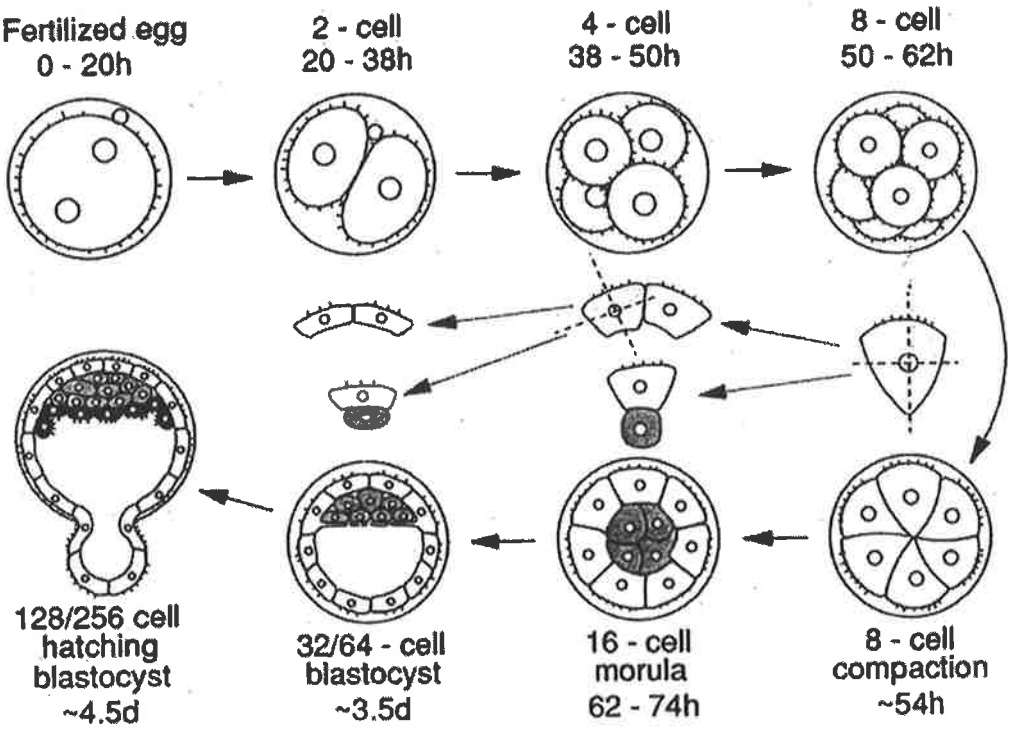
1.3.1.1 Fertilisation and first cleavage divisions

Just prior to ovulation, mouse oocytes complete the first meiotic division with half of their chromosomes being expelled into the first polar body. Following ovulation into the oviduct, fusion of the oocyte and sperm gametes (fertilisation) can occur, instigating the initial events of mouse embryogenesis and completion of meiosis within the oocyte (Figure 1.2). The second polar body is extruded, and the nuclear membrane envelopes the maternal and paternal chromosomes. The first cell cycle is completed 17-20 hours post-fertilisation, culminating with

Figure 1.2: Schematic representation of mouse pre-implantation development

Following fertilisation, the embryonic cells undergo successive cleavages during early development. The embryonic stage and time following fertilisation is denoted. In the centre, the conservative and differentiative cleavage planes (dotted line) of the blastomeres are indicated, which following compaction and outer cell polarisation, lead to differentiation of trophectoderm and the formation of the pluripotent inner cell mass (lightly shaded cells). Following trophectoderm maturation and the subsequent formation of the blastocoelic cavity the pluripotent cells lining the blastocoel differentiate into primitive endoderm (darkly shaded cells), prior to hatching of the embryo from the zona pellucida.

Reproduced from Fleming *et al.* (1992)


Fertilized egg
0 - 20h
2 - cell
20 - 38h
4 - cell
38 - 50h
8 - cell
50 - 62h
128/256 cell
hatching
blastocyst
~4.5d
32/64 - cell
blastocyst
~3.5d
16 - cell
morula
62 - 74h
8 - cell
compaction
~54h

cleavage of the totipotent founder cell into two embryonic daughter cells, termed blastomeres. Following this first cleavage division, rapid degradation of maternally loaded mRNA and protein occurs and the embryonic genome becomes activated, with zygotic transcription detectable by the mid-late 2 cell stage (Taylor and Pilkó, 1987; Telford *et al.*, 1990; Latham *et al.*, 1991; Kidder, 1992). An important consequence of this process is the transfer of developmental control from the oocyte to the embryo. Nonetheless, oocyte derived gene products may play some role in pre-implantation development in the mouse since some maternal protein products persist past the 2 cell stage (Taylor and Pilkó, 1987; Kidder, 1992).

The blastomeres continue to divide for a further 2 divisions, approximately every 20 hours, resulting in an embryo consisting of 8 loosely associated blastomeres contained within the glycoprotein coat of the zona pellucida. Individual blastomeres are totipotent as they are able to give rise to all extra-embryonic and embryonic lineages of the conceptus (Gardner, 1983).

1.3.1.2 Compaction and cell polarisation

At the 8 cell stage the embryo undergoes compaction, resulting in a reduction in the intercellular space between blastomeres, maximisation of cell-cell contacts and establishment of junctional complexes (Fleming *et al.*, 1992). Compaction occurs largely in response to post-translational modifications to gene products such as E-cadherin, synthesised as early as the 2 cell embryo, rather than from transcription/translation processes at the 8 cell stage (Fleming *et al.*, 1992; Kidder, 1992). Gap junctions form between cells and allow cell-cell communication, which is critical for the maintenance of compaction (Lee *et al.*, 1987; Becker *et al.*, 1992). E-cadherin has been shown to be critical for compaction as targeted disruption of this gene results in embryonic lethality attributed to decompaction of morulae following degradation of the maternal store of E-cadherin (Riethmacher *et al.*, 1995).

Following compaction, outer blastomeres become polarised, creating distinct apical and basolateral domains in outer cells. A redistribution of cytoskeletal components, microvilli, and endosomes is seen in apical regions (Fleming and Pickering, 1985; Fleming *et al.*, 1992), whilst a basolateral redistribution of the nucleus and E-cadherin is observed (Vestweber *et al.*, 1987). Asymmetric cell division during the next cleavage to the 16 cell morula results in formation of distinct outer and inner cells (Fleming *et al.*, 1992) (Figure 1.2). The small, apolar inner cells are

destined to remain pluripotent and form the inner cell mass (ICM), whereas the outer cells, which contain an outer surface free of contact with other cells, are destined to differentiate into an epithelial cell layer known as trophectoderm, an extra-embryonic precursor of the chorion and placenta (Figure 1.1).

1.3.1.3 Blastocyst formation

Differentiation of trophectoderm is characterised by the formation of mature cell adhesion and junction systems, basement membrane formation, cytoplasmic reorganisation, and the establishment of cell membrane polarity (Fleming *et al.,* 1992). Concurrent with trophectoderm differentiation, vectorial fluid transport occurs. Sodium ion influx at the apical membrane takes place via several transporters, including a Na^+/H^+ exchanger and a Na^+ ion channel (Manejwala *et al.,* 1989), while sodium ion efflux across the basal membrane is carried out by a Na^+,K^+-ATPase pump, located on basolateral cell membrane of the trophectoderm (Wiley, 1987). This creates a transcellular Na^+ gradient that results in passive water flow to counter the apical to basal sodium ion movement, and the creation of the blastocoelic cavity.

Blastocoelic cavity formation displaces the small mass of non-polarised pluripotent ICM cells to one end of the blastocoel. At approximately 4.0 days *post coitum* (d.p.c.) the blastocyst is approximately the same size as the fertilised egg and during expansion of the blastocoel, the second differentiation event of embryogenesis occurs. A second extra-embryonic epithelial layer differentiates from the ICM cells which line the blastocoelic cavity (Figure 1.2). They are termed primitive endoderm, a progenitor of extra-embryonic parietal and visceral yolk sac endoderm (Figure 1.1). The remaining core of pluripotent cells are now termed either epiblast, primitive ectoderm or embryonic ectoderm, terms which can be used interchangeably to describe the pluripotent cell population present following primitive endoderm differentiation, until they undergo gastrulation, at day 6.5 of embryogenesis.

1.3.2 Peri-implantation Development

1.3.2.1 Blastocyst implantation

Implantation of the blastocyst into the uterine wall occurs around day 4.5 of embryogenesis. Immediately prior to implantation, trophectoderm-derived proteases degrade the

protective zona pellucida coat and the blastocyst hatches from the zona pellucida following expansion and a series of rhythmical contractions (Figure 1.2). The epiblast of hatched embryo flattens and forms the characteristic morphology of the fully expanded blastocyst (Figure 1.3). This embryonic structure now contains three discrete tissues; an outer layer of trophectoderm, the primitive endoderm and pluripotent epiblast situated between the two.

Around this time the uterine wall secretes a number of steroid hormones, including oestrogen and progesterone, and cytokines such as leukaemia inhibitory factor (LIF) (Bhatt *et al.*, 1992), which are required for implantation of the blastocyst into the uterine wall. The requirement for oestrogen in implantation has been demonstrated by ovarectomy at 3.5 d.p.c., which results in dormant blastocysts which are unable to implant within that female. Subsequent treatment with oestrogen re-initiates implantation (Mantalenakis and Ketchel, 1966; Yoshinaga and Adams, 1966). LIF has been demonstrated to increase the *in vitro* activity of extracellular matrix (ECM) metalloproteinases, allowing ECM breakdown and initiation of mural trophectoderm invasion into the uterine wall (Harvey *et al.*, 1995). This suggests that a reduction in the activity of ECM-degrading proteinases in blastocysts may account for the inability of blastocysts to implant into the uterine wall of LIF deficient female mice (Stewart *et al.*, 1992).

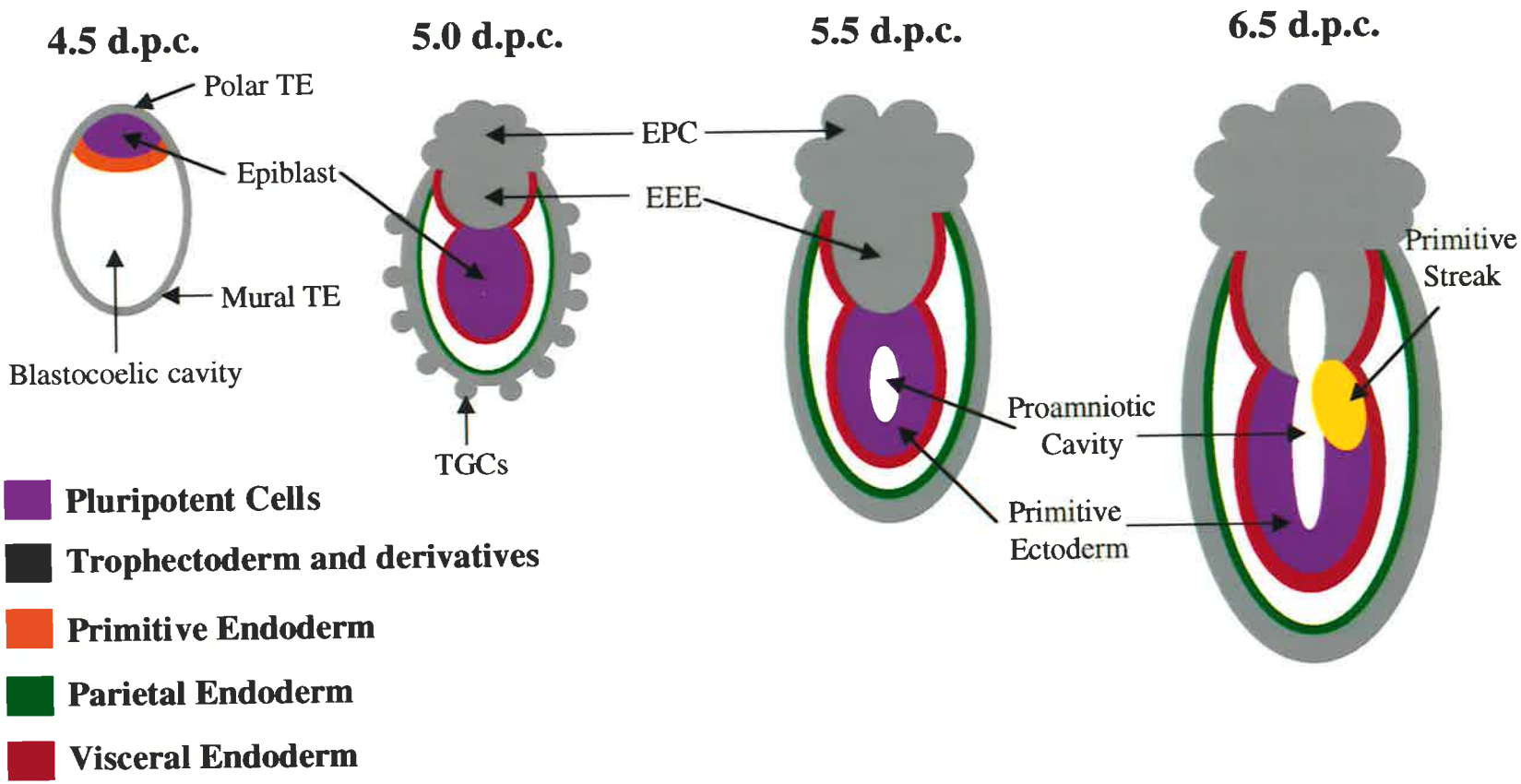
Attachment of the embryo to the uterine wall occurs via the mural trophectoderm, and invasion of these cells into the receptive uterine wall induces a maternal decidual response which consolidates implantation. The decidual reaction, a tissue specific response of the uterus to the embryo, involves stimulation of uterine stroma to deposit a spongy cell mass around the embryo and increased capillarisation.

1.3.2.2 Cellular changes in the embryo

Concomitant with implantation, the mouse embryo initiates a complex developmental program which results in transformation of the solid mass of pluripotent cells into the egg cylinder structure, in readiness for gastrulation (Figure 1.3). The cells enter a poorly defined transition as embryogenesis switches from cleavage based development, in which there is no overall increase in embryo size, to a proliferative based system, which sees coupling of cell growth and division.

Figure 1.3: Diagrammatic representation of post-implantation mouse development from 4.5 to 6.5 d.p.c.

At 4.5 d.p.c. the embryo consists of three distinct cell lineages, the trophectoderm (TE), primitive endoderm and pluripotent epiblast. The trophectoderm lineage can be sub-divided further into two lineages; the polar trophectoderm, which overlays the pluripotent cells, and the mural trophectoderm which abuts the blastocoelic cavity. At 5.0 d.p.c. the embryo consists of a solid mass of epiblast cells (epiblast bud) overlayed by the polar trophectoderm derived extra-embryonic ectoderm (EEE) and ectoplacental cone (EPC), and surrounded by trophoblast giant cells (TGCs) which have differentiated from the mural trophectoderm. The primitive endoderm differentiates into visceral endoderm, directly adjacent to the epiblast, and parietal endoderm, which lines the inner surface of the mural trophectoderm. By 5.5 d.p.c. the proamniotic cavity has formed, and the pluripotent cells have become polarised and organise into a pseudostratified epithelial sheet of primitive ectoderm surrounding the proamniotic cavity, in readiness for the onset of gastrulation at approximately 6.5 d.p.c.


4.5 d.p.c.
5.0 d.p.c.
5.5 d.p.c.
6.5 d.p.c.
Polar TE
Epiblast
Mural TE
Blastocoelic cavity
EPC
EEE
TGCs
Proamniotic Cavity
Primitive Ectoderm
Primitive Streak
Pluripotent Cells
Trophectoderm and derivatives
Primitive Endoderm
Parietal Endoderm
Visceral Endoderm

(i) Cellular changes within the trophectoderm:

Extra-embryonic tissues do not remain as simple epithelia. Cells of the trophectoderm lineage undergo differentiation, proliferative changes and migration. Mural trophectoderm, the trophectodermal cells which line the blastocoelic cavity (Figure 1.3), cease cell division but continue DNA replication to form primary trophoblast giant cells (TGCs), containing polytene chromosomes. These cells play an important role in induction of the decidual reaction during embryo attachment (Gardner and Johnson, 1972).

In contrast, trophectoderm overlying the pluripotent cells, termed polar trophectoderm (Figure 1.3), remains diploid and begin to proliferate rapidly (Gardner and Papaioannou, 1975). Contact with the epiblast is critical for the onset and continued proliferation of the polar trophectoderm as loss of contact between pluripotent cells and cells of the polar trophectoderm suspends proliferation and results in the formation of TGCs (Gardner and Papaioannou, 1975; Rossant and Ofer, 1977). Furthermore, mural trophectoderm can substitute for polar trophectoderm if epiblast tissue is placed in direct contact with the mural trophectoderm prior to differentiation into TGCs (Gardner *et al.,* 1973).

Following implantation some cells of the polar trophectoderm migrate around the embryo, differentiating into secondary TGCs to replace the primary TGC layer. Polar trophectoderm cells which remain in close proximity with the pluripotent cells develop into extra-embryonic ectoderm, an extra-embryonic cell lineage which gives rise to the ectoplacental cone and eventually tissues of the chorion (Gardner *et al.*, 1973; Gardner and Papaioannou, 1975; Rossant and Ofer, 1977).

(ii) Primitive endoderm differentiation:

Following implantation, cells of the primitive endoderm undergo differentiation into two extra-embryonic cell types, visceral and parietal endoderm (Figure 1.3). Intimate contact of endodermal cells with the diploid pluripotent cells appears to be a prerequisite for primitive endoderm differentiation along the visceral endoderm pathway (Gardner, 1983). The visceral endoderm forms a continuous, columnar epithelial layer around the pluripotent cells, and following subsequent morphogenesis, envelops the extra-embryonic ectoderm, and eventually contributes to the visceral yolk sac (Gardner, 1983). The developmental role of visceral

endoderm was traditionally thought to involve supplying nutrition to the embryo, as it displays a morphology appropriate for an absorptive role (Gardner, 1983). However, more recently an additional role for visceral endoderm as a source of instructive molecules which can regulate pluripotent cell development has emerged. Secretion of factor(s) from a visceral endoderm-like cell line, END-2, can induce differentiation of aggregated P19 EC cells, a primitive ectoderm-like cell line, into endoderm and cardiac muscle cells (Mummery *et al.,* 1991; van den Eijnden-van Raaij *et al.,* 1991). Visceral endoderm has also been proposed to secrete a diffusible "death" signal which induces apoptosis of pluripotent cells (Coucouvanis and Martin, 1995). As described in section 1.4.3.4, targeted disruption of visceral endoderm-specific genes can result in pluripotent cell abnormalities (Chen *et al.,* 1994; Spyropoulos and Cappechi, 1994). Finally, the anterior visceral endoderm has been implicated in imparting anterior identity to the primitive ectoderm prior to gastrulation (reviewed by Beddington and Robertson, 1998).

Differentiation of primitive endoderm into parietal endoderm has been proposed to be a result of the surface area of the 4.5 d.p.c. epiblast not increasing in area at a rate proportional with the increased number of primitive endoderm cells. Therefore, following mitotic divisions of the primitive endoderm, one daughter cell remains in contact with the epiblast cells, whilst the other is displaced into a second, superficial layer of endoderm, not in direct contact with the pluripotent cells (Gardner, 1983). Displaced primitive endoderm cells which do not remain in contact with the epiblast differentiate into flattened fibroblast-like cells which migrate around the mural trophectoderm and are termed parietal endoderm (Gardner, 1983). The parietal endoderm produces a thick basement membrane termed Reichert's membrane which constitutes the major barrier between the maternal and embryonic environment, and eventually contributes to the parietal yolk sac (Gardner, 1983).

(iii) Formation of the egg cylinder:

The pluripotent epiblast cells account for the remaining cells of the mature blastocyst at 4.5 d.p.c. and compose the smallest cell population, consisting of approximately 20-25 cells (Snow, 1976). However, between implantation and the onset of gastrulation at 6.5 d.p.c., these cells proliferate more rapidly as cell cycle times decrease to approximately 8 hours and cells increase in number to 120 at 5.5 d.p.c. and 660 by 6.5 d.p.c. (Snow, 1977). A projection of proliferating

extra-embryonic ectoderm displaces the epiblast cells into the blastocoel, such that by 5.0 d.p.c. the embryo consists of an apolar mass of cells, referred to in this thesis as the "epiblast bud", surrounded by visceral endoderm and overlayed by extra-embryonic ectoderm (Figure 1.4).

Formation of a proamniotic cavity within the epiblast bud is initiated at approximately day 5.0 of embryogenesis. A two step, signal-mediated mechanism has been proposed to explain the formation of the cavity (Coucouvanis and Martin, 1995) (Figure 1.4). A visceral endoderm-derived "death" signal diffuses across the epiblast cells, inducing apoptosis. Concomitantly, a second signal, localised to the basement membrane between the epiblast and visceral endoderm, mediates survival of epiblast cells that it contacts, resulting in formation of a monolayer of pluripotent cells.

Following the commencement of proamniotic cavitation epiblast cells redistribute cytokeratins sub-apically to become polarised (Jackson *et al.,* 1981), and reorganise into a pseudostratified epithelial sheet of pluripotent cells. The post-cavitation pluripotent cells appear to have an alternate gene expression to that of the late ICM/early epiblast (see section 1.4.2.1; Rogers *et al.*, 1991; Haub and Goldfarb, 1991; Hébert *et al.*, 1991). Furthermore, post-cavitation pluripotent cells fail to contribute to chimeras when introduced into host blastocysts, suggesting that alterations in developmental potential may exist between pre- and post-proamniotic cavitated pluripotent cells (Rossant, 1977; Beddington, 1983). Therefore, although the terms epiblast, primitive ectoderm or embryonic ectoderm are used interchangeably to describe pluripotent cells from primitive endoderm differentiation until the onset of gastrulation, following proamniotic cavitation the pluripotent cells appear different with respect to gene expression and developmental potential compared with early epiblast cell population present at day 4.5 of embryogenesis. Therefore, for the purpose of this thesis, the pluripotent cell population present from primitive endoderm differentiation until proamniotic cavitation will still be termed epiblast, but following proamniotic cavitation up until gastrulation the pluripotent cell population will be termed primitive ectoderm.

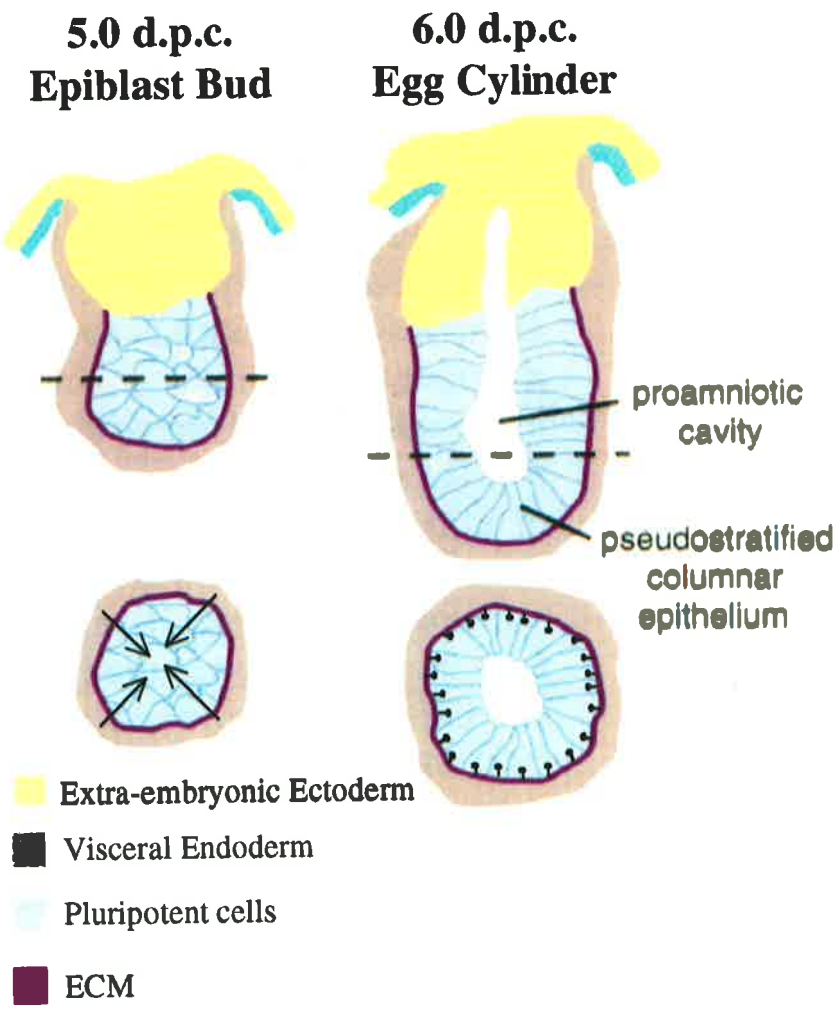
1.3.3 Gastrulation and Axis Patterning

The primitive ectoderm is the substrate for the inductive signals of gastrulation, and through this process, gives rise to all the embryonic and remaining extra-embryonic tissues of the embryo

Figure 1.4: Schematic illustration of the process of proamniotic cavitation

The upper diagrams depict sagittal sections through 5.0 and 6.0 d.p.c. embryos, whilst the lower embryos represent transverse section through the planes indicated by the dotted lines. The proamniotic cavity has been proposed to form following expression of a visceral endoderm derived death signal (arrows), which diffuses across the pluripotent cells and induces apoptosis of inner cells. A second ECM localised signal ( ) mediates survival of pluripotent cells in contact with the basement membrane.

Adapted from Coucouvanis and Martin (1995)


5.0 d.p.c.
Epiblast Bud
6.0 d.p.c.
Egg Cylinder
proamniotic
cavity
pseudostratified
columnar
epithelium
Extra-embryonic Ectoderm
Visceral Endoderm
Pluripotent cells
ECM

proper. Gastrulation commences around 6.5 d.p.c. with the formation of the primitive streak, an apparent cellular thickening on the proximal extra-embryonic/embryonic junction at the posterior side of the embryo. During the next 24 hours the pluripotent cell monolayer is transformed into a multi-layered embryo, consisting of the three primary germ layers; mesoderm, endoderm and ectoderm, arranged according to the basic foetal body plan (Figure 1.5).

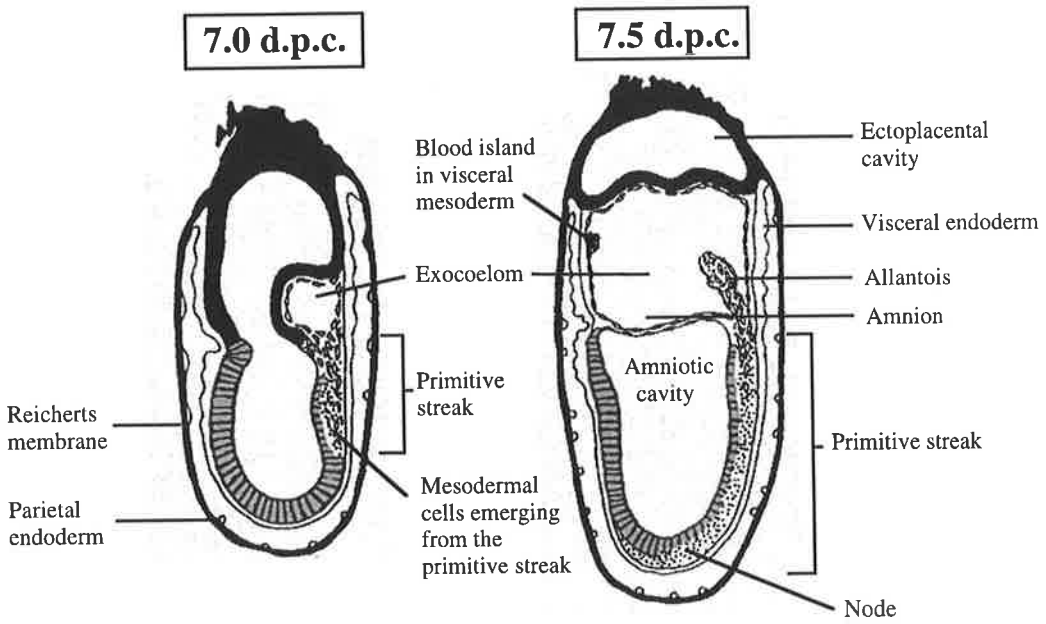
Mesoderm formation occurs within the primitive streak, which is first observed at the posterior aspect of the embryo. However, asymmetry within the embryo appears to be established prior to gastrulation and include the embryonic/abembryonic axis, a result of the blastocoelic cavity displacing the pluripotent ICM to one end of embryo, and the animal/vegetal axis, which is located perpendicular to the embryonic/abembryonic axis. The animal/vegetal axis is established following fertilisation with the extrusion of the second polar body, which remains tethered to a specific blastomere, marking the animal pole of the embryo (Gardner, 1997). However, the proposal that specification of embryonic axes are influenced by determinants present within the pre-implantation embryo remains controversial. Although it can not be ruled out that pattern information present within the early conceptus may influence embryonic axis patterning, evidence of extensive cellular mixing of the epiblast cells prior to gastrulation (Gardner and Cockroft, 1998) suggests that this is not likely to be the case (reviewed in Beddington and Robertson, 1999). A more likely candidate in the establishment of embryonic asymmetry prior to gastrulation is the anterior visceral endoderm, which has recently been implicated in providing instructive cues for anterior specification in the mouse, prior to overt signs of primitive streak formation (reviewed in Beddington and Robertson, 1998).

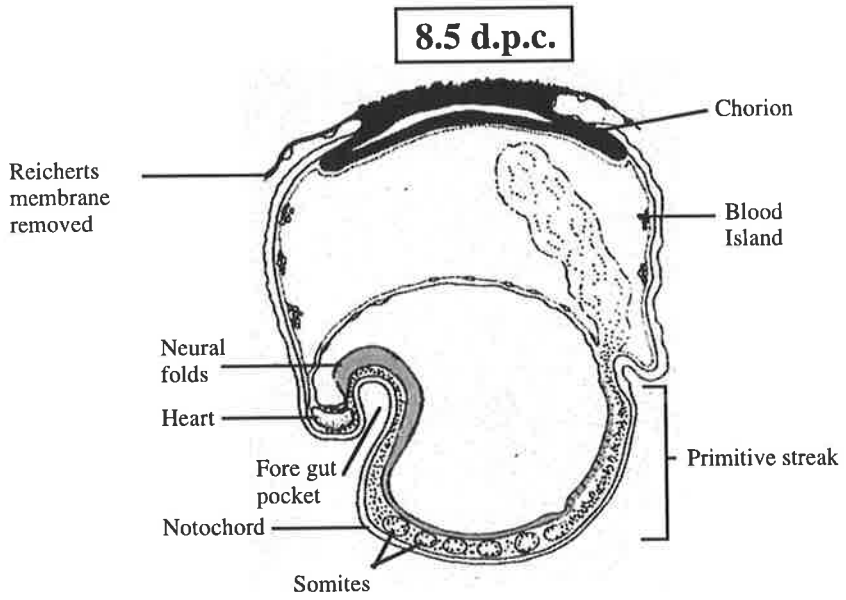
As the primitive streak forms, primitive ectoderm cells assume an accelerated proliferation rate with the average cell cycle time decreasing to 6 hours (Snow, 1977). The epithelial continuity of the primitive ectoderm is lost in the streak region as streak cells undergo an epithelial to mesenchymal cellular transformation. Primitive ectoderm cells delaminate, move toward and accumulate within the streak, which extends distally as gastrulation proceeds. Cells which pass through the streak differentiate into mesoderm and this cell layer emerges between the primitive ectoderm and visceral endoderm. Migration of newly formed mesoderm is dependent on FGF receptor 1 signaling (Ciruna *et al.*, 1997). Early in gastrulation mesoderm migrates proximally, crossing the extra-embryonic ectoderm/primitive ectoderm boundary and forms extra-embryonic

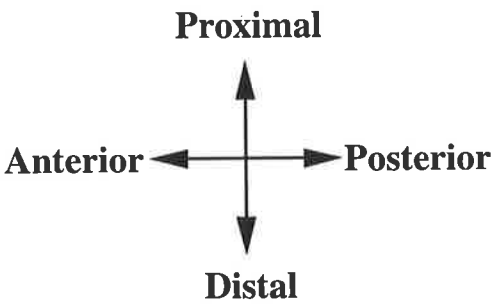
Figure 1.5: Gastrulation in the mouse embryo

Between approximately 6.5 and 7.0 d.p.c. the process of gastrulation is initiated with the formation of the primitive streak in the proximal posterior region of the primitive ectoderm. By 7.5 d.p.c. the primitive streak has progressed distally, with the node forming in the most distal region of the primitive streak. The mesoderm migrates proximally, anteriorly and laterally around the embryo to generate a distinct layer between the primitive ectoderm and visceral endoderm. By 8.5 d.p.c. numerous tissues can be observed along the anterior-posterior axis, including ectoderm derived neural tissue, mesoderm derived notochord, somites and heart, as well as the definitive gut endoderm.

Reproduced from Hogan *et al.* (1986)


7.0 d.p.c.
7.5 d.p.c.
Blood island in visceral mesoderm
Exocoelom
Primitive streak
Reicherts membrane
Parietal endoderm
Mesodermal cells emerging from the primitive streak
Ectoplacental cavity
Visceral endoderm
Allantois
Amnion
Amniotic cavity
Primitive streak
Node


8.5 d.p.c.
Chorion
Reicherts membrane removed
Blood Island
Neural folds
Heart
Fore gut pocket
Notochord
Somites
Primitive streak


Proximal
Anterior
Posterior
Distal

mesoderm. Later mesoderm moves laterally around the embryo, towards the anterior region and forms embryonic mesoderm. Mesoderm cells that emanate from different regions of the streak are fated to form different mesodermal tissue. Extra-embryonic, lateral and paraxial mesoderm emerge from progressively more distal regions of the streak, with axial mesoderm emanating from the node. The node is located at the most anterior aspect of the primitive streak, at the distal tip of the egg cylinder, and is fated to become the headprocess and notochord (Tam and Beddington, 1992). Mesoderm cell types form the precursors of vertebrae, ribs, dermis, kidney, and muscle of the developing embryo (Tam and Beddington, 1992). Definitive endoderm is formed at 7.5 d.p.c. by recruitment of primitive ectoderm cells into the node. They emerge from the node and move anteriorly, incorporating into the visceral endoderm and eventually displacing it to extra-embryonic regions (Tam and Beddington, 1992). Definitive endoderm provides precursors for differentiated tissues including the gut, lung, liver and pancreas (Tam and Beddington, 1992). Primitive ectoderm cells which do not ingress through the streak are the precursors of neurectoderm and epidermis, which give rise to the nervous system and skin respectively (Tam, 1989; Quinlan *et al.,* 1995).

Extensive lineage tracing and transplantation experiments have indicated that primitive ectoderm cells are regionalised with respect to their fate to form differentiated lineages, and have allowed construction of a fate map which is illustrated in Figure 1.6 (Beddington, 1982; Tam and Beddington, 1987; Tam, 1989; Lawson *et al.,* 1991; Tam and Beddington, 1992; Quinlan *et al.,* 1995; Parameswaran and Tam, 1995).

The process of gastrulation is pivotal for establishing the anterior-posterior axis of the embryo and in cell specification within primary germ layers. It depletes the embryo of pluripotent cells, except for a population of primordial germ cells (PGCs) which are partitioned from proximal posterior primitive ectoderm cells. PGCs are first detectable at 7.0 d.p.c. in the extra-embryonic mesoderm at the base of the allantois as approximately 75 rounded cells which express alkaline phosphatase (Ginsburg *et al.,* 1990). Subsequently, they leave the allantois and migrate through the gut endoderm to the genital ridge by 11-11.5 d.p.c. They proliferate during this migration and by day 13 of development approximately 25,000 PGCs colonise the gonad primordia (Tam and Snow, 1981). Upon entering the ovary, female PGCs cease cell division in

Figure 1.6: Fate maps of mouse primitive ectoderm

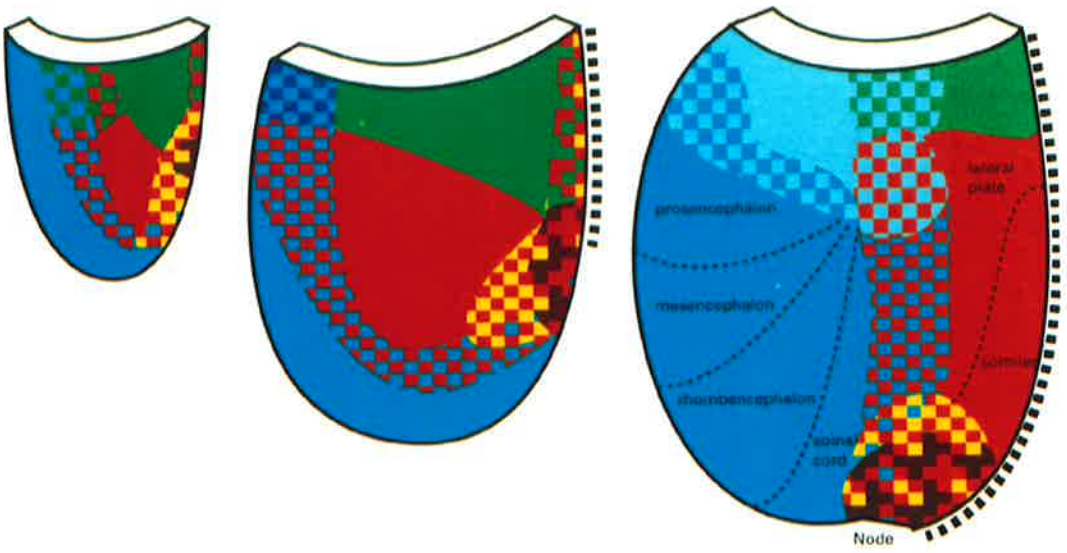
Diagrammatic representation of regionalisation within the primitive ectoderm prior to and during gastrulation with respect to the fate of different regions to form embryonic tissues.

Reproduced from Hogan *et al.* (1994)

PRESTREAK
6.0 d.p.c.

EARLY STREAK
7.0 d.p.c.

LATE STREAK
7.5 d.p.c.



Ectoderm

Mesoderm

Head process/Notochord

Endoderm

Extra-embryonic Mesoderm

Amnion Ectoderm

Surface Ectoderm

Primitive Streak

prophase of the first meiotic division. Male PGCs continue to proliferate in the developing testes until mitotic arrest around day 14 of embryogenesis (Monk and McLaren, 1981; Schultz, 1986).

1.4 EVIDENCE FOR TRANSIENT PLURIPOTENT CELL SUB-POPULATIONS IN THE EARLY MOUSE EMBRYO

Pluripotent cell populations have been broadly classified based on morphology and developmental behaviour (Figure 1.7). Pluripotent cells are found within the ICM lineage which develops from the inner cells of the morula, and in a subsequent population which forms from the ICM, termed either epiblast, primitive ectoderm or embryonic ectoderm. During gastrulation, the epiblast cells lose pluripotency as they differentiate into the three primary germ layers. However, pluripotency is maintained within the PGC lineage, which give rise to subsequent generations, allowing continuation of the species.

The use of complicated and interchangeable nomenclature for the epiblast/primitive ectoderm/embryonic ectoderm cell population has meant that the precise identity with respect to the starting or end points of this pluripotent cell population is not clear. Furthermore, while this standard categorisation of pluripotent cells is consistent with microscopic observations of embryogenesis, several lines of evidence suggest that this classification does not provide a comprehensive description of the pluripotent cell populations present during early mouse development. These are outlined below.

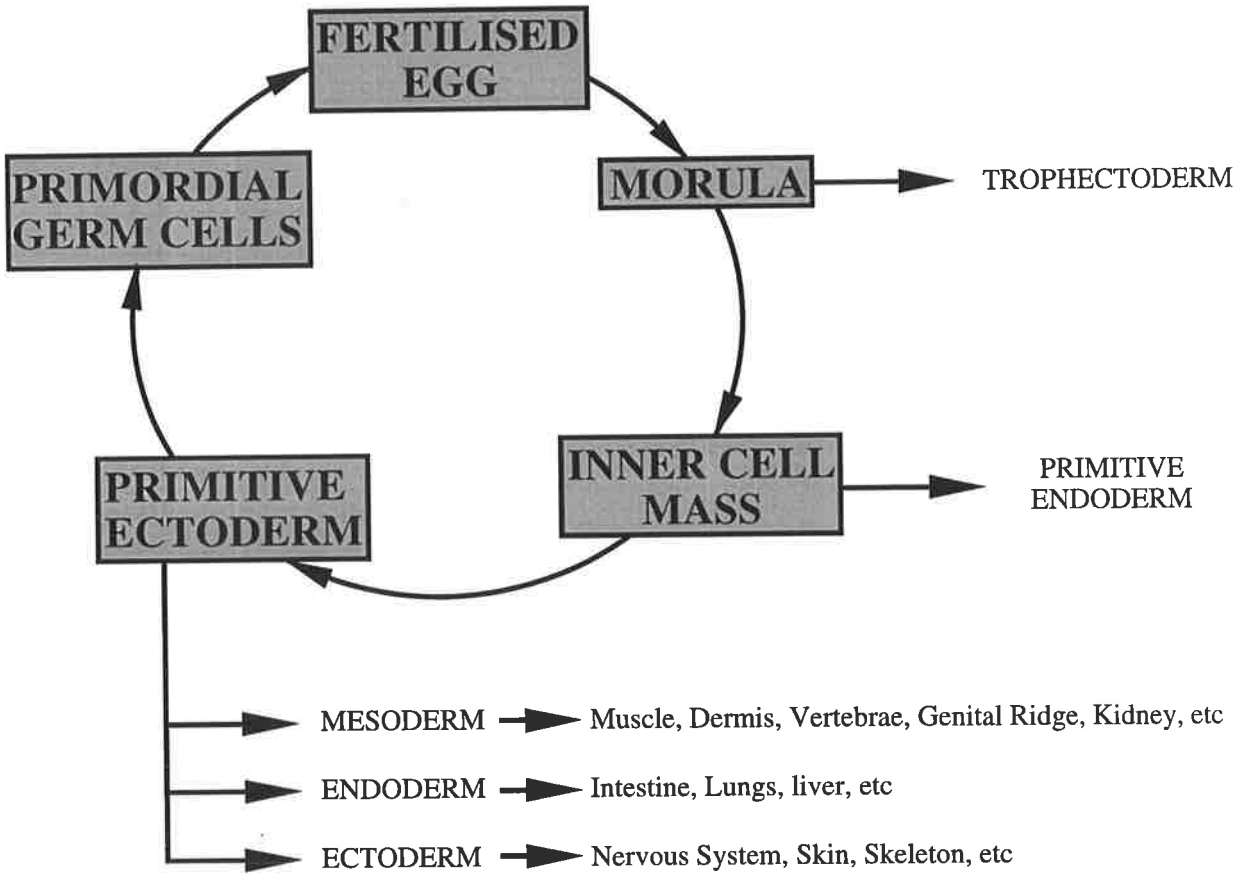
1.4.1 Developmental Potential of Pluripotent Cells

1.4.1.1 Loss of totipotence

Even following compaction and allocation into either ICM or trophectoderm lineages, cells of the morula retain developmental plasticity as commitment to a particular lineage can be altered. Individual blastomeres isolated from the outer portion of the 8 cell morula predominantly form trophectodermal vesicles in culture (Tarkowski and Wroblewska, 1967), indicating commitment to the extra-embryonic lineage. However, the interpretation that individual blastomeres are committed at this point is not consistent with the fact that isolated external or internal cells of the 16 cell morula cultured *in vitro* give rise to both trophectoderm and ICM progeny (Johnson and Ziomek, 1983). Furthermore, aggregation of exclusively external or internal cells and

Figure 1.7: Summary of the maintenance and differentiation of totipotent and pluripotent cell lineages of mouse development

Development begins with the totipotent fertilised egg which, following successive cleavage divisions, gives rise to the totipotent blastomeres of the morula. The pluripotent ICM lineage is formed following compaction and polarisation within the morula, and this population is maintained and expanded during primitive ectoderm formation. Following gastrulation the pluripotent cell lineage is retained within the primordial germ cell lineage. A number of the differentiated tissues which arise from each pluripotent cell stage are also indicated.


FERTILISED EGG
MORULA
TROPHECTODERM
INNER CELL MASS
PRIMITIVE ENDODERM
PRIMITIVE ECTODERM
PRIMORDIAL GERM CELLS
MESODERM
Muscle, Dermis, Vertebrae, Genital Ridge, Kidney, etc
ENDODERM
Intestine, Lungs, liver, etc
ECTODERM
Nervous System, Skin, Skeleton, etc

transplantation *in vivo* results in the formation of normal embryos and viable mice (Ziomek *et al.*, 1982). Moreover, outer cells from late morulae, destined to become trophectoderm, are capable of contributing to both the ICM and trophectoderm derivatives upon isolation and reinjected into the zona pellucida of early morulae (Rossant and Vijh, 1980). Inner cells show a similar totipotency at later stages of development. Whole ICMs isolated by microsurgery or immunosurgery can form TGCs in culture and induce a decidual response *in vivo*, indicating the ability to reform a functional layer of trophectoderm (Rossant and Tamura-Lis, 1979; Nichols and Gardner, 1984; Chisholm *et al.*, 1985), even though this is not the normal fate of the ICM at 3.5 d.p.c. (Rossant, 1975). These studies indicate that loss of totipotence is not concomitant with allocation to extra-embryonic or ICM lineages, or even necessarily with the formation of these cell types.

1.4.1.2 Potency of ICM to form primitive endoderm

Although primitive endoderm normally forms from the ICM at approximately 4.0 d.p.c, the precise time at which the pluripotent cells lose the capacity to form primitive endoderm has not been resolved. Immunosurgically isolated ICMs from fully expanded blastocysts tend to differentiate into primitive endoderm (Hogan and Tilly, 1977; Dziadek, 1979; Pedersen *et al.*, 1977; Rossant and Tamura-Lis 1979; Nichols and Gardner, 1984; Chisholm *et al.*, 1985). However, a somewhat unexpected finding was the ability of individual ICMs to form both trophectoderm and parietal endoderm, even though the appearance of these cell types is temporally distinct *in vivo* (Nichols and Gardner, 1984). Primitive endoderm can also differentiate *in vitro* from teratocarcinomas produced by ectopic injections of 6.0-7.0 d.p.c. primitive ectoderm (Damjanov *et al.*, 1987), and from embryonic carcinoma cell lines derived from these teratocarcinomas (Martin and Evans, 1975a). Furthermore, embryoid body differentiation of primitive ectoderm-like (EPL) cells *in vitro* demonstrated the ability of this pluripotent cell population, downstream of primitive endoderm differentiation, to give rise to parietal endoderm, a primitive endoderm derivative (Lake *et al.*, 2000). These data suggest that competence to form primitive endoderm persists in the pluripotent cells after the formation of this extra-embryonic lineage *in vivo*. This conclusion is however complicated by the observation that microsurgically isolated ICMs at 4.5 d.p.c. do not regenerate primitive endoderm *in vitro*

(Gardner, 1985), and that isolated epiblast cells can not contribute progeny to primitive endoderm derivatives following blastocyst injection (Gardner and Rossant, 1979). However, these observations may reflect variations in the differentiation environment rather than the developmental potential of the cells.

1.4.1.3 Primitive ectoderm developmental potential

Although primitive ectoderm cells are morphologically indistinguishable, primitive ectoderm of pre-streak embryos appears regionalised with respect to developmental fate (Beddington, 1982; Tam and Beddington, 1987; Tam, 1989; Lawson *et al.,* 1991; Quinlan *et al.,* 1995). Spatially distinct primitive ectoderm cells display a more divergent fate the further apart they are situated (Figure 1.6). However, extensive developmental overlap is found between adjacent regions. Proximal posterior primitive ectoderm regions contain precursors of PGCs, extra-embryonic mesoderm and amniotic ectoderm, while lateral and more distal posterior primitive ectoderm is populated by cells predominantly of mesodermal fate and endodermal fate respectively. Anterior primitive ectoderm predominantly forms definitive ectoderm precursors (Figure 1.6).

However, retention of developmental lability within this population is evidenced by the fact that transplantation of primitive ectoderm from 6.5 d.p.c. (Parameswaran and Tam, 1995; Tam and Zhou, 1996) and 7.5 d.p.c. (Beddington, 1982) embryos into heterotopic sites within similar stage embryos results in the transplanted primitive ectoderm tissue adopting a developmental fate appropriate for the new location. Furthermore, analysis of primitive ectoderm cell potency by labelling individual primitive ectoderm cells and following their clonal descendants indicated that single primitive ectoderm cells are not developmentally committed and can contribute to multiple post-gastrulation lineages, including both somatic and germline tissues (Lawson *et al.,* 1991; Lawson and Hage, 1994). These results indicate that although regionalisation exists within areas of the primitive ectoderm, individual cells remain competent to form all post-gastrulation tissue. However, ingression through the primitive streak appears to restrict cell potency as mesoderm which has ingressed through the streak cannot contribute to lateral plate mesoderm following transplantation into the primitive ectoderm (Tam *et al.,* 1997). This suggests that ingressed cells,

which have been exposed to different patterning signals, are no longer able to contribute to the full array of tissues.

1.4.1.4 Summary

The standard classification of pluripotent cells (Figure 1.7) describes changes in morphology and differentiation, but does not reflect in a useful manner the competence of the pluripotent cells to form extra-embryonic lineages, which in many cases persist beyond their initial differentiation. The early potency studies indicated that loss of totipotence does not occur following allocation to the trophectoderm or ICM lineage, or even necessarily with the formation of these cell types. Furthermore, analysis of primitive endoderm formation suggests that competence to form this extra-embryonic cell type is not lost with the initial differentiation event. However, at some point totipotent or pluripotent cells become committed and lose the capacity to form extra-embryonic lineages. This suggests that intermediate populations may be present prior to differentiation of totipotent or pluripotent cells into a particular extra-embryonic lineage. Developmental lability appears to be retained until primitive ectoderm ingression through the primitive streak during gastrulation, when primitive ectoderm cells are exposed to patterning signals, differentiate and lose pluripotence. Retention of this developmental plasticity may underlie the embryo's ability to reprogram development following environmental perturbation, to regenerate lost tissues and allow development to proceed normally. Pluripotent cells which possess an altered developmental potential can not be recognised, neither do they correlate with the classification of pluripotent cells as a simple progression from totipotent blastomeres of the morula to the pluripotent ICM and subsequently to the primitive ectoderm population.

1.4.2 Gene Expression in Pluripotent Cells

Cell specification and axis patterning are dictated by changes in gene expression, such as genes encoding transcription factors, growth factors and receptors. Morphologically indistinguishable cell types can thus be subdivided according to the expression of specific marker genes. Analysis of gene expression in early pluripotent cell populations has been limited however by the small size and inaccessible nature of the mouse embryo *in utero* such that acquiring sufficient, homogeneous starting material for conventional screening approaches has been

inherently difficult. Additionally, while analysis of gastrulation and later embryological processes has greatly benefitcd from the study of lower organisms, fundamental differences in strategies during pre-gastrulation development has meant that extrapolation of gene expression information from lower metazoa to the mouse has not been informative. Despite these difficulties a growing number of genes and gene products which display restricted expression patterns within pluripotent cell populations, prior to entry into the primitive streak, have been isolated and characterised by *in situ* hybridisation or immunohistochemistry.

1.4.2.1 Differential gene expression during pre and early post-implantation development

The POU domain homeobox gene *Oct4* is the best characterised pluripotent cell specific transcript and is expressed by all pluripotent cells, including the fertilised oocyte, morula, ICM, primitive ectoderm, primordial germ cells and germ lines of the testis and ovary (Rosner *et al.*, 1990, Schöler *et al.*, 1990a and b, Yeom *et al.*, 1991). *Oct4* expression is down-regulated in differentiated derivatives of these cells, although Oct4 protein persists in the primitive endoderm (Palmieri *et al.*, 1994). Pluripotent cells also express alkaline phosphatase activity (Johnson *et al.*, 1977; Ginsburg *et al.*, 1990) and the cell surface carbohydrate recognised by the monoclonal antibody, SSEA-1 (Solter and Knowles, 1978). *Fgf4* transcripts and protein have been detected in the ICM of 3.5 d.p.c. blastocysts and the epiblast/primitive ectoderm cell populations prior to gastrulation (Niswander and Martin, 1992; Rappolee *et al.*, 1994). However, the gene becomes expressed in additional sites following gastrulation, and is not expressed by PGCs. The *Fgf4* gene has been shown to be regulated by *Oct4* in a *Sox2* dependent manner (Yuan *et al.*, 1995; Ambrosetti *et al.*, 1997; Fraidenraich *et al.*, 1998)

Expression of a number of genes can be used to identify temporally distinct pluripotent cell populations of the ICM from the primitive ectoderm. Several markers are expressed by the ICM during early development and include activin peptides which are expressed in the ICM at 3.5 d.p.c., with expression decreasing in the pluripotent cells during blastocyst maturation at 4.5 d.p.c. (Albano *et al.* 1993). *Gbx2*, a homeobox gene, is expressed by the ICM at 3.5-4.0 d.p.c., but not in the primitive ectoderm at 6.0 d.p.c. (Chapman *et al.*, 1997). *Rex1*, a zinc finger containing transcription factor whose expression has been shown to be regulated by *Oct4* in EC

cell cultures, is expressed in the epiblast at 4.5 d.p.c. but not in the primitive ectoderm at 6.0 d.p.c. (Rogers *et al.* 1991; Hosler *et al.*, 1993, Ben-Shushan *et al.*, 1998). *Fgf5*, a member of the fibroblast growth factor family, displays an alternate expression pattern with expression not detected in the ICM at 3.5 d.p.c., but detectable in the primitive ectoderm from 5.25-5.5 d.p.c. (Haub and Goldfarb 1991, Hébert *et al.*, 1991). A detailed expression analysis of these genes throughout pre- and peri-implantation development however has not been performed.

1.4.2.2 Regionalised gene expression within pre-gastrulation primitive ectoderm

An increasing number of genes demonstrate regionalised expression within the posterior and anterior primitive ectoderm of pre-gastrulation stage embryos (Figure 1.8). This indicates spatial heterogeneity within this pluripotent cell population and the formation of anterior-posterior asymmetry prior to streak formation (Tam and Behringer 1997).

Cripto, a putative signaling molecule of the EGF gene family, is expressed in the primitive ectoderm of the 5.5 d.p.c. embryo, but becomes asymmetrically distributed within the primitive ectoderm in a proximal to distal fashion in the prestreak embryo at 6.25 d.p.c. (Ding *et al.*, 1998). *Nodal*, a TGF-β signaling molecule, is first detected at low levels throughout the primitive ectoderm and overlying visceral endoderm at 5.5 d.p.c. (Conlon *et al.*, 1994; Varlet *et al.*, 1997). Prior to the onset of gastrulation, *nodal* exhibits graded expression in pluripotent cells with highest levels at the proximal boundary, in a broad zone in the posterior primitive ectoderm, and also a tightly restricted region of the anterior primitive ectoderm (Varlet *et al.*, 1997). *Fgf8* is expressed prior to gastrulation in a broad zone of posterior primitive ectoderm cells at 5.75 d.p.c., while *Evx1*, a homeobox gene, and *Wnt3*, a secreted signaling molecule, are expressed prior to gastrulation at 6.25 d.p.c. in proximal posterior primitive ectoderm cells, at the junction of the embryonic and extra-embryonic ectoderm (Dush and Martin, 1992; Crossley and Martin, 1995; Liu *et al.*, 1999). While these genes appear to share common expression boundaries at the most proximal region of primitive ectoderm, the distal expression boundaries of *Evx1/Wnt3*, *Fgf8* and *nodal* extend further towards the distal tip respectively (Figure 1.8). This indicates the existence of spatial heterogeneity in the prestreak primitive ectoderm and identifies pluripotent cell sub-populations which are either *Evx1/Wnt3/Fgf8/nodal* positive in the proximal posterior

Figure 1.8: Gene expression within the primitive ectoderm prior to the onset of gastrulation

Diagram illustrating regionalised gene expression in the pre-streak egg cylinder (5.75-6.25 d.p.c.). The expression domains of *nodal*, *Fgf8*, *Wnt3* and *Evx1* become progressively more restricted to the proximal posterior aspect of the embryo. The bold line represents expression in the primitive ectoderm, whereas the hatched line represents visceral endoderm expression. The anterior-posterior and proximal (P)-distal (D) axes are also indicated.

Anterior
P
Posterior
VE-1, Cer1,
Hesx1, nodal
nodal
Evx1/Wnt3
Fgf8
nodal
D

primitive ectoderm; *Fgf8/nodal* positive, *Evx1/Wnt3* negative; and *nodal* positive, *Evx1/Wnt3/Fgf8* negative in progressively more distal regions.

Anteriorly regionalised gene expression has also been observed prior to gastrulation within the anterior visceral endoderm and primitive ectoderm, and appears to be important in specifying anterior structures. The VE-1 antigen, recognised by an antibody of unknown epitope specificity, is detected in the anterior visceral endoderm from as early as 5.5 d.p.c. (Rosenquist and Martin, 1995), as is the secreted factor, *cer-1* (Belo *et al.*, 1997) (Figure 1.8). Prior to the onset of gastrulation (6.0 d.p.c.), in addition to proximal posterior asymmetric expression in the primitive ectoderm, the *nodal* growth factor is expressed in a morphologically thickened region of proximal, anterior visceral endoderm and also in the underlying anterior primitive ectoderm (Varlet *et al.*, 1997) (Figure 1.8). Chimeric embryos containing *nodal*-deficient primitive endoderm failed to form anterior neural structures, indicating that anterior-posterior positional information appears to be imparted from the anterior visceral endoderm to anterior primitive ectoderm (Varlet *et al.*, 1997). Additional evidence for the existence of an anterior visceral endoderm derived signal which patterns anterior structures comes from analysis of two homeobox genes, *Hesx1* and *Otx2*. *Hesx1* is expressed in a restricted area of anterior visceral endoderm at the proximal embryonic/extraembryonic boundary at 6.5 d.p.c., with expression expanding 24 hours later to include the primitive ectoderm immediately adjacent to the *Hesx1* positive visceral endoderm cells (Thomas and Beddington, 1996). Removal of this *Hesx1* expressing region of anterior visceral endoderm prevents later *Hesx1* expression in the ectoderm and neurectoderm, indicating that an inductive signal from the visceral endoderm instructs *Hesx1* expression in the underlying ectoderm (Thomas and Beddington, 1996). *Otx2* is expressed throughout the pluripotent cells and visceral endoderm at 5.75 d.p.c. and subsequently becomes restricted to the anterior visceral endoderm by the time the streak forms (Simeone *et al.*, 1993; Acampora *et al.*, 1995). Restricted anterior expression of *Otx2* within the primitive ectoderm is only observed at the subsequent mid-streak stage (7.75 d.p.c.) (Acampora *et al.*, 1995). This suggests that similar to *Hesx1*, *Otx2* expression in the anterior visceral endoderm may result in the generation of a signal which instructs expression of the gene in the adjacent ectoderm, a requirement for correct anterior patterning.

1.4.2.3 Summary

Gene expression analysis has therefore provided a glimpse into the complex nature of pluripotent cell regulation during pre- and early post-implantation development. While differential gene expression delineates ICM from primitive ectoderm within the pluripotent cell population, the temporal expression patterns of the relevant genes have not been mapped in sufficient detail to enable correlation of gene expression with key developmental events *in vivo*, or alterations in developmental potential. Regionalised gene expression is also observed in the posterior primitive ectoderm and in the anterior visceral endoderm and adjacent primitive ectoderm prior to gastrulation and appears important for patterning of the anterior-posterior axis.

1.4.3 Complex Regulation of Pluripotent Cell Populations during Peri-implantation Development

In the period between implantation and the onset of gastrulation the embryonic portion develops from a solid mass of rounded, apolar cells into a pseudostratified epithelial sheet of primitive ectoderm, surrounding a proamniotic cavity. A multitude of cellular events are required for this process, including proliferation, differentiation, proamniotic cavitation, cell polarisation and reorganisation of the newly formed primitive ectoderm into an epithelial sheet. Little is known about the genes involved in this complex array of cellular processes, however, loss of function mutants are beginning to provide valuable insight into their molecular controls (Table 1).

1.4.3.1 Pluripotent cell maintenance

Targeted disruption of the pluripotent cell-specific *Oct4* gene results in early embryonic death (Nichols *et al.*, 1998). *Oct4*$^{-/-}$ embryos undergo normal cleavage development and form blastocysts in the complete absence of Oct4, as maternal stores become depleted by the 2 cell stage. This suggests that *Oct4* is not required for growth or survival of the blastomeres. However, *Oct4* deficient embryos do not form an ICM. Furthermore, the developmental potential of *Oct4* null mutant embryos is restricted as pre-implantation blastocysts grown in culture as embryo outgrowths produced only TGCs, in comparison to wild type or heterozygous embryo cultures which give rise to ICMs which increased in size during cell culture, trophectodermal derivatives of extra-embryonic ectoderm and TGCs, migratory parietal endoderm cells and often

Gene	Gene Product	Pluripotent Cell Expression	Mutant Phenotype	Reference
β1 Integrin	ECM receptor	mRNA detected by RT-PCR in all pre-implantation stages, Expression N/D in PE	Retarded growth of ICM and no PEn differentiation	Sutherland *et al.*, 1993; Stephens *et al.*, 1995; Fässler and Meyer, 1995
Brca1	Breast cancer susceptibility gene	N/D	Defects in pluripotent cell proliferation from 4.5 d.p.c.; fail to form organised egg cylinders	Hakem *et al.* 1996; Liu *et al.* 1996; Ludwig *et al.* 1997
Brca2	Breast cancer susceptibility gene involved in DNA repair	Ubiquitous expression in all cells at 6.5 d.p.c.	Variable penetrance, impaired proliferation of PE by 5.5 d.p.c.	Ludwig *et al.* 1997; Sharan *et al.* 1997 Suzuki *et al.* 1997;
Bmp4	TGF-β growth factor	No pluripotent cell expression reported prior to gastrulation	Variable phenotype (arrest between 6.5-9.5 d.p.c.); defects in PE proliferation/survival & mesoderm induction	Winnier *et al.* 1995
Bmpr	Type 1 serine/threonine kinase TGF-β receptor which binds BMP2 & 4	No pluripotent cell expression reported prior to gastrulation	Defect in proliferation of PE, and no mesoderm formation	Suzuki *et al.* 1994; Mishina *et al.* 1995
EGFR	Tyrosine kinase receptor which binds EGF and TGF-α	Low level ICM expression reported, expression high in TE	Defects in pluripotent cell proliferation *in vitro,* no PEn differentiation and disorganised embryonic tissue	Dardik et al., 1992; Helden, 1995; Threadgill *et al.*, 1995
Evx1	even-skipped class homeobox gene	Proximal posterior PE at 6.25 d.p.c.	epiblast/early PE degeneration; may result from defect in VE differentiation	Dush and Martin 1992; Spyropoulos and Cappechi 1994
Fgf4	Fibroblast growth factor	Pluripotent cells of the blastocyst and PE but not PGCs	Defect in pluripotent cell proliferation following blastocyst implantation	Niswander and Martin 1992; Rappolee *et al.* 1994; Feldman *et al.* 1995

Table 1: Summary of loss of function mutants which display an effect on pluripotent cells prior to gastrulation, and there corresponding expression within pluripotent cells.

Gene	Gene Product	Pluripotent Cell Expression	Mutant Phenotype	Reference
Fgfr1	Tyrosine kinase linked FGF receptor	Low level expression in PE at 6.5 d.p.c.	Growth retardation by 6.5 d.p.c.; delayed and/ or defected mesoderm induction	Orr-Urtreger *et al.* 1991; Yamaguchi *et al.* 1992; Deng *et al.* 1994; Yamaguchi *et al.* 1994
Fgfr2	Tyrosine kinase linked FGF receptor	Low level expression in PE at 6.5 d.p.c.	Defect in pluripotent cell proliferation and differentiation from 4.5 d.p.c.	Orr-Urtreger *et al.* 1991; Arman *et al.* 1998
Fug1	Yeast RNA1 homologue	Expression induced in all cells after 6.5 d.p.c.	Impaired growth and reorganisation ability of PE	DeGregori *et al.* 1994
GATA 6	Zinc finger TF	No pluripotent cell expression, expressed by PEn derivatives at 6.5 d.p.c.	Smaller PE, increased apoptosis and defective VE differentiation	Morrisey *et al.*, 1998
Hnf4	Hepatocyte TF of the steroid hormone receptor superfamily	No pluripotent cell expression, expressed by VE from 4.5 d.p.c.	Increased apoptosis in PE by 6.5 d.p.c.; delayed and defective gastrulation	Duncan *et al.* 1994; Chen *et al.* 1994; Duncan *et al.* 1997
Oct4	Pou domain containing homeobox gene	All pluripotent cells	Essential for ICM pluripotency; ICM developmental potential restricted to TE lineage	Rosner *et al.* 1990; Schöler *et al.* 1990a,b; Yeom *et al.* 1990; Nichols *et al.* 1998
Rad51	*E. coli* RecA homologue	N/D	Defect in pluripotent cell proliferation by 5.5 d.p.c.	Lim and Hasty 1996
Smad4	Intracellular effector molecule which mediates TGF-β signalling.	Ubiquitous expression throughout embryo at 6.5 d.p.c.	Lethal between 6.5-8.5 d.p.c.; proliferation and differentiation defects in pluripotent cells in culture	Yang *et al.* 1998
Stat3	Signal transducer and TF	N/D	Smaller embryos at day 6.0 and impaired ICM outgrowth in culture	Takeda *et al.*, 1997

Abbreviations: TE = trophectoderm, ICM = Inner Cell Mass, PE = primitive ectoderm, PEn = primitive endoderm, VE = visceral endoderm, EEE = extra-embryonic ectoderm, PGC = primordial germ cells, d.p.c. = days *post coitum*, ECM = extracellular matrix, N/D = not determined, TF = transcription factor

visceral endoderm cells. Moreover, *Oct4* mutant cells in an internal position differentiate into trophectoderm as assayed by immunoreactivity using Troma-1, a monoclonal antibody specific for trophectoderm. Thus, *Oct4* appears essential for the formation and pluripotency of the ICM as loss of *Oct4* results in divergence to the trophectoderm lineage.

1.4.3.2 Paracrine signaling between ICM and extra-embryonic cells

Targeted disruption of *Fgfr2* (Arman *et al.*, 1998), and *Brca1* (Hakem *et al.*, 1996; Liu *et al.*, 1996; Ludwig *et al.*, 1997) results in deficiencies in trophectoderm development. When both *Brca1* and *Fgfr2* mutant embryos are cultured as *in vitro* outgrowths, trophectoderm-like cells and TGCs form in these cultures, indicating that initial trophectoderm differentiation is not compromised. Furthermore, *Brca1* mutant embryos are capable of invading the uterine wall and inducing decidualisation, a marker of the functional integrity of trophectoderm. However, *Brca1* mutant embryos display abnormal extra-embryonic ectoderm development with an absence of diploid trophoblast cells and elevated levels of TGCs.

Fgfr2 deficient embryos are able to invade the uterine wall and induce increased capillarisation, but mutant embryos fail to trigger a complete decidual response, indicating a loss of functional integrity of trophectoderm in $Fgfr2^{-/-}$ embryos. Similarly, expression of a dominant negative FGF receptor (dnFGFR) in the extra-embryonic ectoderm, which abrogates all FGF signaling, results in extra-embryonic ectoderm cells becoming post-mitotic and migratory (Chai *et al.*, 1998), suggesting differentiation into TGCs.

These defects in trophoblast development may result from either a direct, cell autonomous requirement of these genes during trophoblast differentiation, or indirectly from an absence of trophic signals from an impaired ICM, which are required for the continued propagation and diploid status of polar trophectoderm (Gardner *et al.*, 1973; Gardner and Papaioannou, 1975; Rossant and Ofer, 1977). Recent studies have implicated FGF signaling in the trophic ability of the ICM on polar trophectoderm, as increased growth of extra-embryonic ectoderm in blastocyst outgrowths following exogenous addition of FGF4 has been reported (Chai *et al.*, 1998; Nichols *et al.*, 1998). The very similar peri-implantation lethal phenotype of *Fgf4* and *Fgfr2* nullizygous mutants (Feldman *et al.*, 1995; Arman *et al.*, 1998) and the reciprocal expression profiles of *Fgf4* and *Fgfr2* by the epiblast and trophectoderm respectively (Niswander and Martin, 1992; Arman *et*

al., 1998; Haffner-Krausz *et al.,* 1999), suggest that binding of pluripotent cell derived FGF4 to FGFR2 on polar trophectoderm mediates the increased proliferation of the extra-embryonic ectoderm. Consistent with this, Tanaka *et al.,* (1998) have isolated trophoblast stem cell lines *in vitro* in the presence of FGF4. Upon removal of FGF4 from the culture media the cells differentiate into TGCs displaying polyploid DNA contents.

Therefore, interaction of FGF4 with FGFR2 expressed by the polar trophectoderm appears to be a component of a pluripotent cell derived, paracrine signal required for proliferation and maintenance of the undifferentiated/diploid status of polar trophectoderm, indicating a role for the ICM in directing trophectoderm development. The fact that *Brca1*$^{-/-}$ embryos display an absence of diploid trophoblast cells and increased numbers of TGCs suggests that *Brca1* may also play a role in this signaling pathway. However, further analysis, such as the sites of embryonic expression of *Brca1*, will be required to determine its role in trophoblast development.

Fgf4, Fgfr2, β1 integrin, and *EGFR* nullizygous mutants have been reported to display impaired primitive endoderm differentiation in culture (Fässler and Meyer, 1995; Feldman *et al.,* 1995; Stephens *et al.,* 1995; Threadgill *et al.*, 1995; Arman *et al.*, 1998). However, further analysis of these genes in primitive endoderm differentiation and morphogenesis was not undertaken in the reported studies. Given that *Fgf4/Fgfr2* signaling is thought to mediate paracrine signaling between pluripotent cells and polar trophectoderm, subsequent lack of primitive endoderm differentiation in these mutants may indicate that signaling between the pluripotent cells and the primitive endoderm is also required for primitive endoderm formation.

1.4.3.3 Pluripotent cell proliferation

Following blastocyst implantation but prior to gastrulation, pluripotent cells undergo more rapid proliferation (Snow, 1977). Pluripotent cell proliferation is critical for embryogenesis since a threshold number of cells is required for the initiation and progression of gastrulation. Disruptions to pluripotent cell proliferation results in a delay in the onset of gastrulation, whereas doubling the numbers of pluripotent cells by aggregating embryos results in a reduced rate of proliferation, rather than the early onset of gastrulation (Snow and Tam, 1979; Tam, 1988; Power and Tam, 1993). These results indicate that pluripotent cell proliferation is a coordinated,

flexible process, capable of modulation in response to environmental factors (Tam and Behringer, 1997).

(i) Early proliferative pluripotent cell defects:

A number of genes, including *Fgf4*, *Fgfr2*, *Brca1*, *β1 integrin*, and *EGFR* (Table 1) have been implicated in controlling the proliferation of pluripotent cells in peri-implantation development via analysis of loss-of-function mutations (Fässler and Meyer, 1995; Feldman *et al.*, 1995; Stephens *et al.*, 1995; Threadgill *et al.*, 1995; Hakem *et al.*, 1996; Liu *et al.*, 1996; Ludwig *et al.*, 1997; Arman *et al.*, 1998). Embryos from all of these null mutations develop normally to the blastocyst stage, but degenerate shortly after implantation. Egg cylinders fail to form and the embryonic component of these embryos is reduced in size. The mutants display impaired ICM proliferation when mutant blastocysts are cultured *in vitro,* with the impaired ICM proliferation of *Fgf4*$^{-/-}$ embryos rescuable by the addition of exogenous FGF4 (Feldman *et al.*, 1995).

The cellular basis of these proliferation defects can be difficult to interpret as direct measurement of proliferation has not been analysed in many of these mutants. Nevertheless, reduced pluripotent cell numbers in *Brca1*$^{-/-}$ embryos do not result from elevated levels of apoptosis, as measured using TUNEL assays (Gavrieli *et al.*, 1992), but from reduced proliferation determined by 5-bromo-2'-deoxyuridine (BrdU) incorporation, and an accompanying reduction in cyclin E levels (Hakem *et al.*, 1996; Liu *et al.*, 1996).

The similar arrest phenotypes observed for *Fgf4* and *Fgfr2* knockouts, raises the possibility that proliferation of pluripotent cells is mediated in some part by FGF4 signaling transmitted by FGFR2 (Arman *et al.*, 1998). A more definitive role for FGF signaling in early embryonic cell proliferation emerged from experiments using transgenic expression of a dominant negative FGF receptor (dnFGFR) (Chai *et al.*, 1998). Expression of the dnFGFR inhibits all FGF signal transduction, removing the complicating effects of persisting maternal FGFs and compensatory effects of other FGFs and FGF receptor family members. In blastocysts expressing the dnFGFR in all cells, the ICM forms and appears morphologically normal. However, cell division ceases after the fifth cell division, prior to implantation, with the ICM containing one third the number of normal cells. Increased pyknotic nuclei were not observed, indicating that FGF signaling had an effect on proliferation and not in preventing cell death.

(ii) Proliferation defects in pluripotent cells following egg cylinder formation:

Other targeted mutations affect pluripotent cell proliferation at later stages of mouse development, following formation of an egg cylinder. In addition to other defects observed, mutation of *Fgfr1* (Deng *et al.*, 1994; Yamaguchi *et al.*, 1994), *Brca2* (Ludwig *et al.*, 1997; Sharan *et al.*, 1997; Suzuki *et al.*, 1997), *Rad51* (Lim and Hasty, 1996), *Bmp4* (Winnier *et al.*, 1995), *Bmpr* (Mishina *et al.*, 1996), *Smad4* (Yang *et al.*, 1998), *Stat3* (Takeda *et al.*, 1997) and *Fug1* (DeGregori *et al.*, 1994) all result in reduced numbers of pluripotent cells at 6.5 d.p.c., and delayed or failed gastrulation, with many also displaying disorganised egg cylinders (Table 1). Outgrowth of pluripotent cells from *Fgfr1*$^{-/-}$, *Rad51*$^{-/-}$, *Brca2*$^{-/-}$, *Smad4*$^{-/-}$ and *Stat3*$^{-/-}$ blastocysts cultured *in vitro* was impaired. Direct measurement of primitive ectoderm proliferation *in vivo* by BrdU incorporation revealed reduced proliferation at day 5.5, 6.5, and 7.5 of embryogenesis in *Rad51*$^{-/-}$ embryos, at day 6.5 of embryogenesis in *Bmpr*$^{-/-}$ embryos, and at day 7.5 of embryogenesis in *Brca2*$^{-/-}$ embryos. *Smad4*$^{-/-}$ embryos also displayed only weak staining for proliferating cell nuclear antigen, a marker of proliferation, at 6.5 d.p.c. compared with strong staining in wild type and heterozygous litter mates.

(iii) Summary:

Therefore, there appear to be two stages of embryogenesis in which null mutations in an array of genes results in pluripotent cell proliferation defects; the first around the time of, or shortly following implantation, and the second following egg cylinder formation. FGF signaling appears to be important in both temporally distinct proliferation phenotypes, indicating that two separate FGF pathways may be important in regulating pluripotent cell proliferation during early mammalian development. The temporal distinction between *Fgf4/Fgfr2* and *Fgfr1* null phenotypes with respect to proliferation defects, indicates that a switch from FGFR2 to FGFR1 signaling prior to gastrulation may occur. Early pluripotent cell proliferation appears dependent on FGF4 binding to FGFR2. This may represent a direct autocrine or paracrine effect of FGF4 on pluripotent cells given that *Fgf4* expression at this stage of development is restricted to the pluripotent cells and that *Fgfr2* is detectable at low levels in the epiblast population (Niswander and Martin, 1992; Haffner-Krausz *et al.*, 1999). The contribution of *Fgfr1* signaling during peri-implantation development is not known as expression of *Fgfr1* prior to gastrulation has not been

characterised. Recent studies suggest a role for FGF4 in mediating paracrine signaling to polar trophectoderm (Chai *et al.*, 1998; Nichols *et al.*, 1998; Tanaka *et al.*, 1998). An indirect effect of FGF4 on pluripotent cell proliferation via trophectoderm signaling can therefore not be discounted, as loss of a polar trophectoderm derived signal required for pluripotent cell proliferation may occur in response to loss of FGF4/FGFR2 signaling in the polar trophectoderm.

Targeted disruption of *Bmp4,* a TGF-β growth factor gene, *Bmpr,* which encodes a transmembrane receptor with specificity for BMP2 and BMP4 (Suzuki *et al.*, 1994) and *Smad4*, an intracellular effector molecule which mediates TGF-β signals (Massague, 1996), resulted in similar proliferation defects in the primitive ectoderm following formation of the egg cylinder. Therefore, signals mediated by BMP molecules binding to *Bmpr* may also be important in the control of pluripotent cell proliferation following egg cylinder formation.

1.4.3.4 Primitive ectoderm formation and maintenance; an instructive role for visceral endoderm

During peri-implantation development, an organised epithelial layer of primitive ectoderm forms from the epiblast bud. Concomitant with this transition, primitive endoderm differentiates into either visceral endoderm, if cells remain in contact with pluripotent cells, or parietal endoderm, if cells lose contact with pluripotent cells (Gardner, 1983). Recently, visceral endoderm has emerged as a key regulator controlling primitive ectoderm formation and maintenance.

Prior to its expression in the primitive ectoderm at 6.25 d.p.c., the homeobox gene *Evx1* has been detected at low levels throughout the visceral endoderm (Spyropoulos and Capecchi, 1994). Ablation of *Evx1* results in degeneration of pluripotent and extra-embryonic cells around 5.0 d.p.c.. Mutant embryos implant, but show very little growth or differentiation thereafter, are disorganised and undergo resorption by day 6.5 of embryogenesis. Defective developmental progression of the pluripotent cells is interpreted to result from perturbation of communication between the visceral endoderm and pluripotent cells (Spyropoulos and Capecchi, 1994), although a direct role for *Evx1* in primitive ectoderm maintenance has not been established experimentally.

Targeted deletion of *Hnf4*, a winged helix transcription factor which is expressed in primitive and visceral endoderm between 4.5 and 7.5 d.p.c (Duncan *et al.*, 1994) also results in perturbation of pluripotent cell development with increased apoptosis within the primitive ectoderm by 6.5 d.p.c., delayed gastrulation, and death between 9.5 and 10.5 d.p.c. (Chen *et al.*, 1994). Chimeric tetraploid embryos, formed from $Hnf4^{-/-}$ embryos recombined with $Hnf4^{+/+}$ visceral endoderm, produced healthy egg cylinders and gastrulated normally (Duncan *et al.*, 1997). This suggests that expression of *Hnf4* within the visceral endoderm is essential for primitive ectoderm survival. Targeted disruption of *GATA6*, a zinc finger transcription factor, expressed within the visceral endoderm from 6.5 d.p.c., displayed a similar phenotype to that of $Hnf4^{-/-}$ embryos, with smaller embryonic regions and increased apoptosis within the primitive ectoderm (Morrisey *et al.*, 1998). This primitive ectoderm defect appeared to be a result of incomplete visceral endoderm differentiation as reduced expression of visceral endoderm markers such as *Hnf4* and *GATA4* at 6.5 d.p.c., and *Hnf4*, *GATA4*, *HNF3β* and *AFP* in embryoid bodies formed from $GATA6^{-/-}$ ES cells was reported. Interestingly, GATA6 was capable of inducing *Hnf4* expression, suggesting that GATA6 may lie upstream of HNF4 in a transcriptional pathway that regulates visceral endoderm differentiation (Morrisey *et al.*, 1998). Taken together, these results indicate that signaling from visceral endoderm is required for pluripotent cell survival.

Embryoid bodies (1.5.3.2) have been used as an experimental model to investigate the role of visceral endoderm in proamniotic cavity formation within the epiblast bud and subsequent survival of the primitive ectoderm. Coucouvanis and Martin (1995) propose that a diffusible "death" signal, emanating from the visceral endoderm, induces apoptosis of inner cells, whilst an ECM localised signal mediates survival of a layer of pluripotent cells in contact with the ECM. Therefore, spatially localised signals appear to induce alternate cellular functions within the pluripotent cells as they develop and form primitive ectoderm, indicating heterogeneity within this population. BMP-mediated signaling has been implicated in controlling the proamniotic cavity formation process (Coucouvanis and Martin, 1999), as blocking all BMP signaling via overexpression of a dominant negative BMP receptor prevents cavitation of inner cells in embryoid bodies. These embryoid bodies also fail to express the visceral endoderm marker, *Hnf4*, indicating that visceral endoderm differentiation does not occur. Addition of exogenous

BMP protein to embryoid bodies which are unable to undergo cavitation induces expression of visceral endoderm markers and cavitation. These results suggest a role for BMP signaling between pluripotent cells and the visceral endoderm lineage, which may be important in visceral endoderm differentiation and subsequent proamniotic cavitation. Compensatory actions by other BMP family members may account for the lack of a similar phenotype being observed in either *Bmp2* or *Bmp4* knockouts.

1.4.3.5 Summary

The analysis of gene knockout phenotypes demonstrates the existence of both temporally and spatially localised signals which emanate from both pluripotent cells and extra-embryonic lineages, to influence cellular decisions such as the maintenance of pluripotency, proliferation, differentiation and survival. This analysis provides circumstantial evidence for the presence of transient pluripotent cell sub-populations during early embryogenesis, which respond to temporally and spatially localised signals. Further analysis of the molecular controls and developmental behaviour will be fundamental to understanding the processes that establish and regulate the precursors of gastrulation. It is anticipated that pluripotent cell sub-populations will express an array of functionally important, sub-population specific genes. Therefore, isolation of such genes and close expression mapping during peri-implantation development *in vivo* may allow recognition of pluripotent cells of a diverse developmental potential and may recognise the pluripotent cells which respond to the autocrine or paracrine signals of early embryogenesis.

1.5 *IN VITRO* MODELS OF EARLY EMBRYOGENESIS

Considerable insight into stem cell biology has been obtained from the analysis of *in vitro* models, most notably in the haematopoiesis field (Metcalf, 1989; Metcalf, 1991). Hence, a particularly profitable route for investigation of pluripotent cell biology during early stages of mouse embryogenesis has been the establishment of *in vitro* culture systems of early pluripotent cells.

1.5.1 Embryonal Carcinoma Cells

Early analysis of pluripotent cells was carried out using embryonal carcinoma (EC) cells which provided the first *in vitro* model system for studying pluripotent cell gene expression and differentiation potential. EC cells were isolated from the multipotential cells of complex germ cell tumours termed teratocarcinomas, which are surrounded by differentiated endodermal, ectodermal and mesodermal tissues (Martin and Evans, 1974). EC cell lines are similar to either ICM cells or primitive ectoderm cells in morphology, cell surface marker expression and protein synthesis profiles. In addition, many EC cell lines have the capacity to differentiate into multiple cell lineages *in vitro* and *in vivo* (Martin and Evans, 1975a and b; Papaioannou *et al.*, 1975; Martin *et al.*, 1978; Solter and Knowles, 1978; Martin, 1980; Hosler *et al.*, 1989). However, when induced to differentiate, specific EC cell lines display only a partial differentiation potential. For instance, F9 EC cells are restricted to forming primitive endodermal lineages when chemically induced by retinoic acid (RA) or di-butyryl cAMP (Strickland and Mahdavi, 1978; Strickland *et al.*, 1980; Hogan *et al.*, 1983; Mummery *et al.*, 1990) and P19 EC cells give rise to muscle and neuronal lineages when differentiated as aggregates in the presence of dimethyl sulphoxide (DMSO) or RA respectively (Jones-Vileneuve *et al.*, 1982; McBurney *et al.*, 1982). It has been suggested that developmentally restricted EC cells represent transient pluripotent cell intermediates, frozen by transformation or *in vitro* isolation procedures. Many EC cell lines can contribute to chimera formation following reintroduction into mouse blastocyst, but are limited in their developmental potential, as apart from a single report (Stewart and Mintz, 1981), EC cells are unable to colonise the germ line. This difference between EC cells and pluripotent cells of the early embryo may be a result of the accumulation of mutations, including karyotypic alterations, following extended proliferation *in vivo* and *in vitro,* which are incompatible with gametogenesis (Papaioannou and Rossant, 1983).

Despite these limitations, investigation of EC cell lines has allowed the identification of biologically relevant genes and signaling molecules (Mummery *et al.*, 1990; Hosler *et al.*, 1989; Coucouvanis and Martin, 1995; Coucouvanis and Martin, 1999). The use of EC cell lines for investigation of later aspects of mouse development, such as neural development, has also been important. Finally, isolation and analysis of EC cells provided the initial foundations required for the isolation of ES cell lines directly from pre-implantation mouse blastocysts.

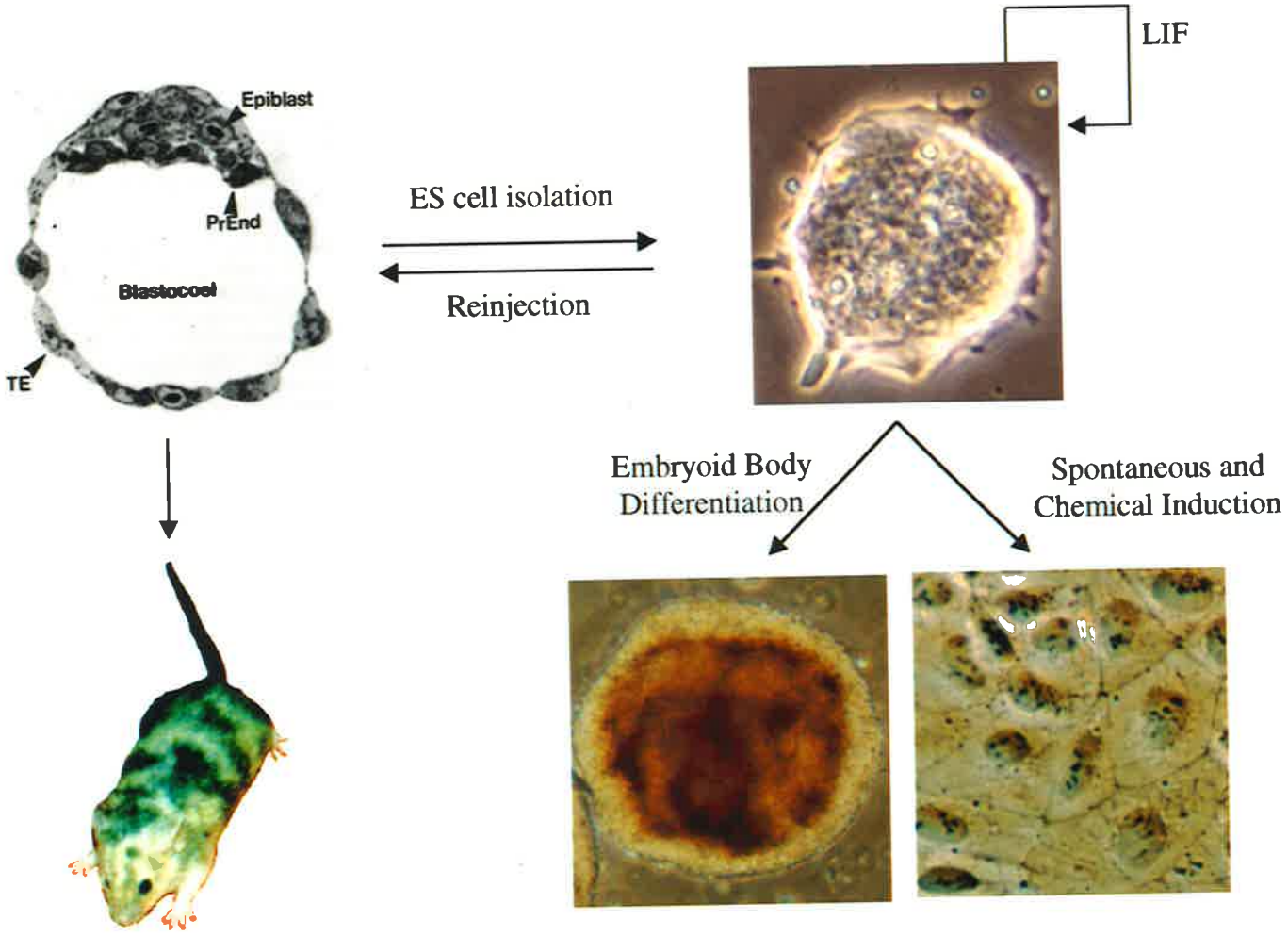
1.5.2 Isolation of Embryonic Stem Cells

Cell lines with properties more akin to the pluripotent cells of early development were obtained from ICM cells of the pre-implantation blastocyst of the inbred 129 mouse strain (Evans and Kaufman, 1981; Martin, 1981). These cells have been termed embryonic stem (ES) cells and grow as compact, three dimensional colonies of small cells with large nuclei (Figure 1.9). ES cells retain characteristics of the pluripotent cell population from which they were isolated with respect to gene expression and differentiation potential, even after long periods in culture (Beddington and Robertson, 1989). They continue to express markers of pluripotent cell types, such as *Oct4*, *Rex1*, SSEA-1 antigens, and alkaline phosphatase activity in culture (Solter and Knowles, 1978; Martin, 1981; Hahnel *et al.,* 1990; Rosner *et al.,* 1990; Schöler *et al.,* 1990b; Rogers *et al.,* 1991; Yeom *et al.,* 1991). ES cells are maintained in an undifferentiated state by growth on feeder layers of mitotically inactivated fibroblasts or in the presence of leukaemia inhibitory factor (LIF) or other members of the IL-6 cytokine family (Moreau *et al.,* 1988; Smith *et al.,* 1988; Williams *et al.,* 1988; Gearing and Bruce, 1992; Conover *et al.,* 1993; Piquet-Pellorce *et al.,* 1994; Pennica *et al.,* 1995).

ES cells retain the ability, even after prolonged periods of culture, to contribute functional progeny to all tissues of the embryo proper, including the germ line, upon reintroduction into host blastocysts (Bradley, *et al.,* 1984; Beddington and Robertson, 1989). ES cells therefore retain the ability to generate and respond to the regulatory signals of mouse development, and thus provide an *in vitro* model system for the investigation of pluripotent cell developmental decisions. The demonstration of germ line transmission, in addition to the *in vitro* cell culture of ES cells, has enabled effective manipulation of the mouse genome by allowing selection for rare events, such as homologous recombination, and functional analysis of the genomic perturbations *in vivo*. Following precise modification of a gene of interest by homologous recombination in ES cells, altered ES cells can be reintroduced into the blastocyst where they participate in development and contribute to all cell types. Provided the ES cells contribute to the germ line, the resulting chimeric mice can be bred to homozygosity for the modified locus, and mutant mice analysed for functional defects. Targeted removal or alterations in specific genes enables *in vivo* analysis of the gene of interest during mouse development, including early developmental decisions, gastrulation, axis patterning and organogenesis, as well as gene function in the adult,

Figure 1.9: Schematic representation of the ES cell model of early mouse development

ES cells are isolated from the ICM of the mouse blastocyst and maintained indefinitely in culture in the presence of Leukaemia Inhibitory Factor (LIF). ES cells grown *in vitro* are capable of participating normally in mouse development following re-injection into host blastocysts. ES cells can be differentiated *in vitro* in a variety of ways which mimic varying aspects of development by removal of LIF, addition of chemical inducers, or embryoid body formation.


Epiblast
PrEnd
Blastocoel
TE
ES cell isolation
Reinjection
LIF
Embryoid Body
Differentiation
Spontaneous and
Chemical Induction

and allows the development of mouse models of human disease (reviewed by Brandon *et al.*, 1995 a, b and c; Shearwin-Whyatt *et al.*, 1999).

The precise embryonic cell type represented by ES cells *in vitro*, if indeed there is one, has not been unequivocally determined and remains controversial (Smith, 1992). ES cells were isolated from the pluripotent cells of the pre-implantation stage blastocyst. Some researchers believe that ES cells are more akin to the ICM at this stage as following reintroduction into host blastocysts ES cells are capable of contributing to trophectoderm, although rarely, and only following injection of multiple ES cells and not single ES cells (Beddington and Robertson, 1989). Furthermore, ES cells are capable of differentiation into trophectoderm derivatives in culture (Beddington and Robertson, 1989). In contrast, others suggest that ES cells are more closely related to 4.5 d.p.c. epiblast cells, as single 4.5 d.p.c. epiblast cell suspensions produce approximately twice as many ES cell lines compared with intact pre-implantation blastocysts (Brook and Gardner, 1997). No trophoblast giant cell differentiation is observable in epiblast derived ES cell lines, suggesting that differentiation of trophoblast giant cells in ES cell cultures isolated from pre-implantation blastocysts is more likely to arise from contaminating trophoblast stem cells within the ES cell culture than from the ES cells themselves (Brook and Gardner, 1997).

Since contradictory information regarding the exact relationship of ES cells to the ICM or later epiblast population is apparent, other researchers have suggested that a direct embryonic equivalent of ES cells does not exist. Instead ES cells may represent a generic *in vitro* pluripotent cell population (Rossant, 1993). ES cells may represent pluripotent cells which have undergone reprogramming or de-differentiation, and have lost developmental commitment. Therefore, they demonstrate a greater versatility than their *in vivo* counterparts, which enables them to respond to all the signals of the host blastocyst following reinjection (Smith, 1992). Alternatively, ES cells may represent an adapted pluripotent cell type, redirected from normal development due to *in vitro* isolation procedures and continued culturing in a proliferation based program (Rossant, 1993). Embryonic germ (EG) cells can be isolated from the PGC population in the 8.5 d.p.c. and 12.5 d.p.c. embryo of the 129 strain of mouse in the presence of LIF, steel factor and basic FGF (Matsui *et al.*, 1992; Resnick *et al.*, 1992). These cells posses an ES cell-like morphology, are pluripotent, and like ES cells can participate in normal development and transmit through the

germ line. Pluripotent cells isolated from post-proamniotic cavitated primitive ectoderm in the presence of a hepatocellular carcinoma conditioned media and LIF also revert to an ES cell-like phenotype and can be maintained in culture for extended periods of time (J. Rathjen, unpublished data). This ability of pluripotent primitive ectoderm and PGCs to convert to an ES cell-like intermediate in culture suggests that reprogramming has occurred and may represent a generalised feature of a pluripotent cell state with the ability to revert through a common intermediate (Rossant, 1993).

1.5.3 Differentiation of ES Cells *in vitro*

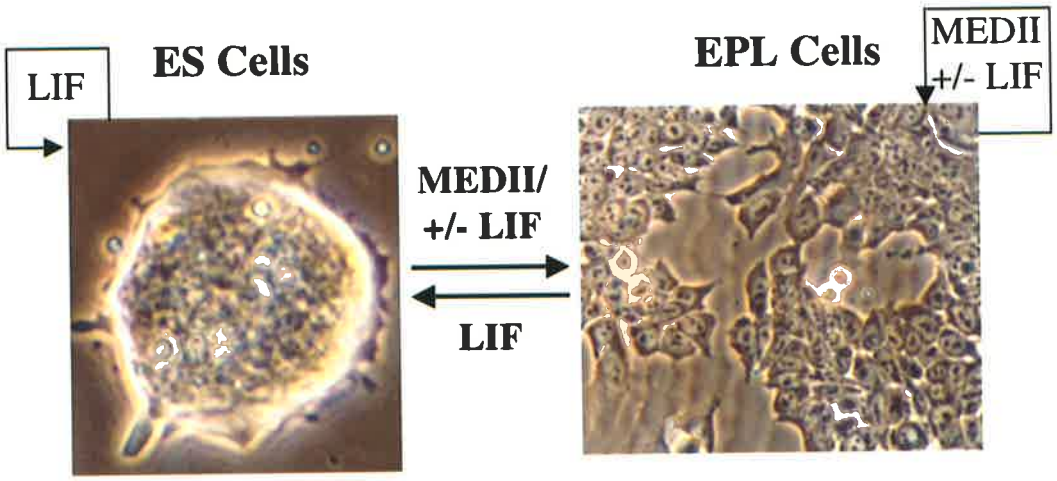
The ability to maintain ES cells in the absence of feeder layers by the addition of LIF has allowed the investigation of ES developmental potential and the establishment of differentiation strategies and *in vitro* models of development. ES cells can be differentiated by a number of means as illustrated in Figure 1.9, including differentiation as monolayer cultures by withdrawal of LIF or addition of chemical inducers, as aggregates or embryoid bodies which mimic early mouse development, and more recently by lineage specific differentiation regimes (Figure 1.10).

1.5.3.1 Spontaneous and chemical induction of ES cell differentiation

ES cells can be differentiated in monolayer culture *in vitro* by the withdrawal of LIF which results in spontaneous ES cell differentiation, and by the addition of chemical inducers, such as RA, DMSO or methoxybenzamide (MBA) (Smith, 1991). Spontaneous ES cell differentiation results in the appearance of a number of poorly defined cell lineages and is often incomplete, with formation of residual stem cell nests surrounded by terminally differentiated progeny. This may occur as a result of a feedback mechanism involving increased LIF expression by differentiated cells, to ensure that a stem cell population is maintained (Rathjen *et al.*, 1990b). ES cells can be differentiated in a more specific manner using RA or MBA, generating a fibroblast-like cell type and parietal yolk sac-like cells, and a relatively uniform epithelial-like cell type, respectively (Smith, 1991). However, the restricted array of cell types produced, and the doubtful biological significance of differentiation by chemical induction has limited the usefulness of these differentiation regimes for investigating embryogenesis, although developmentally regulated

Figure 1.10: Conversion of ES cells to EPL cells, an *in vitro* model of the ICM to primitive ectoderm transition *in vivo*

Schematic representation of the formation of EPL cells from ES cells *in vitro*. ES cells can be induced to differentiate into a uniform, morphologically distinct cell type, termed EPL cells, by the addition of conditioned media from a human hepatocellular carcinoma cell line (MEDII), in the presence or absence of LIF (+/- LIF). Maintenance of the EPL cell phenotype requires continued culture in MEDII. EPL cells can also revert to an ES cell like state following withdrawal of MEDII in the presence of LIF.


LIF
ES Cells
MEDII/
+/- LIF
LIF
EPL Cells
MEDII
+/- LIF

genes such as *Rex1* (Hosler *et al.*, 1989) have been isolated upon chemical induction of the F9 EC cell line.

1.5.3.2 Embryoid body differentiation

A more accurate *in vitro* model of early embryogenesis utilises the ability of ES cells to differentiate as embryoid bodies, ES cell aggregates cultured in the absence of LIF and in an environment which prevents adherence to a substratum. Embryoid bodies undergo differentiation in an ordered and predictable manner reminiscent of early mouse development (Doetschman *et al.*, 1985; Robertson, 1987). Initially, the outer cells of the ES cell aggregates differentiate to form primitive endoderm, mimicking the formation of primitive endoderm by ICM cells exposed to the blastocoelic cavity, and over a period of some days the primitive endoderm differentiates further into visceral and parietal endoderm. The inner cells of the aggregate remain pluripotent and undergo a complex, organised cellular progression, involving proliferation, differentiation and selective cell apoptosis and survival, to form a cavitated monolayer of primitive ectoderm, recapitulating early developmental decisions of the mouse embryo. After the formation of primitive ectoderm, the pluripotent cells of the embryoid body begin to differentiate as evidenced by a loss in *Oct4* expression, indicative of the loss of pluripotent cells, and up-regulation of mesoderm specific markers, such as *brachyury* and *goosecoid* and the appearance of terminally differentiated lineages from the three primary germ layers (Doetschman *et al.*, 1985; Wilkinson, *et al.*, 1990; Blum *et al.*, 1992; Shen and Leder, 1992; Keller *et al.*, 1993; Johansson and Wiles, 1995). Although many of the processes of differentiation within the embryoid body mimic the process of gastrulation *in vivo*, embryoid bodies lack the organisational information and axis patterning present within the embryo.

Therefore, differentiation of ES cells in a complex embryoid body environment provides an effective *in vitro* system which recapitulates many of the differentiation events undertaken by pluripotent cells of early mouse embryogenesis, including primitive endoderm differentiation, the formation of a primitive ectoderm epithelium, and mesoderm formation. Embryoid bodies have been utilised as a model system to study embryological events such as identification of the signals and events involved in proamniotic cavitation (Coucouvanis and Martin, 1995), cardiac development (Lyons *et al.*, 1995), and neural differentiation (Strübing *et al.*, 1995).

1.5.3.3 Lineage specific ES cell differentiation

While factors which prevent differentiation of ES cells *in vitro* have been identified, biologically significant factors which induce ES cell differentiation in culture in a manner reminiscent of normal embryogenesis have not. Recently it has been demonstrated that ES cells can be differentiated in a lineage specific manner, into a uniform, morphologically distinct cell type by the addition of HepG2 conditioned media (MEDII) in the presence or absence of LIF (Rathjen *et al.,* 1999) (Figure 1.10). This cell type shared morphological similarities with the primitive ectoderm-like P19 EC cell line (Rudnicki and McBurney, 1987), such as growth as a monolayer with distinguishable cells, visible nuclei and prominent nucleoli. Analysis of gene expression (Bettess, 1993; Rathjen *et al.,* 1999) demonstrated that this cell type expressed the pluripotent cell markers, *Oct4*, SSEA-1 antigen and alkaline phosphatase activity, indicating that these cells were pluripotent (Solter and Knowles, 1978; Johnson *et al.,* 1977; Hahnel *et al.,* 1990; Rosner *et al.,* 1990; Schöler *et al.,* 1990a; Yeom *et al.,* 1991). E-cadherin expression within this cell type indicated that they did not represent PGCs (Sefton *et al.,* 1992), and consistent with the pluripotent cell marker expression, EPL cells did not express *H19*, a marker of extra-embryonic lineages (Poirier *et al.,* 1991), *alphafetaprotein*, a marker of visceral endoderm (Dziadek and Adamson, 1978), nor primitive streak and nascent mesoderm markers (*Evx1*, Dush and Martin, 1992; *brachyury*, Wilkinson *et al.,* 1990; Herrmann, 1991; and *goosecoid*, Blum *et al.*, 1992). However, these cells displayed a distinct gene expression profile when compared with ES cells. ES cells demonstrated high *Rex1* and *Gbx2,* but barely detectable or no *Fgf5* expression, a gene expression profile akin to the ICM/early epiblast *in vivo* (Rogers *et al.,* 1991; Haub and Goldfarb, 1991; Chapman *et al.,* 1997). In contrast, EPL cells demonstrated progressively more elevated levels of *Fgf5* expression and downregulated *Rex1* and *Gbx2* expression over a period of 4-6 days in culture, a gene expression profile reminiscent of the early primitive ectoderm (Rogers *et al.,* 1991; Haub and Goldfarb, 1991; Hébert *et al.,* 1991; Chapman *et al.*, 1997). Therefore, this novel cell type expresses an array of marker genes which is shared only by the primitive ectoderm population within the embryo and consequently the cells were termed EPL cells (Early Primitive ectoderm-Like cells).

Pluripotent cells from pre-implantation embryos have the ability to contribute to embryo development following blastocyst injection, however, pluripotent cells from post-implantation

embryos do not (Gardner, 1971; Rossant, 1977; Beddington, 1983; Brook and Gardner, 1997). EPL cells, in contrast to ES cells, were unable to contribute to chimeras following blastocyst injection, consistent with EPL cells representing an early primitive ectoderm like cell population (Rathjen *et al.,* 1999). Preliminary studies also demonstrated that EPL cells, but not ES cells, were capable of differentiating in response to cytokines such as activin A and basic FGF, which in *Xenopus* are capable of inducing mesoderm, indicating that EPL cells were more akin to the primitive ectoderm, the substrate for gastrulation (Lake, 1996). Finally, EPL cells were competent to form embryoid bodies and differentiated in a coordinated manner similar to early embryogenesis, confirming the pluripotent nature of EPL cells. Comparative analysis of ES and EPL cell differentiation using embryoid bodies demonstrated that EPL cells formed mesoderm earlier and at higher levels, as indicated by the expression of nascent mesoderm markers, *brachyury* and *goosecoid* (Wilkinson *et al.,* 1990; Herrmann, 1991; Blum *et al.,* 1992) and the formation of differentiated mesoderm cell types such as cardiac muscle and haematopoietic lineages (Lake *et al.*, 2000). This was consistent with the supposition that EPL cells were analogous to a pluripotent cell type downstream of ES/ICM cells. However, EPL cells were restricted in their ability to form visceral endoderm and neurons in embryoid bodies compared to ES cells (Lake *et al.,* 2000). Nevertheless, combination of ES cells and β-galactosidase tagged EPL cells into heterologous embryoid bodies, or induction of neural differentiation in EPL cell embryoid bodies using RA, demonstrated that this restriction in developmental potential was not due to an intrinsic deficiency in the EPL cells' ability to form visceral endoderm or neurons respectively. Rather, an altered differentiation environment in which visceral endoderm and neuron formation require inductive signals from earlier pluripotent cell types which are not present in the EPL cell embryoid body appears likely (Lake *et al.*, 2000).

This evidence suggests that EPL cells represent an alternate pluripotent cell state *in vitro*. Maintenance of EPL cells requires the presence of MEDII. Removal of MEDII in the absence of LIF results in spontaneous differentiation, whilst removal of MEDII in the presence of LIF reverts the cells to an ES cell-like phenotype as assessed by morphology, gene expression and differentiation potential *in vitro* and *in vivo*, following reintroduction into host blastocysts (Rathjen *et al.,* 1999; Lake *et al.*, 2000). This interconvertability of ES and EPL cell states is reminiscent of that observed following isolation of pluripotent cells from PGCs, which adopt an

ES cell-like morphology in culture (EG cells) (Matsui *et al.,* 1992). The potential experimental advantage of pluripotent cell lines representative of alternate states *in vivo* becomes evident following experiments utilising stable, semi-differentiated intermediates of haematopoietic stem cells which have allowed cytokines which regulate their differentiation *in vitro* (Metcalf, 1989) and very often *in vivo* (Metcalf, 1991) to be identified and examined.

The formation of EPL cells from ES cells *in vitro* mimics the ICM to primitive ectoderm conversion *in vivo,* in terms of both gene expression and differentiation potential, and offers several advantages for investigating pluripotent cell developmental progression. The use of biologically derived factors which induce a homogeneous population of EPL cells provides a more easily controllable and manageable system, with respect to experimental manipulation, free of contaminating extra-embryonic lineages which are known to have inductive capabilities. Since the ES to EPL cell transition mimics the gene expression changes which occur during the formation of primitive ectoderm from the ICM, this model is likely to provide a powerful mechanism for identifying genes which are differentially expressed within the predicted pluripotent cell sub-populations that exist between the ICM and primitive ectoderm.

1.6 A SYSTEM FOR THE GENERATION OF PLURIPOTENT CELL MARKER GENES

The haematopoietic system has benefited greatly from the identification of numerous genes and cell surface markers which are differentially expressed within distinct lineages of the haematopoietic system and has led to the functional dissection of stem cell renewal and differentiation during haematopoiesis (Brady *et al.,* 1995). In a similar fashion, several lines of evidence from investigation of pluripotent cell developmental potential (1.4.1), gene expression (1.4.2) and null mutants (1.4.3) have suggested that pluripotent cells of early mouse development may undergo gradual developmental progression via transient, intermediate sub-populations. Therefore, generation and expression mapping of marker genes for pluripotent cell populations *in vivo* could allow the identification of the proposed pluripotent cell sub-populations during early mouse development, and provide the first molecular evidence of their existence. Furthermore, the gene products of the identified markers may be functionally important in developmental processes within these sub-populations.

1.6.1 Differential Message Display PCR

Advances in molecular biology techniques have resulted in mRNA expression analysis becoming an accessible and efficient route to marker gene identification. However, difficulties arise when attempting to modify the use of techniques to early embryonic cell types. With the exception of one report (Harrison *et al.*, 1995) in which subtractive hybridisation of libraries constructed from the three primary germ layers was used to isolate differentially expressed genes, an indepth study of differential gene expression has not been undertaken in early mouse development. Homology based screening has also been limited due to the unique mammalian development strategy during the peri-implantation period. Expression screening approaches using embryonic tissue are difficult due to problems associated with obtaining sufficient starting material free of contaminating extra-embryonic tissues, as embryos are small and fragile and dissections at this stage are intricate and time consuming. Furthermore, such approaches require validation of potential markers by expression *in vivo*, which is difficult during the peri-implantation stage of development. However, differential hybridisation screening of *in vitro* pluripotent cell lines, such as the F9 EC cell line, compared with a library constructed from F9 cells differentiated by RA was utilised to isolate the *Rex1* gene (Hosler *et al.*, 1989), and allowed *in vitro* validation of potential markers. These difficulties explain the paucity of pluripotent cell markers expressed transiently during early mouse development.

Differential message display polymerase chain reaction (dd-PCR) is a method which allows simultaneous comparison of mRNA transcripts from related cell types (Liang and Pardee, 1992). Since most of the displayed transcripts are common, subtle molecular differences between the alternate cell populations can be identified. This method when combined with the gradual progression of ES to EPL cells *in vitro* provides an assay system for the identification of genes which have the potential to be differentially expressed during the ICM to primitive ectoderm transition *in vivo*. Therefore, dd-PCR of ES cell RNA, various passage EPL cell RNA and differentiated ES cell RNA was carried out to identify candidate marker genes for the recognition of different pluripotent cell populations.

1.6.2 dd-PCR of ES and EPL Cell RNA

To compare gene expression in ES, EPL and differentiated ES cells both modified dd-PCR (Schulz, 1996) and conventional dd-PCR (S. Sharma, unpublished) were used. The modified method (id-PCR) involved the utilisation of a degenerate 3' primer for reverse transcription and PCR, replacing the anchored 3' oligo d(T) primer. Although this modification decreased the number of transcripts analysed by each combination of primers, it also reduced spurious background amplification and enhanced specific reamplification and cloning. Importantly, when fragments were primed from within transcript open reading frames (ORF), rapid identification of clones by sequencing was often achievable, a limiting feature of 3' untranslated region (UTR) amplification resulting from application of the original method (Schulz, 1996).

In a screen for transcripts differentially expressed between ES and EPL cells, idPCR using 40 PCR primer combinations was performed (Schulz, 1996). Individual products were identified with patterns consistent with ES cell restricted expression with the expression profile of one product, termed *Psc1* (peri-implantation stem cell 1), indicated in Figure 1.11 A. *Psc1* was expressed by ES cells and EPL cells grown for 6 days in culture, and was downregulated in later passage EPL cells grown for 8 days in culture. Downregulated expression was also seen in ES cell cultures differentiated in the presence of RA, with the low level expression detected reflecting the presence of residual undifferentiated stem cells within the RA treated culture as analysed by *in situ* hybridisation (Schulz, 1996).

Following extensive idPCR analysis of ES and various passage EPL cell RNA, no EPL cell specific products were identified. In an effort to isolate late passage EPL cell specific products the original dd-PCR technique was carried out using an anchored 3' oligo d(T) primer along with degenerate 5' 10 mers (S. Sharma, unpublished). After the use of only two different degenerate 5' 10 mers in combination with the oligo d(T) 3' primer, a number of ES and EPL cell specific products were identified. The expression profiles of two products, *L17* and *K7*, are shown in Figure 1.11 B. *L17* was expressed by ES cells, and EPL cells following 2 days of culture in the presence or absence of LIF and 4 days in culture in the presence of LIF, but not by EPL cells following culture for 4 days in the absence of LIF, or 6 days in the presence or absence of LIF. Low level *K7* expression was detected in ES cells and EPL cells grown for 2 days in the presence or absence of LIF and 4 days in the presence of LIF, but elevated expression was detected in EPL

Figure 1.11: Differential display PCR analysis of ES and EPL cells

A) 2 μg of cytoplasmic RNA, isolated from ES cells, EPL cells grown for 4-6 days in the presence of LIF (+), and ES cells differentiated with retinoic acid (RA) was reverse transcribed with an internal 3'-hp primer, before dilution 1/20 for dd-PCR, incorporating (^{33}P) dATP, using OPB-04 5' primer (Operon technologies inc.) and the 3'-hp primer. Reactions were resolved by electrophoresis on a 6% (v/v) denaturing polyacrylamide gel prepared from 40% sequagel (Schulz, 1996). The *Psc1* band is indicated.

5' OPB-04 primer: GGACTGGAGT

3' hp primer: TAGAATTCCGTNCGTAGTTCTTGAGAACCA

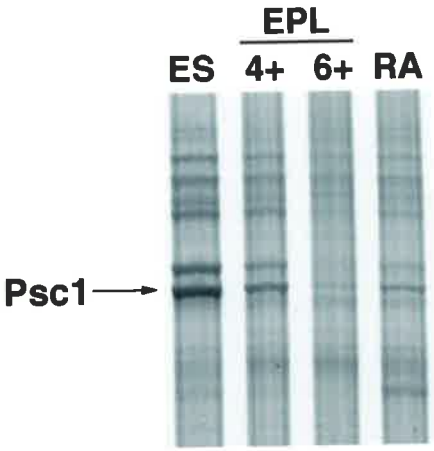
(N = A, C, G or T)

B) 2 μg of cytoplasmic RNA isolated from ES cells, EPL cells grown for 2-6 days in the presence (+) or absence (-) of LIF, and spontaneously differentiated ES cells (S), was reverse transcribed with a 3'-oligo d(T) primer, before dilution 1/20 for dd-PCR, incorporating (^{33}P) dATP, using OPA-01 5' primer (Operon technologies inc.) and the 3'-oligo d(T) primer. Reactions were resolved by electrophoresis on a 6% (v/v) denaturing polyacrylamide gel prepared from 40% sequagel (S Sharma, unpublished). The *L17* and *K7* bands are indicated.

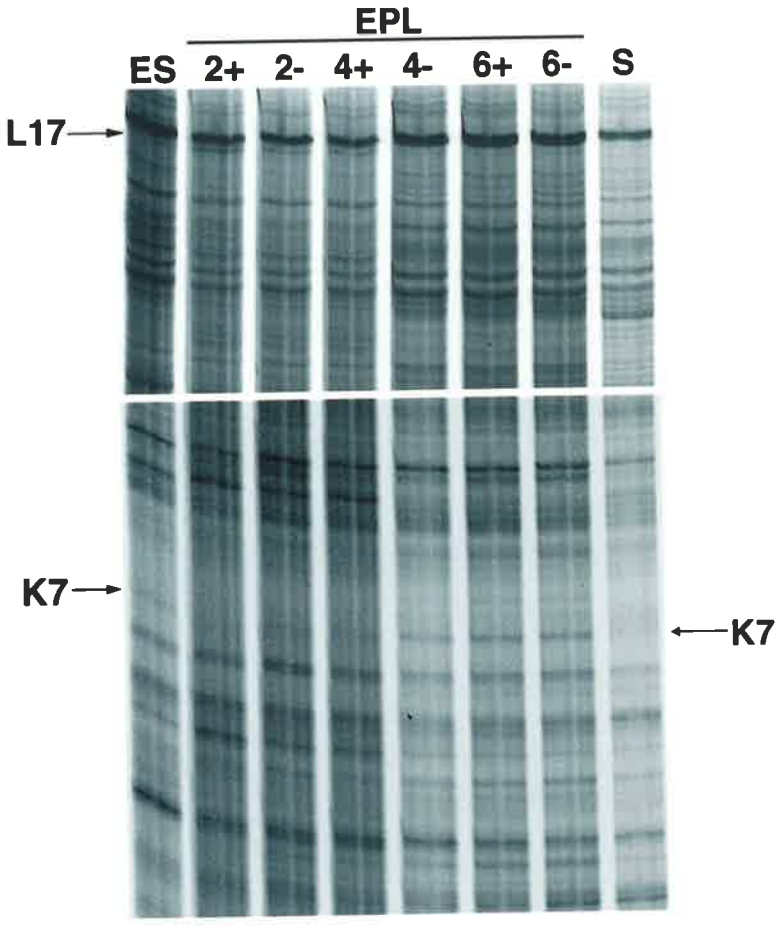
5' OPA-01 primer: CAGGCCCTTC

3' oligo d(T) primer: TTTTTTTTTTTCA

A



B



cells following culture for 4 days in the absence of LIF or 6 days in the presence or absence of LIF (S. Sharma, unpublished).

An inherent problem with dd-PCR arises in the isolation of the cDNA fragments which recapitulate the original dd-PCR profile following excision, reamplification from dried acrylamide gels and cloning of the fragments. Prospective clones were screened by Northern analysis of ES cell and various passage EPL cell RNA, in comparison to controls *Oct4* and/or *mGap*. Correct *Psc1* clones demonstrated high ES cell expression, with expression downregulating following EPL cell formation, with *Psc1* transcripts barely detectable following 8 days in culture in the presence or absence of LIF (Schulz, 1996) (Figure 1.12 A; Figure 1.13 A). Correct clones for *L17* demonstrated highest expression in ES cells, with expression downregulated following formation of EPL cells. Further downregulation was observed over a succession of passages in MEDII in the presence of LIF every 2 days for up to 6 days (S. Sharma, unpublished; Rodda, 1998) (Figure 1.12 B; Figure 1.13 B). *K7* however demonstrated a different expression pattern, with expression detected at low levels in ES cells and EPL cells grown for 2 days in culture, but upregulated in EPL cells grown for 4-8 days in culture in the presence of LIF (S. Sharma, unpublished) (Figure 1.12 C; Figure 1.13 C).

Digoxygenin *in situ* hybridisation to ES and/or EPL cell cultures was performed to determine that expression of *Psc1*, *L17* and *K7* was specific for pluripotent cells and not a consequence of sporadic differentiation within the culture. Antisense probes specifically detected expression of these markers in undifferentiated cells (Schulz, 1996; S. Sharma, unpublished). Representative *in situ* hybridisation patterns for *Psc1*, *L17* and *K7* expression in ES cells and EPL cells grown for 4 days in culture are illustrated in Figure 1.14 (Schulz, 1996; S. Sharma, unpublished).

1.6.3 Sequence of *L17*, *Psc1* and *K7* Markers

The *L17*, *Psc1* and *K7* partial cDNAs isolated were sequenced, and consisted of 736 bp, 458 bp and 232 bp respectively (Schulz, 1996; S. Sharma, unpublished). Data base searches of *L17* and *Psc1* revealed that neither sequence displayed any significant homology at the DNA level to any known genes. Data base searches using the *K7* sequence revealed significant homology to a human cDNA for the *KIAA0165* gene of unknown function, isolated from the human immature

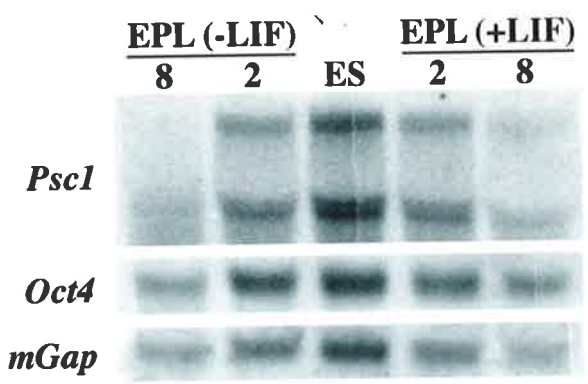
Figure 1.12: Northern blot analysis of *Psc1*, *L17* and *K7* in ES and EPL cell cultures

A) 10 μg of poly-$(A)^+$ RNA, isolated from ES cells and EPL cells grown for 2 and 8 days in culture in the presence (+LIF) or absence (-LIF) of LIF, was electrophoresed on a 1.3% agarose gel and transferred to nylon membrane. The filter was hybridised with a 458 bp *Psc1* DNA probe (2.3.23). The expression of each gene was determine by phosphorimager analysis following exposure of filters overnight in phosphorimager cassettes. The filter was subsequently stripped, re-probed as described above, with *Oct4* for quantitation against the number of pluripotent cells present in the culture, and then *mGap* as a general loading control (Schulz, 1996). Transcript sizes were 5.5, 3.7 and 3.5 kb for *Psc1*, 1.6 kb for *Oct4* and 1.5 kb for *mGap*.

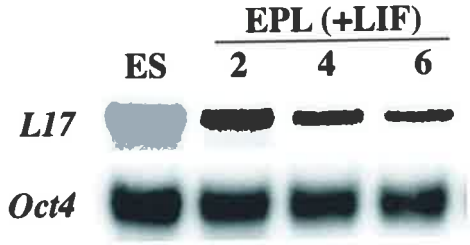
B) 5 μg of poly-$(A)^+$ RNA isolated from ES cells and EPL cells grown for 2-6 days in culture in the presence of LIF (+LIF), was electrophoresed on a 1% agarose gel and transferred to nylon membrane. The filter was hybridised overnight with a 737 bp *L17* DNA probe (2.3.23) prior to washing and overnight exposure in a phosphorimager cassette. The filter was subsequently stripped and re-probed with *Oct4* as described above, for quantitation of *L17* expression against the pluripotent cells present in the culture (S Sharma, unpublished). The transcript size for *L17* was approximately 9 kb.

C) 25 μg of cytoplasmic RNA isolated from ES cells and EPL cells grown for 2-8 days in culture in the presence of LIF (+LIF), was electrophoresed on a 1% agarose gel and transferred to nylon membrane. The filter was hybridised overnight with a 232 bp *K7* DNA probe (2.3.23) prior to washing and exposure for 3 days in a phosphorimager cassette. The filter was subsequently stripped and re-probed with *mGap* as a general loading control, as described above (S Sharma, unpublished). Transcript size for *K7* was approximately 7 kb.

A


EPL (-LIF)
EPL (+LIF)
8
2
ES
2
8
Psc1
Oct4
mGap

B


EPL (+LIF)
ES
2
4
6
L17
Oct4

C

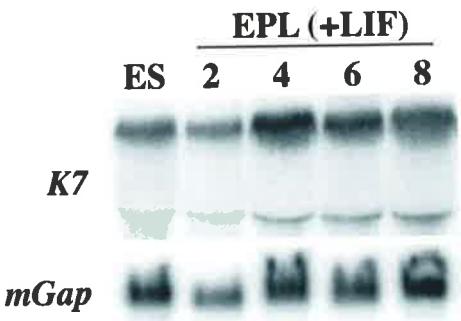
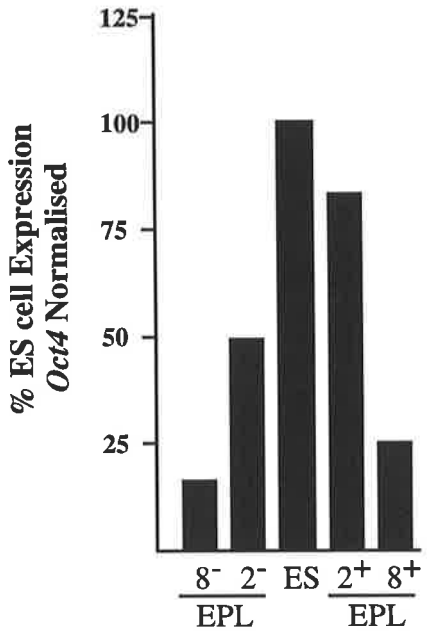

EPL (+LIF)
ES
2
4
6
8
K7
mGap

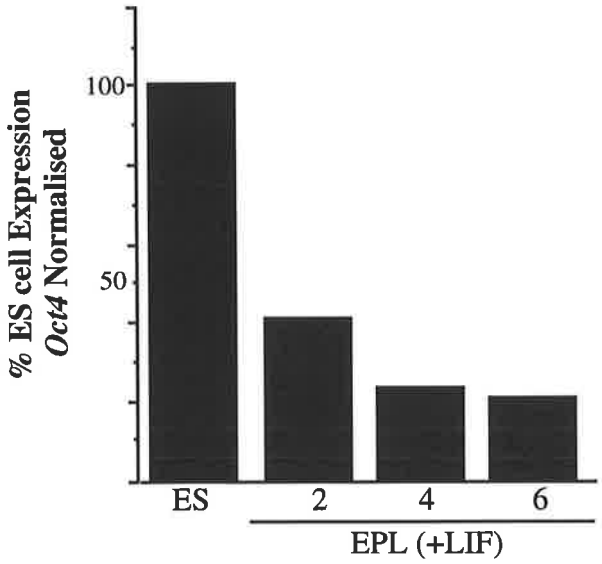
Figure 1.13: Quantitation of *Psc1*, *L17* and *K7* northern analysis

Graphical representation of *Psc1* (**A**), *L17* (**B**) and *K7* (**C**) expression in ES and EPL cells, as analysed by northern blot (Figure 1.12). The expression levels of each gene were quantitated by volume integration and normalised against *Oct4* expression for *Psc1* and *L17,* and *mGap* expression in the case of *K7*. Expression of each transcript in ES cells was designated 100%, and the expression level of each gene in EPL cell culture depicted as a percentage of ES cell expression.

A


% ES cell Expression
Oct4 Normalised
125
100
75
50
25
8⁻ 2⁻ ES 2⁺ 8⁺
EPL
EPL

B


% ES cell Expression
Oct4 Normalised
100
50
ES 2 4 6
EPL (+LIF)

C

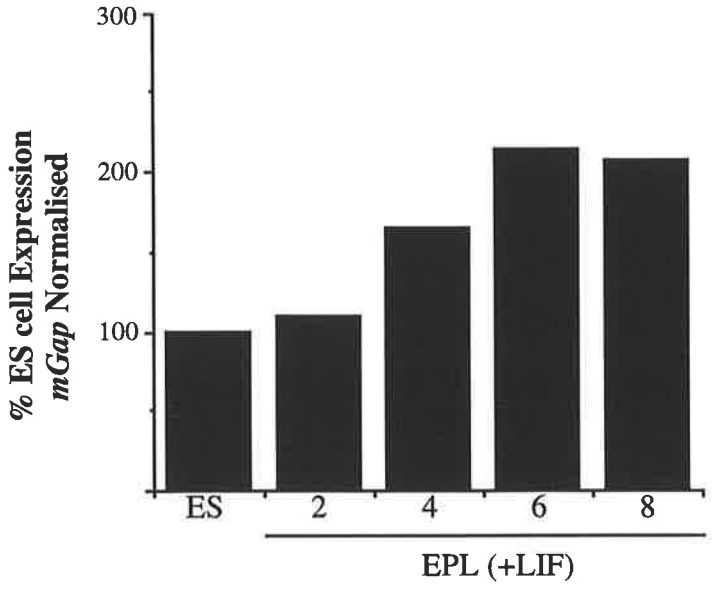
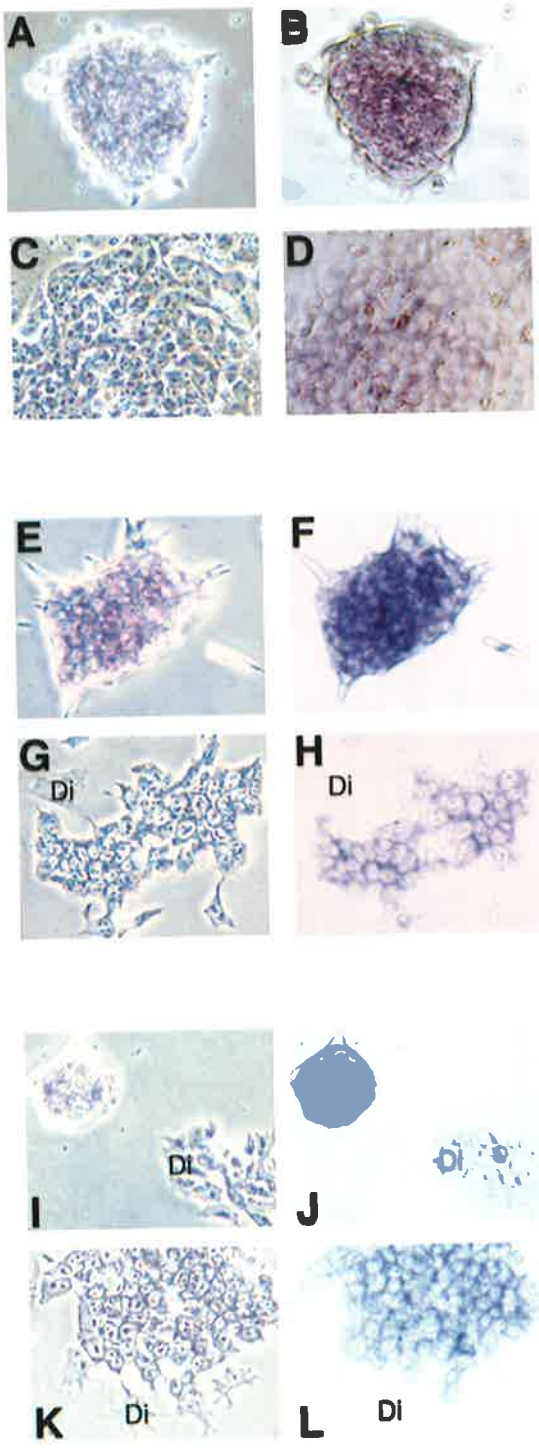

% ES cell Expression
mGap Normalised
300
200
100
ES 2 4 6 8
EPL (+LIF)

Figure 1.14: Localisation of marker gene expression to pluripotent cell

In situ hybridisation of ES cells, and EPL cells grown for 4 days in the presence of LIF. *Psc1* expression: **A, B, C** and **D**; *L17* expression: **E, F, G** and **H**; *K7* expression: **I, J, K** and **L**. Phase contrast microscopy: **A, C, E, G, I** and **K**; Bright field microscopy: **B, D, F, H, J** and **L**.

Magnification 100X

Di = Differentiated cells


A
B
C
D
E
F
G
Di
H
Di
Di
I
J
Di
K
Di
L
Di

myeloid cell line, KG-1 (Nagase *et al.,* 1996). Two regions of *K7* sequence showed homology to *KIAA0165* at 88% and 90% respectively, from position 10-39 bp and 141-185 bp.

No potential ORFs were present in the *L17* or *K7* cDNAs, presumably reflecting the expected origin of these sequences within the 3' UTR region. However *Psc1,* which was isolated using an internal degenerate 3' primer, contained a region of similarity to a predicted *C. elegans* ORF from the cosmid BO336 (Schulz, 1996). A 96 amino acid sequence in the *Psc1* ORF, spanning positions 25-120, contained 44.8 % identity to positions 348-443 of the predicted CELB0336.3 10 *C. elegans* ORF (Wilson *et al.,* 1994; Accession number U32305).

1.7 AIMS AND APPROACH

The availability of marker genes which allow unequivocal discrimination between different pluripotent cell populations would have important implications for our understanding of early embryogenesis. In particular, identification of distinct pluripotent cell sub-populations and correlation with embryonic development may allow the molecular analysis of pluripotent cell developmental progression. Ultimately, marker genes may allow the relationship between pluripotent cell sub-populations *in vivo* and *in vitro* pluripotent cell populations, such as ES and EPL cells, to be established, allowing more precise definition of their embryonic equivalents. Therefore, the initial aim of this project was the investigation of pluripotent cell progression *in vivo* via detailed mapping of pluripotent cell marker gene expression during the ICM to primitive ectoderm transition. The temporal and spatial boundaries of *Oct4*, *Rex1*, *Fgf5*, *L17*, *Psc1* and *K7* were identified.

Furthermore, the gene products of marker genes may be functionally important in developmental processes within these sub-populations. *K7* demonstrated a particularly interesting expression pattern *in vivo*, with expression detected in a short period of early mouse development from approximately day 4.75 to day 5.25. This time period was proposed to be coincident with a number of cellular events, including the commencement of more rapid pluripotent cell proliferation, proamniotic cavitation and reorganisation of apolar epiblast cells into a polarised, pseudostratified epithelium of primitive ectoderm. *K7* was therefore analysed further for possible functions by cDNA screening, homology searches and further expression analysis.

CHAPTER 2:

MATERIALS AND METHODS

2.1 ABBREVIATIONS

Abbreviations are as described in "Instructions to authors" (1978) *Biochem. J.* **169**:1-27. Additional abbreviations are as follows:

AP	alkaline phosphatase
APS	ammonium persulphate
bp	base pair
BCIG	5-bromo-4-chloro-3-indolyl-beta-D-galactosidase
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BSA	bovine serum albumin
CHAPS	3-[(Cholamidopropyl)dimethlyammonio]-1-propane-sulphonate
CIP	calf intestinal phosphatase
Ci	curie
c.p.m.	counts per minute
d.p.c.	days *post coitum*
DIG	digoxygenin
DMEM	dulbecco's modified eagle medium
DMF	dimethylformamide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTPs	deoxyribonucleotide triphosphates
DOC	sodium deoxycholate
DTE	dithioerythritol
DTT	dithiothreitol
E. coli	Escherichia coli
EGFP	enhanced green flourescent protein
EtBr	ethidium bromide
EDTA	ethylene diamine tetra acetic acid
FD	Farraday
FCS	foetal calf serum
FITC	fluorescein isothiocynate

FLB	formamide load buffer
GLB	gel loading buffer
Hyb	hybridisation buffer
HEPES	N-2-hydroxyethyl piperazine-N-ethane sulphonic acid
HRP	horse radish peroxidase
IMDM	Iscove's modified dulbecco's medium
IP	immunoprecipitation
IPTG	isopropyl-β-D-thiogalactopyranoside
kDa	kilodalton
LB	luria broth
M	molar
mA	milliamperes
min.	minute(s)
mM	millimolar
MOPS	3-[N-morpholino]propane sulphonic acid
Mr.	relative molecular weight
MQ H_20	reverse osmosis filtered water passed through a Milli-Q™ ion-exchange matrix
NBT	nitro blue tetrazolium chloride
NP-40	nonidet-P 40
OD_{600}	optical density at 600 nanometres
O/N	overnight
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBT	phosphate buffered saline + 0.1% Tween-20
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFA	paraformaldehyde
PIPES	piperazine-N,N'-bis'(2-ethanesulfonic acid)
PMSF	phenyl methyl sulfonyl chloride

PSB	phage storage buffer
RNA	ribonucleic acid
RNAse	ribonuclease
RNAsin	RNAse inhibitor
r.p.m.	revolutions per minute
rNTP	ribonucleotide triphosphate
RT	room temperature
SD Buffer	standard digest buffer
SDS	sodium dodecyl sulphate
sec.	second
ssDNA	sheared herring sperm DNA
TBS	tris buffered saline
TBST	tris buffered saline + Tween-20
TEMED	N, N, N', N'-teramethyl-ethenediamine
TRITC	Tetramethylrhodamine B isothiocyanate
tRNA	transfer RNA
Tween-20	polyoxyethylenesorbitan monolaurate
U	unit(s)
UV	ultra violet
V	volts
W	watts

2.2 MOLECULAR MATERIALS

2.2.1 Chemicals and Reagents

Chemicals and reagents used during the course of this work were of analytical grade and supplied as follows:

Sigma: agarose (Type 1), ampicillin, BSA, CHAPS, coomassie brilliant blue, DOC, EtBr, EDTA, heparin, kanamycin, MOPS, PMSF, rATP, SDS, TEMED, Tris base, Hoechst-33258, PIPES, propidium iodide

Roche: BCIG, BCIP, 10X DIG labelling mix, DTE, DTT, glycogen, ssDNA, IPTG, NBT, dNTPs, tRNA (from brewers yeast), and 10X transcription buffer, complete™ protease inhibitor tablets

BDH Chemicals: APS, DMF, NP-40, phenol, and PEG 6000

Difco: Bacto-tryptone, yeast extract, Bacto-agar, yeast nitrogen base, casamino acids

Merck: PFA

Pharmacia: Oligo d(t)-cellulose Type 7, Sepharose CL-6B, glutathione sepharose 4B, reduced glutathione, protein-G sepharose

ICN: Gluteraldehyde, Tween-20

National Diagnostics: Sequagel 6 and protogel

Pierce: SuperSignal™ Substrate

2.2.2 Radiochemicals

[α-^{32}P] dATP (3000 Ci/mmol), [α-^{33}P] dATP (1500 Ci/mmol), and [α-^{32}P] UTP (3000 Ci/mmol) were supplied by Geneworks Ltd.

2.2.3 Kits

Bresaclean kit: Geneworks Ltd.

BRESAspin plasmid mini kit: Geneworks Ltd.

Gigaprime DNA labelling kit: Geneworks Ltd.

QIAfilter plasmid midi kit: Qiagen

2.2.4 Enzymes

Restriction endonucleases were supplied by Pharmacia and New England Biolabs. Other enzymes were obtained from the following sources:

Roche: CIP, DNAse 1, RNAse T1, T7 and T3 RNA polymerases

Geneworks Ltd: *E. coli* DNA polymerase I (Klenow fragment), Taq DNA Polymerase, RNAsin, T4 DNA ligase, and SP6 RNA polymerase

Sigma: RNAse A

Gibco/BRL: Superscript II reverse transcriptase

Merck: Proteinase K

2.2.5 Buffers and Solutions

Acetate solution: 3M KAc, 2M HOAc, pH 5.8

AP buffer: 100 mM NaCl, 50 mM $MgCl_2$, 100 mM Tris HCl pH 9.5, 0.1% (v/v) Tween-20

Bacterial lysis buffer: 3% (w/v) SDS, 5% (v/v) β-mercaptoethanol

10X CIP buffer: 500 mM Tris HCl pH 8.5, 10 mM EDTA (stored at 4°C)

Coomassie Destain: 45% (v/v) methanol, 45% (v/v) MQ H_20, 10% (v/v) acetic acid

Coomassie stain: 25 5 (w/v) coomassie brilliant blue, dissolved in destain

Denaturing Solution: 0.5 M NaOH, 1.5 M NaCl

Denhardt's solution: 0.1% (w/v) Ficoll, 0.1% (w/v) polyvinylpyrolidone, 0.1% (w/v) BSA

Dissection Buffer: 80% ES DMEM, 20% FCS, 10 mM HEPES pH 7.4

Elution Buffer: 500 mM NH_4Ac, 10 mM MgCl2, 0.1% (w/v) SDS

FLB: 95% (v/v) deionised formamide, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol

10X GLB: 50% (v/v) glycerol, 0.1% (w/v) SDS, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol

5X Hyb: 2 M NaCl, 200 mM PIPES pH 6.4, 5 mM EDTA

Pre-hyb solution: 5.85g NaCl, 40% (v/v) deionised formamide, 50 mM Tris HCl pH 7.4, 16.5% (v/v) PEG, 1% (v/v) SDS, 5% (v/v) Denhardts solution (2.2.5), 100 µg/ml ssDNA

IP Lysis Buffer: 50 mM Tris-HCl pH7.5, 150 mM NaCl, 10% (v/v) glycerol, 1% Triton X-100, 10 mM EDTA, 200 µM Na-orthovanadate, 1 complete™ protease inhibitor tablet

In situ Buffer 1: 50% formamide, 5X SSC, 0.1% Tween-20, 0.5% (w/v) CHAPS

In situ Hyb: 50% (v/v) deionised formamide, 5X SSC pH 4.5, 50 µg/ml heparin, 0.1% (v/v) Tween-20, 100 µg/ml tRNA, 100 µg/ml ssDNA

In situ substrate mix: 1 ml AP buffer, 4.5 µl NBT (75 mg/ml in 70% (v/v) DMF), 3.5 µl BCIP (50 mg/ml in 100% (v/v) DMF)

10X Klenow buffer: 500 mM Tris HCl pH 7.6, 100 mM $MgCl_2$

5X Ligation Buffer: 250 mM Tris HCl pH 7.5, 25% (w/v) PEG 6000, 50 mM $MgCl_2$, 5 mM rATP, 5 mM DTT

10X Ligation Buffer: 500 mM Tris HCl pH 7.5, 100 mM $MgCl_2$, 10 mM rATP, 100 mM DTT

Lysis solution: 0.2 M NaOH, 1% (v/v) SDS

10X MOPS: 23 mM MOPS pH 7.0, 50 mM NaAc, 10 mM EDTA

Na-TES: 0.5 M NaCl, 10 mM Tris HCl pH 7.5, 1 mM EDTA, 0.1% (v/v) SDS

Neutralising solution: 0.5 M Tris HCl pH 8.0, 1.5 M NaCl

Plating solution: 50% (w/v) PEG 6000, 10% (w/v) lithium acetate, 1% (w/v) 1 M Tris HCl pH 7.5, 0.2% (w/v) 0.5 M EDTA

NP-40 Lysis Buffer: 50 mM Tris HCl pH 7, 150 mM NaCl, 1% (v/v) NP-40, 1 mM PMSF, 1 mM EDTA

PBS: 136 mM NaCl, 2.6 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4` pH 7.4.

PSB: 10 mM Tris HCl pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 0.05% (w/v) gelatin

RNAse digestion buffer: 300 mM NaCl, 10 mM Tris HCl pH 7.5, 5 mM EDTA, 40 µg/ml RNAse A, 2 µg/ml RNAse T1

PBT: PBS + 0.1% (v/v) Tween-20

RIPA: 150 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl pH 8.0, 1% (v/v) NP-40, 0.5% (v/v) DOC, 0.1% (v/v) SDS

10X SD buffer: 330 mM Tris-HAc pH 7.8, 625 mM KAc, 100 mM MgAc, 40 mM spermidine, 5 mM DTE

2X SDS load buffer: 125 mM Tris HCl pH 6.8, 4% (v/v) SDS, 20% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 5% (v/v) β-mercaptoethanol

SDS-PAGE buffer: 25 mM Tris-Glycine, 0.1% (w/v) SDS

Solution D: 4 M guanidinium thiocyanate, 25 mM sodium citrate pH 7, 0.5% (v/v) sarcosyl, 0.1 M β-mercaptoethanol

Solution 1: 25 mM Tris HCl pH 8, 10 mM EDTA pH 8, 15% (w/v) sucrose

SSC: 150 mM NaCl, 15 mM sodium citrate, pH 7.4

STET: 50 mM Tris HCl pH 8, 50 mM EDTA, 8% (w/v) sucrose, 5% (v/v) Triton-X-100

TAE: 40 mM Tris-acetate, 20 mM NaAc, 1 mM EDTA, pH 8.2

1X Taq Buffer: 67 mM Tris-HCl pH 8.8, 16.6 mM $(NH_4)_2SO_4$, 0.2 mg/ml gelatin, 0.45% (v/v) Triton X-100

TBE: 90 mM Tris, 90 mM boric acid, 2.5 mM EDTA, pH 8.3

TBS: 25 mM Tris HCl pH 8, 150 mM NaCl

TBST: TBS + 0.1% (v/v) Tween-20

TE: 10 mM Tris HCl pH 7.5, 1mM EDTA

TEN Buffer: 40 mM Tris HCl pH 7.4, 1 mM EDTA, 150 mM NaCl

TES: 10 mM Tris HCl pH 7.5, 1 mM EDTA, 0.1% (v/v) SDS

Tfb 1: 30 mM KAc, 100 mM RbCl, 10 mM $CaCl_2$, 15% (v/v) Glycerol, pH 5.8 (adjusted with 0.2 M Acetic Acid)

Tfb 2: 10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% (v/v) Glycerol, pH 6.5 (adjusted with 1 M KOH)

TNM: 30 mM Tris HCl pH 7.6, 150 mM NaCl, 15 mM $MgCl_2$, 0.4% (v/v) NP40

Transcription buffer: 40 mM Tris HCl pH 8, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT

4X Tris-SDS buffer: 1.5 M Tris HCl pH 8.8, 0.4% (w/v) SDS

TUNES: 10 mM Tris HCl pH 8.0, 7 M urea, 0.35 M NaCl, 1 mM EDTA, 2% (w/v) SDS

Western Transfer buffer: 39 mM glycine, 48 mM Tris HCl pH 8.3, 0.037% (w/v) SDS, 20% (v/v) methanol

2.2.6 Plasmids

2.2.6.1 Cloning Vectors

pBluescript II KS (Stratagene) was a kind gift from Dr. Tom Schulz

pGemT-Easy was obtained from Promega

pYES2 (Invitrogen) was a kind gift from Dr. Steve Dalton

pGex2T (Pharmacia) was a kind gift from Dr. Grant Booker

pEGFP-C2 (Clontech) was a kind gift from Dr. Murray Whitelaw

pXMT2 (Rathjen *et al.,* 1990a) was a kind gift from Dr. Bryan Haines

2.2.6.2 Cloned DNA Sequences

mGap: The mouse *glyceraldehyde phosphate dehydrogenase (mGap)* cDNA clone in pGEM3Z was a kind gift from Prof. Peter Rathjen and contained a 300 bp *Hind*III/*Pst*I fragment from the 5' end of the mouse gene (Rathjen *et al.,* 1990b).

Oct4: The *Oct4* cDNA clone in pBluescript was a kind gift from Dr. Hans Schöler and contained a 462 bp *Stu*I cDNA fragment of positions 491 to 953 of the *Oct4* cDNA sequence (Schöler *et al.,* 1990b).

Rex1: An 848 bp *Rex1* containing fragment in pCR™II was a kind gift from Dr. N. Clarke (Hosler *et al.,* 1989).

Fgf5: An 800 bp *Fgf5* full length coding region fragment in pBluescript was a kind gift from Dr. J. Hébert (Hébert *et al.,* 1991).

Nkx-2.5: A 1596 bp *Nkx2.5* containing cDNA clone in pBluescript SK$^+$ was a kind gift from Dr. Richard Harvey (Lints *et al.*, 1993).

L17 and *K7*: A 736 bp containing *L17* fragment and a 232 bp containing *K7* fragment, both from the 3' UTR regions of these genes, in pBluescript II KS were provided by Dr. Shiwani Sharma.

Psc1: A 458 bp containing *Psc1* fragment (corresponding to positions 1661-2041 of the *Psc1* cDNA) in pBluescript II KS was provided by Dr. Tom Schulz.

ESP1: A 7 kb *ESP1* containing fragment, in the Yep13 yeast expression vector was a kind gift from Dr. Breck Byers (McGrew *et al.*, 1992).

2.2.7 Oligonucleotides

DNA primers were synthesised by Geneworks Ltd.

2.2.7.1 General sequencing primers

T7: TAATACGACTCACTATAGGGAGA

T3: ATTAACCCTCACTAAAGGGA

USP: GTTTTCCCAGTCACGAC

RSP: CAGGAAACAGCTATGAC

pGEX2T 5': GGGCTGGCAAGCCACGTTTGGTG

pEGFP 5': CATGGTCCTGCTGGAGTTCGTG

2.2.7.2 Gene specific sequencing primers

7AIF: GACTATGAGGCCGTGACC

7A2R: ACTGGCAGCTGTCGTACA

93F: GAGACAAGGAGCTACAAC

H24F: CTCACAGAGGTGCTGACC

S15R: CAGATGTCTGATTCAGGA

S16F: GTCTGGAAGAGGAAGGGC

S17R: TGCAGATGCTCTGCCAGT

5' ESP1 6175: AGTTCATCTCTCATTTCC

3' K7 5702: CGTCCTACTGCCTCGCTC

2.2.7.3 PCR Primers

5' Ab1: ATGGATCCTCCAGATCCCGGACTGAT

3' Ab1: TAGAATTCAGAGCTACCACCCTCACC

5' Ab2: ATGGATCCCAGGCCCTTGCAAAGCTG

3' Ab2: TAGAATTCCTTGCTCTTGGTGGGGTG

5' *Sac*I: TACGAGCTCCATGAGGAACTTCA

3' *Xba*I: TCATCTAGACCTTGTCCTGTA

RT1: CAGAGTCTTAGTAGGTGA

RT2: CAGATCTGCAACCGTCCACT

5' *Nco*I: TATACCATGGAGCCTTGGCGTGTGGGCTG

5' *Aat*II: TATAGACGTCAGACCAAAAAGAAAACAGC

3' *Sac*I: TATAGAGCTCAGTCACACAGCTTCTACC

5' PTTG: AGTAATCCAGAATGGCTACTC

3' PTTG: GAAGTTTAAATATCTGCATCG

5' *Eco*RI PTTG: TATGAATTCAGAATGGCTACTCTTATC

3' FLAG PTTG: TTAGAATTCTCACTTGTCATCGTCGTCCTTGTAGTCAATATCTGCATC GTAACA

2.2.8 Antibodies

Antibodies, dilutions for use and suppliers were as follows:

Anti-γ-tubulin (clone GTU88) (1/10000):	Sigma
Anti-GFP polyclonal antibody (1/5000):	Gift from Dr. Pam Silver
Anti-FLAG monoclonal antibody (1/200-1/500):	Kodak IBI
Anti-DIG F_{ab} - alkaline phosphatase conjugate (1/2000):	Roche
Goat anti-rabbit IgG (whole molecule) - FITC conjugate (1/50):	Sigma
Rabbit anti-mouse IgG (whole molecule) - FITC conjugate (1/50):	Dako
Goat anti-mouse IgG (whole molecule) - TRITC conjugate (1/50):	Sigma
Goat anti-rabbit IgG (whole molecule) - HRP conjugate (1/4000):	Sigma
Goat anti-mouse IgG (whole molecule) - HRP conjugate (1/4000):	Silenus

2.2.9 Bacterial Strains and Media

All recombinant plasmids were maintained in *E. coli* strain DH5α. Library screening was carried out in *E. coli* strain BB4 (Clontech). Lambda zapping was carried out in *E. coli.* strain XL1-Blue. Protein expression from recombinant plasmids in bacteria were carried out in *E coli.* strain BL21.

DH5α: *supE*44 delta*lac* U169 (phi80 *lacZ*deltaM15) *hsdR*17 *recA*1 *endA*1*gyrA*96 *thi*-1*relA*1

BB4: *supF*58 *supE*44 *hsdR*514 *galK*2 *galT*22 *trpR*55 *metB*1 *tonA* delta*lac*U169 F'[*proAB*$^+$ *lacI*^{q} *lacZ*deltaM15 Tn*10*(*tet* ^{r})]

BL21: *hsdS gal* (λc*Its857 ind*1 *S*am7 *nin*5 *lac*UV5-T7 gene *1*)

XL1-Blue: *supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thi relA*1 *lac*$^-$ F'[*proAB*$^+$ *lacI*^{q} *lacZ* deltaM15 Tn*10*(tet^r)]

Growth media was prepared in double distilled water and sterilised by autoclaving. Antibiotics and other labile chemicals were added after the media solution had cooled to 50°C. Ampicillin (100 µg/ml) was added for growth of transformed bacteria to maintain selective pressure for all recombinant plasmids, apart from pEGFP vectors which were selected using kanamycin (25 µg/ml).

LB: 1% (w/v) Bacto-tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.0 (adjusted with NaOH).

LMM broth: 1% (w/v) Bacto-tryptone, 0.5% (w/v) NaCl, 0.4% (w/v) maltose, 0.2% $MgSO_4$.

Psi broth: 2% (w/v) Bacto-tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) $MgSO_4$, pH 7.6 (adjusted with 1 M KOH).

Solid Media: Agar plates were prepared by supplementing LB media with 1.5% (w/v) Bacto-agar and the appropriate antibiotic if required. LMM agarose for library screening was prepared by dissolving 0.7 g agarose/100 ml LMM broth.

2.2.10 Yeast Strains and Media

Wild type strain (W303-1a): *MATa*, *ura3-52*, *trp1-1*, *ade2-1*, *lys2-801*, *leu2-3*, *his3-11*, *can1-100* (*psi+*)

ESP1^{ts} strain (S252): *MATa*, esp1-1, *ura3-52*, *leu2-3*, *112 can-100*, *lys2*

Yeast growth media were prepared in double distilled water and sterilised by autoclaving. Labile components were added after the media solution had cooled to 50°C.

YEPD Media: 1.1% yeast extract, 2.2% bactopeptone, 100 µg/ml adenine, 55 µg/ml tyrosine, 2% (w/v) glucose

Ura-/Glu+ Plates: 0.8% yeast nitrogen base, 1.1% casamino acids, 100 µg/ml adenine, 55 µg/ml tyrosine, 50 µg/ml tryptophane, 50 µg/ml leucine, 2% (w/v) glucose, 2% (w/v) bacto-agar

Ura-/Gal+ Plates: 0.8% yeast nitrogen base, 1.1% casamino acids, 100 µg/ml adenine, 55 µg/ml tyrosine, 50 µg/ml tryptophane, 50 µg/ml leucine, 2% (w/v) galactose, 2% (w/v) bacto-agar

2.2.11 DNA Markers

*Hpa*II digested pUC19 markers were purchased from Geneworks Ltd.

Band sizes (bp): 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34, 26.

*Eco*RI digested SPP-1 phage markers were purchased from Geneworks Ltd.

Band sizes (kb): 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36.

2.2.12 Protein markers

SDS-7 protein molecular weight markers were purchased from Biorad.

Band sizes (Mr. in kDa): 66, 45, 36, 29, 24, 20.1, 14.2.

Benchmark™ prestained protein ladder was purchased from Gibco/BRL.

Band sizes (Mr. in kDa): 187, 118, 85, 61, 50, 38, 26, 20, 15, 9.

2.3 MOLECULAR METHODS

2.3.1 Rapid Small Scale Preparation of DNA (Mini-prep)

2 ml LB (2.2.9) containing ampicillin (100 µg/ml) was inoculated with a single transformed colony and incubated O/N at 37°C, with shaking. Cell cultures were transferred to eppendorf tubes and spun for 1 min. The bacterial pellets were resuspended in 200 µl STET buffer (2.2.5) and lysed at 100°C for 1 min. Samples were centrifuged for 15 min. and the cell debris and chromosomal DNA removed with a sterile toothpick. Plasmid DNA was precipitated by addition of 200 µl of isopropanol and incubation at -20°C for 5 min. Plasmid DNA was pelleted by centrifugation for 15 min., washed with 70% (v/v) ethanol, pellets dried and resuspended in 20 µl MQ H_2O. All eppendorf centrifugation steps were carried out in a eppendorf centrifuge (5415C).

2.3.2 Rapid Mid Scale Preparation of DNA (Midi-prep)

50 ml LB (2.2.9) plus ampicillin (100 µg/ml) was inoculated with a single transformant and incubated O/N at 37°C with shaking. Cell cultures were transferred to oakridge tubes, pelleted by centrifugation at 5000 r.p.m. for 5 min. at 4°C (SS-34 rotor, Sorvall RC-5), and resuspended in 3 ml of solution 1 (2.2.5). 6 ml of fresh lysis solution (2.2.5) was added, mixed by gentle

inversion and placed on ice for 5 min. 4.5 ml of acetate solution (2.2.5) was subsequently added, mixed by gentle inversion and placed on ice for 5 min., before a brief, vigorous mix and returning to ice for a further 15 min. The cell debris was pelleted by centrifugation at 13000 r.p.m. for 15 min. at 4°C (SS-34 rotor, Sorvall RC-5) and the supernatant decanted into a new tube. Nucleic acids were precipitated at RT for 5 min. by addition of 8 ml isopropanol. Plasmid DNA was pelleted by centrifugation at 12000 r.p.m. for 10 min. at 4°C (SS-34 rotor, Sorvall RC-5), the pellet drained and resuspended in 400 µl MQ H_2O, before being transferred to a eppendorf tube. Cellular RNA and protein were removed by the respective addition of 2 µl RNAse A (10 mg/ml) and incubation at 37°C for 30 min., followed by addition of 8 µl of 10% (w/v) SDS and 2.5 µl Proteinase K (20 mg/ml) and incubation at 37°C for 15 min. Plasmid DNA was extracted twice with an equal volume of 1:1 phenol:chloroform, once with chloroform and re-precipitated with 100 µl 7M NH_4Ac and 1 ml absolute ethanol (-20°C, 1 hour). Plasmid DNA was recovered by centrifugation, washed with 70% (v/v) ethanol, dried and resuspended in 100 µl MQ H_2O. The concentration of DNA was determined by spectrophotometric analysis at 260 nm.

2.3.3 Large Scale Plasmid Preparation for Transient Transfection Analysis

High quality plasmid DNA for transient transfection into mammalian cells was produced by two methods:

i) CsCl preparation of DNA:

500 ml LB (2.2.9) plus ampicillin (100 µg/ml) was inoculated with a single bacterial colony and incubated overnight at 37°C in an orbital shaker. The cells were harvested by centrifugation at 5000 r.p.m. for 5 min. at 4°C, (GS-3 rotor, Sorvall RC-5) and the bacterial pellets drained. The pellet was resuspended in 6.5 ml of solution 1 (2.2.5) before the addition of 13 ml of freshly made lysis solution (2.2.5). After gently mixing and incubation on ice for 5 min., 6.5 ml of acetate solution (2.2.5) was added and the solution mixed thoroughly, before incubating on ice for a further 10 min. Cell debris were pelleted by centrifugation at 15000 r.p.m. for 15 minutes at 4°C (SS-34 rotor, Sorvall RC-5). The supernatant was transferred to a new tube before precipitation with 0.6 volumes of isopropanol. Plasmid DNA was recovered by centrifugation at 10000 r.p.m. for 10 min. at 4°C (SS-34 rotor, Sorvall RC-5), the pellet drained

and gently resuspended in a solution containing 7 ml TE (2.2.5) with 7g CsCl. Once the DNA pellet was thoroughly resuspended, 700 µl EtBr (10 mg/ml) was added, and any remaining debris removed by centrifugation at 3000 r.p.m. for 10 min. (SS-34 rotor, Sorvall RC-5). The supernatant was transferred to a 10 ml Nalgene Oakridge tube and balanced with paraffin oil. After ultracentrifugation at 45000 r.p.m. for 24 hours at 20°C (Ti50 rotor, Beckman L8-M Ultracentrifuge), plasmid DNA was visualised under long wavelength UV light and recovered using a pasteur pipette. EtBr was removed by repeated extraction with NaCl/TE saturated isopropanol. The DNA solution was diluted 1 in 6 with MQ H_2O before precipitation with 2 volumes of ethanol and chilling at -20 for 45 min. Plasmid DNA was recovered by centrifugation (9000 r.p.m./30 min./4°C in a SS-34 rotor, Sorvall RC-5) and resuspended in 400 µl MQ H_2O. The DNA solution was transferred to an eppendorf tube and reprecipitated before final resuspension in MQ H_2O. Plasmid yields were calculated from the absorbance at 260 nm determined by spectrophotometric analysis.

ii) QIAfilter Plasmid Midi Kit (2.2.3):

DNA was prepared according to the manufacturers instructions.

2.3.4 Restriction Endonuclease Digestion of DNA

Plasmid DNA was digested in 1X SD buffer (2.2.5) with 4U of enzyme/µg DNA for 1-3 hours at the appropriate temperature. Complete digestion of DNA was assayed by agarose gel electrophoresis (2.3.5).

2.3.5 Agarose Gel Electrophoresis

Agarose gel electrophoresis (0.8% to 3% (w/v) agarose in TAE or TBE) was carried out using horizontal mini-gels prepared by pouring 10 ml of agarose onto a 7.5 cm x 5.0 cm glass microscope slide. Agarose mini-gels were electrophoresed in TAE or TBE (2.2.5) at 90 mA. Nucleic acid was visualised by staining gels for 5 min. with EtBr (0.25 mg/ml in water), destaining for 5 min. in MQ H_2O, and exposure to medium wavelength UV light.

2.3.6 Plasmid DNA Southern Blot

Agarose minigels for southern blot analysis were prepared as described in section 2.3.5, with gels being photographed adjacent to a ruler for scaling purposes. DNA was transferred to Hybond-N^+ membrane (Amersham) using capillary transfer. Two pieces of 3MM chromatography paper (Whatman) was pre-wetted in 0.4 M NaOH and placed over a glass plate so that the edges were submerged in 0.4 M NaOH. The gel was placed wells facing down on the damp Whatman paper (avoiding bubbles) and parafilm placed around the gel to avoid short circuiting. The membrane was pre-wetted in MQ H_20 and then 2X SSC for 5 min., and then carefully placed over the gel such that no air bubbles were trapped between the filter and the gel. Two pieces of Whatman paper, pre-wetted in 2X SSC, were then placed on top of the membrane, followed by a 2 cm stack of dry Whatman paper. A 5 cm stack of paper towels was then added and the glass plate placed on top of the paper towels. A 0.1-0.5 kg weight was placed on a glass plate and the gel transferred for at least 24 hours. Following transfer, the DNA was cross-linked to the filter by exposure to 120 mJ of UV radiation in a UV Stratalinker™ (Stratagene). The Southern filter was then placed on nylon mesh DNA side up, rolled up, and prehybridised for at least 4 hours at 42°C with rotation, in a Hybaid cylinder containing pre-hyb solution (2.2.5).

2.3.7 Purification of Linear DNA Fragments

DNA bands were excised from preparative agarose TAE gels (2.3.5), under long wave UV irradiation using a sterile scalpel blade. Linear DNA fragments greater then 250 bp were purified using the Bresaclean kit (2.2.3) according to the manufacturer's instructions. Small DNA fragments between 100-250 bp were eluted in 400 µl of elution buffer (2.2.5) at 37°C O/N, before ethanol precipitation at -20°C for at least 1 hour by the addition of 1/10th volume of 3M NaAc pH 5.2, 2 volumes of absolute ethanol and 1 µl of glycogen.

2.3.8 Blunting DNA Fragments

5' overhangs on DNA fragments were filled in by incubation at 37°C for 30 min. in 0.2 mM dNTPs, 1X klenow buffer (2.2.5) and 6 U of klenow enzyme. 3' overhangs were removed from DNA fragments by incubation in klenow buffer and 6 U of klenow enzyme for 5 min. at

37°C, before the addition of 0.4 mM dNTPs and incubation at 37°C for a further 20 min. Fragments were subsequently gel purified (2.3.7) in readiness for blunt cloning.

2.3.9 Preparation of Vector DNA

5 µg midi-prep DNA (2.3.2) was digested using appropriate enzyme(s) in 1X SD buffer as described in section 2.3.4. Vectors were ethanol precipitated and resuspended in 45 µl of MQ H_2O before the addition of 1 U of CIP, 5 µl 10X CIP buffer (2.2.5) and incubation at 37°C for 30 min. Vectors were isolated and purified as described previously (2.3.7).

2.3.10 DNA Ligation Reactions

Ligation reactions contained 25-50 ng purified vector, 75-150 ng DNA insert, the appropriate ligation buffer (10X ligation buffer for complementary end ligations and 5X ligation buffer for blunt end ligations; 2.2.5) and 1 U T4 DNA ligase. Reactions were incubated at RT for 1-4 hours for complementary end ligations and 4-16 hours at 16°C for blunt end ligations.

2.3.11 Preparation of Competent Cells

A single *E. coli* colony was used to inoculate 2 ml Psi broth (2.2.9) and grown in an orbital shaker O/N at 37°C. 0.5 ml of the O/N was subcultured into 15 ml Psi broth (2.2.9) and grown at 37°C until the culture reached an OD_{600} of 0.6. 5 ml of the subculture was then added to 95 ml Psi broth (2.2.9) and grown at 37°C until the OD_{600} was between 0.5 and 0.6. The bacterial cells were chilled on ice for five min. and harvested by centrifugation at 6000 r.p.m. for 5 min. at 4°C (SS-34 rotor, Sorvall RC-5). The cells were resuspended in 40 ml of Tfb 1 (2.2.5) and chilled on ice for a further 5 min. The cells were re-harvested by centrifugation and resuspended in 4 ml of Tfb 2 (2.2.5). After incubation on ice for 15 min., 50 µl aliquot's of cell suspension were transferred to eppendorf tubes and snap frozen in a dry ice/ethanol bath. The cells were stored at -80°C until required.

2.3.12 Transformation of Competent Bacterial Cells

Approximately 10 ng of ligated DNA (2.3.10) was added to 50 µl of thawed competent cells and placed on ice for 20 min. The cells were heat shocked at 42°C for 2 min., incubated on

ice for 10 min., and then 900 µl of LB added for growth at 37°C for 45 min. while shaking. The cells were pelleted by centrifugation and all but approximately 100 µl of supernatant decanted. The cells were resuspended in this residual supernatant, spread onto LB ampicillin (100 µg/ml) plates (2.2.9) for all ligations, apart from those using pEGFP vector. In this case selective pressure was obtained by growth on LB kanamycin (25 µg/ml) plates. Transformant colonies were grown O/N at 37°C. If blue/white colour selection was required, plates were prepared by spreading with 50 µl IPTG (50 µg/ml) and 20 µl BCIG (50 µg/ml in DMF) at least 15 min. prior to bacterial spreading.

2.3.13 Sequencing of Plasmid DNA

Plasmid DNA was prepared for cycle sequencing using the BRESAspin plasmid mini kit (2.2.3) according to the manufacturers instructions. PCR products were prepared for cycle sequencing by gel purification (2.2.7). Sequencing reactions were carried out in 0.5 ml PCR tubes containing 8 µl of either Dye Terminator reaction mix or Big Dye reaction mix (Perkin-Elmer), 100 ng of primer and 1 µg of plasmid DNA, or 150 ng of gel purified PCR product, and made up to 20 µl with MQ H_2O. PCR sequencing reactions were carried out in a Programmable Thermal Controller (PCT-100) PCR machine (MJ Research Inc.), using PCR conditions of denaturation at 96°C for 30 sec., annealing at 50°C for 15 sec., and extension at 60°C for 4 min. for a total of 25 cycles. The reaction mixes were transferred to eppendorf tubes and precipitated with 1/10th volume of 3M NaAc pH 4.6 and 2 volumes of absolute ethanol at -80°C for 30 min. Tubes were then spun at 12000 r.p.m. for 20 min., washed with 70% (v/v) ethanol, and dried. Reactions were analysed on an ABI PRISM 277 DNA sequencer (Perkin-Elmer).

2.3.14 Library Screening

2.3.14.1 Preparation of plating bacteria

10 ml LMM broth (2.2.9) was inoculated with a single bacterial colony and grown O/N at 37°C in an orbital shaker. Bacterial cultures were centrifuged at 5000 r.p.m. for 5 min. at 4°C (SS-34 rotor, Sorvall RC-5) and resuspended in 10 ml 10 mM $MgSO_4$.

2.3.14.2 Library plating

First round library screening involved adding 5 x 10^4 recombinant phage to 200 µl plating bacteria (*E. coli* BB4; 2.2.9) and incubating at 37°C for 15 min. 10 ml LMM agarose (2.2.9) at 50°C was then transferred to the phage/bacteria mixture and overlayed onto 15 cm L agar plates. Second and third round library screening was carried out by plating appropriate dilutions of first round positive plugs (which were picked with the wide end of a pasteur pipette and eluted overnight in 1 ml of PSB (2.2.5)) as described above with two exceptions. Firstly, infections were carried out using 100 µl of plating bacterial and secondly, 5 ml of LMM agarose (2.2.9) was used to plate phage onto 10 cm L agar plates (2.2.9). Bacteriophage plaques were formed by incubating plates at 37°C for 5-7 hours. When the plaques had grown to a sufficient size and density, the plates were cooled at 4°C for at least 1 hour before plaque lifts.

2.3.14.3 Plaque lifts

Genescreen filters (DuPont) were placed onto first round library plates for 1 min. (1st lift) or 3 min. (2nd lift) and transferred to 3MM chromatography paper (Whatman) to air dry. Filters were then autoclaved for 2 min. and nucleic acid was bound to the filter by UV crosslinking. Nitrocellulose filters (Scheicher and Schell) cut to the size of 10 cm bacteriological dishes were used for second and third round screen lifts. After lifting, nitrocellulose filters were transferred to 3MM chromatography paper (Whatman) and soaked in denaturing solution (2.2.5) for 5 min., and then transferred to 3MM chromatography paper (Whatman) soaked in neutralising solution (2.2.5) for 5 min., before rinsing in 2X SSC for 5 min. Filters were air dried for 30 min. and then UV crosslinked. Filters were prehybridised at 42°C for a minimum of 4 hours in petri dishes with a sufficient amount of pre-hyb solution (2.2.5) to cover all filters.

2.3.14.4 High Titre Stock Production

Isolated duplicate positive plaques from second or third round screening were picked with the narrow end of a pasteur pipette and phage were eluted overnight at 4°C in 400 µl of PSB (2.2.5). The eluate was serially diluted and 1 µl of each dilution incubated with BB4 *E. coli* (2.2.9) before plating onto L agar plates (2.2.9). Plates observed to have confluent lysis were

soaked in 1-2 ml of PSB (2.2.5) for 3 hours with shaking. The eluate was transferred to eppendorf tubes and a drop of chloroform was added to kill any bacteria present.

2.3.15 Lambda Zapping

Purified cDNA clones were isolated as plasmids in pBluescript II SK from second or third round positive plaques lambda zapping according to the manufacturers instructions (Clontech).

2.3.16 Reverse Transcription

10 µg total RNA was reverse transcribed using superscript II™ reverse transcriptase according to the manufacturer's instructions (Gibco/BRL). cDNA was diluted to 100 µl with MQ H_2O prior to use in PCR reactions.

2.3.17 PCR

PCR mixes contained either 10 ng plasmid template or 5 µl of diluted reverse transcribed cDNA, 100 ng of each primer, 2 U of Taq polymerase, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, and 1X Taq buffer (2.2.5). PCR reactions were carried out in a capillary thermal sequencer (Corbett Research) using PCR conditions of denaturation at 94°C for 5 sec., annealing at 50°C for 5 sec., and extension at 72°C for 1 min per 1 kb to be polymerised for a total of 35 cycles.

2.3.18 Isolation of Cytoplasmic RNA from Cultured Cells

Cytoplasmic RNA was isolated using the method of Edwards *et al.*, (1985). Cells were harvested by trypsinisation and washed twice with PBS. Cell pellets were thoroughly resuspended in 2 ml ice cold TNM (2.2.5) and lysed by pipetting vigorously. After incubation on ice for 5 min., the nuclei were pelleted by centrifugation at 3000 r.p.m. for 5 min. (Jouan, C412), and the supernatant removed and mixed thoroughly with 2 ml TUNES (2.2.5). This solution was extracted twice with phenol/chloroform (1:1) and the aqueous layer was transferred to a corex tube. RNA was precipitated by addition of 1/10th volume NaAc pH 5.2 and 2.5 volumes RNAse free ethanol, and incubated at -80°C for 30 min. or O/N at -20°C. RNA was pelleted by centrifugation at 9000 r.p.m. for 30 min. at 4°C (SS-34 rotor, Sorvall RC-5), resuspended in 450 µl of sterile water, and transferred to an eppendorf tube for re-precipitation as

described above. RNA was resuspended in an appropriate volume of water, and concentration determined by spectrophotometric analysis at 260 nm.

2.3.19 Isolation of Polyadenylated RNA

Isolation of polyadenylated (poly-$(A)^+$) RNA from cultured cells was carried out using the method of Celano *et al.* (1993). Oligo d(t) cellulose was hydrated by 3 washes in Na-TES (2.2.5) at RT to give a final concentration of 100 mg/ml. 1 mg of total RNA was made up to 600 µl with TES (2.2.5) in an eppendorf tube and heated at 65°C for 5 min., before cooling on ice. 1/10th the volume of 5 M NaCl and an equal volume of hydrated oligo d(t) cellulose was added and mixed well, before incubation at 37°C for 10 min with occasional mixing by gentle inversion. The cellulose was then pelleted, the supernatant removed, and the pellet washed by resuspension in 1 ml Na-TES (2.2.5). The cellulose was washed again with 1 ml of ice cold MQ H_2O before RNA was eluted from the cellulose by resuspension in 400 µl of MQ H_2O heated at 55°C for 5 min. This elution step was repeated a further 2 times before pooling eluates and precipitating with 2 M NH_4Ac and 2 volumes of ethanol at -20°C O/N. The recovered poly-$(A)^+$ RNA was washed with 70% (v/v) ethanol and resuspended in an appropriate volume of MQ H_2O before spectrophotometric analysis at 260 nm to determine RNA concentration.

2.3.20 Isolation of RNA from Mouse Tissues

RNA from tissue samples was isolated using the guanidium thiocyanate method (Chomczynski and Sacchi, 1987). Tissues were harvested and immediately frozen at -80°C until required. Tissues were thawed on ice in 1 ml of solution D (2.2.5) and homogenised using a glass teflon homogeniser. The solution was transferred to a 10 ml polypropylene tube and 0.1 ml 2 M NaAc pH 4, 1 ml of phenol and 0.2 ml chloroform:isoamylalcohol (49:1) added sequentially and mixed by gentle inversion, before vigorous mixing once all reagents were added. The homogenate was incubated on ice for 15 min., before centrifugation at 3800 r.p.m. (Jouan, C412) and the aqueous layer transferred to a corex tube. RNA was precipitated by the addition of 1 ml of isopropanol and incubation at -20°C for a minimum of an hour. RNA was recovered following centrifugation at 9000 r.p.m. for 30 min. at 4°C (SS-34 rotor, Sorvall RC-5) and the pellet resuspended in 400 µl of solution D. Following transfer to an eppendorf tube, RNA was

re-precipitated, washed with 75% (v/v) ethanol, dried and resuspended in an appropriate volume of MQ H_2O. RNA concentration was determined by spectrophotometric analysis at 260 nm.

2.3.21 Ribonuclease Protection Analysis

2.3.21.1 Riboprobe templates and probe synthesis

Linearised riboprobe templates were prepared by digesting 20 µg of DNA with the appropriate enzyme (2.3.4), followed by phenol/chloroform extraction, ethanol precipitation and resuspension in 20 µl of MQ H_2O. *K7* riboprobes for RNAse protection analysis were generated from a 370 bp *Hind*III/*Sty*I fragment from the *K7* cDNA (nucleotide position 4197-4567 bp). Antisense templates were generated by *Eco*RI digestion, and riboprobes produced by transcription with T7 RNA polymerase. *mGap* riboprobes for RNAse protection were generated from the 300 bp *mGap* cDNA clone (2.2.6.2). The antisense template was generated by *Bam*HI digestion and transcribed with SP6 RNA polymerase. Transcription reactions (in 15 µl) contained 125 µCi of [α-^{32}P]-rUTP for *K7* riboprobes and 40 µCi [α-^{32}P]-rUTP for *mGap* riboprobes, transcription buffer (2.2.5), 2.25 µl of 3.3 mM rATP/rCTP/rGTP mix, 0.5 µl of 1 mM rUTP, 1 µg DNA template, 10 U RNAsin and 10 U of the appropriate RNA polymerase. Reactions were incubated at 37°C for 75 min., before the addition of 20 U of DNAse 1 and incubation at 37°C for a further 15 min. Unincorporated label was removed using a sephadex G-50 column and spinning at 3000 r.p.m. for 5 min. (Jouan, C412).

2.3.21.2 RNAse protection assay

RNAse protection analysis was carried out using the method of Krieg and Melton (1987). 150,000 c.p.m. of single stranded antisense *K7* riboprobe and 37,000 c.p.m. of single stranded antisense *mGap* riboprobe were added to each RNA sample. Probe/RNA samples were precipitated, and the recovered pellet resuspended in 6 µl 5X Hyb (2.2.5) and 24 µl deionised formamide, before denaturation at 85°C for 3 min., and hybridisation O/N at 45°C. The following day, samples were digested at 37°C for 30 min. in RNAse digestion buffer (2.2.5), incubated for a further 15 min. with 2 µl proteinase K and 10 µl of 10% (w/v) SDS, and then phenol/chloroform extracted. RNA digestion products were re-precipitated and resuspended in 1 µl of MQ H2O and 4 µl of FLB (2.2.5), heated to 85°C for 3 min., snap cooled and analysed by

electrophoresis at 60 W on a 6% (v/v) denaturing polyacrylamide gel prepared from 40% sequagel 6. Following electrophoresis, the gel was transferred to 3MM chromatography paper (Whatman), and dried at 70°C under vacuum. Radioactivity was detected by exposure to phosphorimager screens (Molecular Dynamics).

2.3.22 Northern Blot Analysis

1% agarose gels for northern blot analysis were prepared by dissolving 2.5 g agarose in 210 ml MQ H_20. Once the gel solution had cooled to 60°C, 25 ml 10X MOPS (2.2.5) and 15 ml 20% (w/v) formaldehyde (freshly prepared by dissolving 4 g paraformaldehyde in 20 ml MQ H_2O) were added before pouring the gel.

Poly-$(A)^+$ RNA samples (4 µg), or total RNA samples (25 µg) were prepared for electrophoresis as follows:

11.25 µl	RNA + MQ H_2O
5 µl	10X MOPS
8.75 µl	formaldehyde (37%, pH 4.5)
25 µl	deionised formamide

RNA samples were denatured by heating at 65°C for 15 min., snap cooled on ice, before the addition of 5 µl of RNAse free 10X GLB (2.2.5). 5 µg *Eco*RI digested SPP-1 phage markers (2.2.10) were also loaded. Northern gels were run at 80 V in 1X MOPS (2.2.5) until the bromophenol blue dye had run to within 3/4 of the bottom of the gel. The DNA marker containing lane was removed from the gel, stained in EtBr for 45 min. and destained in water O/N before photographing under medium wavelength UV light. The remainder of the gel was blotted onto Hybond-N^+ membrane (Amersham) using capillary transfer as described in 2.3.6, with the exception of the transfer being carried out in 20 X SSC. RNA was cross-linked to the filter using a UV Stratalinker™ (Stratagene). The Northern filter was then placed on nylon mesh RNA side up, rolled up, and prehybridised in a Hybaid cylinder at 42°C for 4-16 hours in 10 ml of pre-hyb solution (2.2.5).

2.3.23 Synthesis of Radioactive DNA Probes

Radioactive probes were synthesised from DNA fragments digested from cloned DNA sequences described in section 2.2.6.2 as follows:

mGap: A 300 bp *mGap* fragment was obtained following digestion with *Hind*III/*Pst*I.

Oct4: An 462 bp *Oct4* fragment was produced following excision with *Xho*I/*Hind*III.

Rex1: An 848 bp *Rex1* fragment was generated following digestion with *Eco*RI.

Fgf5: An 800 bp *Fgf5* DNA fragment was obtained following digestion with *Eco*RI/*Bam*HI.

L17: A 736 bp containing *L17* fragment was excised with *Xho*I/*Hind*III.

K7: An 232 bp containing *K7* fragment was produced following digestion with *Bam*HI/*Hind*III.

Psc1: A 458 bp *Psc1* fragment was excised using *Bam*HI/*Hind*III.

DNA probes were prepared using the Gigaprime labelling kit (2.2.3) with 50 µCi [α-^{32}P] dATP according to the manufacturer's instructions. The reactions were terminated by dilution to 100 µl with MQ H_2O and incubation at 65°C for 10 min. Unincorporated label was removed from the probe by centrifugation through a Sepharose CL-6B column at 1800 r.p.m. for 3 min. (Jouan C412).

2.3.24 Hybridisation of Radioactive DNA Probes to Nylon and Nitrocellulose Filters

DNA probes were boiled for 5 min., snap-cooled on ice and added to filters in fresh prehyb solution. Plasmid Southern blots and Northern filters were hybridised at 42°C in hybaid cylinders, whilst plaque lifts were hybridised at 42°C in petri dishes (up to 20 per dish). The following day filters were removed from hybaid cylinders or petri dishes and washed with increasing stringency (2X SSC/0.1% SDS at 42°C, to 0.1X SSC/0.1% SDS at 65°C) until non-specifically bound counts had been removed. Filters were subsequently sealed in plastic bags and radioactivity detected by exposure to phosphorimager cassettes (Molecular Dynamics) or exposure to Biomax autoradiographic film (Kodak).

2.3.25 Phosphorimager Analysis and Image Manipulation

Gels and filters exposed to phosphorimager screens (Molecular Dynamics) were processed using a Molecular Dynamics phosphorimager, running the ImageQuant software package.

Quantitation by volume integration was carried out using ImageQuant. Phosphorimager files were manipulated using the Adobe Photoshop™ and PowerPoint programs. Autoradiographic film and photographic slides were scanned and manipulated using Adobe Photoshop™.

2.3.26 Transformation of Yeast Strains

10 ml of YEPD media (2.2.10) was inoculated with a single colony of the appropriate yeast strain and grown at 25°C overnight O/N. This culture was subcultured 1/40 in YEPD media (2.2.10) and grown to an OD_{600} of 0.8-1.0. 1 ml of cells was transferred to an eppendorf tube and harvested by pulse centrifugation. The supernatant was decanted and the cell pellet resuspended in the residual supernatant. 2 µg of ssDNA was added and mixed before the addition of 2 µg of plasmid DNA. 750 µl of plating solution (2.2.5) was then added and mixed thoroughly by gentle pipetting, and incubated O/N at RT. 500 µl of the transformation was plated onto appropriate selective media and incubated at 25°C until transformants appeared (2-3 days).

2.3.27 Cos-1 Cell Lysis

On the second day following transient transfection (2.6.7ii) Cos-1 cells, grown in 6 cm petri dishes, were wash in PBS and then harvested by scraping with a rubber policeman. Cells transferred to a 1.5 ml eppendorf tube, and pelleted for 2 min. Cell pellets were lysed in 100 µl of NP-40 lysis buffer (2.2.5) on ice for 30 min. Cell debris was removed following centrifugation for 2 min. at 12000 r.p.m. in an eppendorf centrifuge, and 10 µl of supernatants, containing the detergent soluble cytoplasmic and nucleoplasmic proteins, were analysed for protein expression by SDS-PAGE (2.3.29) and western analysis (2.3.30), following addition of 10 µl of 2X SDS-load buffer (2.2.5).

2.3.28 Immunoprecipitation

On the second day following transient transfection (2.6.7ii) Cos-1 cells, grown in 6 cm petri dishes, were washed in PBS and lysed for immunoprecipitation by the addition of 1 ml of IP lysis buffer (2.2.5) and incubation of ice for 30 min.. Cells were harvested by scraping with a rubber policeman and transferred into a 1.5 ml eppendorf tube. Cell debris was removed by centrifugation for 10 min. at 10000 r.p.m. in an eppendorf centrifuge, and the supernatant

transferred to a new tube. Lysates were pre-cleared of proteins which bind non-specifically to protein G-sepharose by incubating in 30 µl/ml of protein G-sepharose O/N at 4°C. Protein G-sepharose was removed by centrifugation and the lysate split in two 500 µl samples. 1 µg of anti-GFP polyclonal antibodies were added to one 500 µl lysate sample and 13.2 µg of anti-FLAG monoclonal antibodies was added to the other sample and incubated for 3 hours at 4°C while rotating. 30 µl/ml of protein G-sepharose was subsequently added and incubated at 4°C for 2 hours while rotating. The sepharose were pelleted, unbound protein removed by washing IP lysis buffer (2.2.5), followed by two washed with PBS. The sepharose was pelleted, resuspended in 30 µl of 1X SDS-load buffer, boiled for 10 min. and analysed by SDS-PAGE (2.3.29).

2.3.29 SDS-PAGE Analysis

SDS-polyacrylamide gels, containing 1X Tris-SDS buffer (2.2.5), 0.1% (w/v) APS and 0.1% (v/v) TEMED, were poured using 0.75 mm spacers and allowed polymerise for approximately 30 min. beneath an overlay of distilled water. After polymerisation, the water was removed and a 4% stacker gel, containing 1X Tris-SDS buffer, 0.1% APS (w/v) and 0.1% (v/v) TEMED was applied. 10 well combs (0.75 mm) were inserted and the gel left to polymerise for another 30 min. Gels were electrophoresed using a PAGE minigel apparatus (Biorad) in SDS-PAGE buffer (2.2.5) at 200 V. Gels were either stained in coomassie stain (2.2.5) for a minimum of 30 min. at 37°C, and destained (2.2.5) overnight to visualise the protein bands, or transferred to nitrocellulose by western blot (2.3.30).

2.3.30 Western Analysis

Proteins were transferred from SDS-polyacrylamide gels to nitrocellulose (Schneider and Schell) in western transfer buffer (2.2.5), using a semi-dry western apparatus according to the manufacturers instructions (Hoefer Pharmacia Biotech). Filters were air dried on 3MM chromatography paper (Whatman) and blocked by incubation in 5% (w/v) milk powder in PBT overnight. An appropriate dilution of primary antibody (2.2.8) in PBT was added to the filter and incubated for 1 hour. Filters were washed using 4 x 15 min. washes in PBT before incubation with the appropriately diluted HRP-conjugated secondary antibody (2.2.8) for 1 hour. Following a further 6 x 10 min. washes in PBT, the blot was developed by bathing in enhanced

chemiluminescence reagents for 5 min. (SuperSignal™ Substrates), drained and exposed on Biomax autoradiographic film (Kodak) for an appropriate time (1 sec-5 min.).

2.3.31 Immunofluorescence

Cos-1 cells were seeded onto coverslips (Crown Scientific) in 6 well dishes and transiently transfected (2.6.7) with 1 μg of the appropriate mammalian expression plasmid. Cells were washed twice with PBS, and fixed by immersing in methanol for 2 min. at -20°C and re-hydrating for 15 min. in PBS. For immunolocalisation of centrosome localised γ-tubulin protein, cells were fixed and permealised in two ways. Cells were either incubated for 2 min. in 0.1% Triton X-100 in 80 mM PIPES, 5 mM EDTA and 1 mM $MgCl_2$ at RT, prior to methanol fixation for 3 min., or fixed with 3.7% PFA for 5 min. at RT, washed twice with PBS and cell membranes permealised using 0.1-0.3% Triton X-100 in PBS. Primary antibodies were diluted appropriately (2.2.8) in 3% (w/v) BSA/PBT and incubated with the cells for 60 min. at RT. Unbound primary antibody was removed following 3 x 5 min. washes in PBT. Cells were incubated with the appropriate secondary antibodies (2.2.8), diluted 1/50 in 3% (w/v) BSA/PBT, at RT for 60 min. in the dark, before being washed 3 times in PBT for 5 min. Nuclei were visualised by staining with 0.5 μg/ml Hoechst 33258 Trihydrochloride (BisBenzamide) for 1 min., or 2 μg/ml of propidium iodide for 1 min. before washing twice for 5 min. with PBT. Coverslips were viewed using either a Zeiss Axioplan microscope equipped for 3 channel fluorescence (Zeiss filter sets II, IX, and XV) and photographed with a Zeiss MC 100 camera attachment using 35 mm ektachrome 160T film (Kodak) or a Nikon Eclipse TE300 inverted microscope equipped with a TE-FM Epi-Fluorescence attachment, and photographed with a Nikon F70/F70D camera attachment using ektachrome 100 ASA slide film (Kodak). Images from confocal laser scanning microscopy were acquired in the equatorial plane using a 100X water immersion objective (numerical aperture, 1.4) using a MRC1000UV unit (Biorad) in combination with a Diaphot 300 inverted Nikon microscope. Overlaid images were prepared using Confocal Assistant 4.0.

2.3.32 Flow Cytometric Analysis of Transiently Transfected Cos-1 Cells

Cos-1 cells, grown in 10 cm dishes, were transfected with the appropriate expression plasmid (2.6.7ii). Cells were harvested 30-36 hours post-transfection by washing twice with PBS followed by incubation in TEN buffer (2.2.5) for 5 min. and vigorous pipetting. Cells were pelleted by centrifugation at 1200 r.p.m. for 4 min., before washing once with PBS.

2.3.32.1 Flow cytometric analysis of the entire population of transfected Cos-1 cells

To determine the DNA content of an entire population of Cos-1 cells which had been transiently transfected with mammalian expression constructs, cells were harvested as described above, resuspended in 300 µl of PBS/1% FCS and fixed by drop-wise addition of 900 µl of absolute ethanol while vortexing. Cells were incubated on ice for 10 min. before washing twice with PBS. Cell pellets were resuspended in a final volume of 250 µl of PBS, prior to addition of 50 µg/ml RNAse A to digest RNA. DNA was subsequently stained with 10 µg/ml of propidium iodide and incubated in the dark for 20 min. before processing in the cytometer. Flow cytometric analysis was performed using an EPICS XL-MCL cytometer (Coulter). Proportions of cells in each phase of the cell cycle was calculated with the use of algorithms incorporated into the flow cytometer XL program.

2.3.32.2 Flow cytometric analysis of transfected Cos-1 cells within the population

Simultaneous detection of the DNA content in Cos-1 cells overexpressing EGFP-tagged proteins (ie. the DNA content of transfected cells within a population) was carried out following fixation overnight in 80% ethanol. Cells were washed twice with PBS and incubated with 50 µg/ml of RNAse A for 30 min. at RT. Cells were then incubated with 10 µg/ml of propidium iodide for 20 min. prior to processing in the cytometer. Determination of the DNA content in Cos-1 cells overexpressing FLAG-tagged proteins was carried out using a modified method of Gong *et al.* (1994). Harvested cells were fixed in 80% (v/v) ethanol at -20°C for 2-24 hours, washed twice with PBS and treated with 0.25% (v/v) Triton X-100 in PBS for 5 min. on ice. Cells were washed twice with PBS and then incubated at 4°C O/N with the anti-FLAG

monoclonal antibody diluted 1/400 in 1% (w/v) BSA/PBS (2.2.8). Unbound antibody was removed following two washes in PBS prior to incubation with an anti-mouse FITC-conjugated secondary antibody, diluted 1/50 in 1% (w/v) BSA/PBS, for 1 hour at 4°C (2.2.8). Cells were washed twice in PBS, treated with 50 µg/ml RNAse A for 30 min. at RT and then 10 µg/ml propidium iodide for 20 min. prior to processing in the cytometer. Flow cytometric analysis was performed using an EPICS XL-MCL cytometer. FITC/EGFP and propidium iodide fluorescent signals were detected simultaneously using the FL1 and FL3 filters respectively. Proportions of cells in each phase of the cell cycle was calculated with the use of algorithms incorporated into the flow cytometer XL program.

2.3.33 Sequencing Software and Database Searches

Contiguous sequence alignments were carried out using the MacDNASIS program. Homology searches and alignments were carried out using the BLAST program of Altschul *et al.* (1990). cDNA sequences were compared to the non-redundant nucleotide database and expressed sequence tag database, or translated in six reading frames and compared to the non-redundant nucleotide database translated in all reading frames using the National Centre for Biotechnology Information web site (www.ncbi.nlm.nih.gov/blast). Alignment of *K7* and *KIAA0165* DNA sequences was carried out using the ALIGN program (Genestream network IGH France; www2.igh.cnrs/bin/align-guess.cgi). Alignment of K7 and KIAA0165 protein sequences was carried out using BLAST 2 program of Tatusova and Madden (1999) (www.ncbi.nlm.nih.gov/blast). Multiple sequence alignments were carried out using the CLUSTALW program of Thompson *et al.* (1994) (dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html) and identities and similarities shaded using the Boxshade program (www.ch.embnet.org/software/BOXform.html). The predicted K7 protein sequence was examined for motifs using the PROSITE database (www.expasy.ch/prosite). However, as weakly defined sequence motifs appear frequently and thus are unlikely to be informative, amidation sites, N-glycosylation sites, cAMP- and cGMP-dependent protein kinase phosphorylation sites, casein kinase II phosphorylation sites, N-myristylation sites, protein kinase C phosphorylation sites and tyrosine phosphorylation sites were excluded from the analysis. Protein sequences were also analysed using the protein structure prediction server of

Fischer and Eisenberg (1996) (www.doe-mbi.ucla.edu/people/frsvr/frsvr.html) which combines sequence structure compatibility with secondary structure prediction.

2.3.34 Containment Facilities

All manipulations involving viable organisms which contained recombinant DNA were carried out in accordance with the regulations and approval of the Australian Academy of Science Committee on Recombinant DNA and the University Council of the University of Adelaide.

2.4 EMBRYOLOGICAL METHODS

2.4.1 Mouse Strains and Matings

Embryos to be used for *in situ* hybridisation were generally produced from natural Swiss matings. Mice were maintained on a 12 hour light/12 hour dark cycle and supplied with food and water *ad libitum.*

2.4.2 Embryo Processing Baskets

For *in situ* hybridisation procedures, pre-implantation and post-implantation embryos were processed in siliconised plastic processing baskets. Baskets were made (in a fume hood) by removing the bottom of a screw capped eppendorf tube with a heated scalpel blade. The top of the tube was briefly melted over a Bunsen burner and held down on a piece of fine grade nylon mesh (pore size approximately 20-50 μm), until the plastic had reset. Excess mesh was removed and the basket was siliconised.

2.4.3 Embryo Isolation and Dissection

Embryo stages are represented as number of days *post coitum* (d.p.c.), assuming fertilisation occurred at midnight preceding the discovery of the vaginal plug. Slight variation in developmental progression occurred within litters and between litters such that individual embryos were reclassified according to idealised developmental time points (Kaufman, 1992; Hogan *et al.,* 1994; 3.2.1). Isolation and dissection of pre-implantation and post-implantation embryos was carried out according to Hogan *et al.* (1994) with the following modifications:

i) Pre-implantation embryos:

Uteri were removed and embryos were flushed from the uteri with dissection buffer (2.2.5). The zona pellucida was not removed.

ii) 5.0-5.5 d.p.c. embryos:

Uteri were removed, placed into dissection buffer (2.2.5), and individual implantation sites isolated by cutting the uterus on either side of the implantation site. The mesometrial side of the early decidua was prised open and embryos were "shelled" out of the decidua with fine forceps.

2.4.4 Embryo Fixation and Dehydration

Dissected pre-implantation and post-implantation embryos were handled using siliconised glassware. Embryos were rinsed once in PBS and once in 4% PFA/PBT. Embryos were transferred to siliconised processing baskets and fixed in 4% (w/v) PFA/PBT for 4-16 hours at 4°C. Embryos were rinsed twice on ice with PBT and dehydrated on ice by successive 5 min. washes in 25% (v/v) methanol/PBT, 50% (v/v) methanol/PBT, and 75% (v/v) methanol/PBT and 2 final washes in 100% methanol. Embryos were stored in 2.5 ml 100% methanol at -20°C in processing baskets for up to 6 months in sealed 10 ml polypropylene tubes.

2.4.5 Digoxygenin Labelled RNA Probes

2.4.5.1 Preparation of templates

Templates for DIG riboprobe synthesis were prepared as described in section 2.3.21.1, by digesting plasmids described in section 2.2.6.2 with the appropriate restriction enzymes to linearise the plasmids. Specific DIG labelled riboprobes were generated using the following templates and RNA polymerases:

L17: The sense template was generated by *Bam*HI digestion, and riboprobes transcribed using T7 RNA polymerase. Antisense templates were generated by *Xho*I digestion, and riboprobes produced with T3 RNA polymerase.

K7: The sense template was generated by *Bam*HI digestion, and riboprobes transcribed using T3 RNA polymerase. Antisense templates were generated by *Hind*III digestion, and riboprobes produced with T7 RNA polymerase.

Psc1: The sense template was generated by *Hind*III digestion, and riboprobes transcribed using T7 RNA polymerase. Antisense templates were generated by *Bam*HI digestion, and riboprobes produced with T3 RNA polymerase.

Oct4: The sense template was generated by *Xho*I digestion, and transcripts polymerised with T7 RNA polymerase. The antisense template was generated by *Eco*RI digestion, and transcripts produced with T3 RNA polymerase.

Rex1: Sense and antisense riboprobes were produced using *Bam*HI or *Xba*I linearised templates and T7 and Sp6 RNA polymerases respectively.

Fgf5: Sense and anti-sense *Fgf5* probes were generated by linearising the plasmid with *Eco*RI or *Bam*HI, and transcribing with T7 and T3 RNA polymerases respectively.

Nkx2.5: Antisense riboprobes were synthesised using a *Hind*III linearised template and T3 RNA polymerase.

2.4.5.2 DIG labelled riboprobe synthesis

Riboprobes were synthesised in transcription reactions containing 1 µg linearised plasmid, transcription buffer (2.2.5), 1X DIG labelling mix (10 mM each of rATP/rCTP/rGTP, 6.5 mM UTP, 3.5 mM DIG-UTP), 20 U of RNAsin, and 20 U of the appropriate RNA polymerase. Transcription reactions were incubated at 37°C for 2 hours before the template was removed by addition of 40 U RNAse free DNAse 1 and incubation at 37°C for 15 min. Reactions were precipitated at -20°C (1 hour-O/N) after addition of 60 µl MQ H_20, 20 µl 100 mM EDTA, 10 µl 3M NaAc pH 5.2, and 250 µl ethanol. After centrifugation for 15 min., DIG probes were resuspended in 100 µl of RNAse free MQ H_2O containing 40 U of RNAsin. A 5 µl sample of resuspended probe was used to assess riboprobe yield and quality by agarose gel electrophoresis (2.3.5).

2.4.6 Embryonic *in situ* Hybridisation

In situ hybridisation to pre-implantation and post-implantation mouse embryos was carried out using the whole mount method of Rosen and Beddington (1993) with the following modifications. Embryos were treated with 2 ml solution volumes in 24 well trays using processing baskets (2.4.2) to minimise embryo loss. Dehydrated embryos in 100% methanol

were rehydrated on ice with successive 5 min. washes in 75% (v/v) methanol/PBT, 50% (v/v) methanol/PBT, and 25% (v/v) methanol/PBT before rinsing twice with PBT for 5 min. at RT. All further washes were carried out at RT unless specified. Embryos were bleached with 6% (v/v) H_2O_2/PBT for 1 hour and rinsed 3 times with PBT for 5 min. Embryos were washed with RIPA buffer (2.2.5) 3 times for 20 min., rinsed twice in PBT for 5 min., and postfixed in 4% (w/v) PFA / 0.2% (v/v) gluteraldehyde in PBT for 20 min. Embryos were then washed 3 times for 5 min. in PBT, for 5 min. in 1:1 *in situ* Hyb:PBT, and for a further 5 min. in *in situ* Hyb (2.2.5) before being prehybridised in fresh *in situ* Hyb for 1-4 hours at 65°C. Prehybridisation and hybridisation steps were carried out in a sealed humidified box containing paper towels soaked in 50% (v/v) formamide. DIG-labelled riboprobes were diluted 1/200, denatured at 80°C for 10 min., snap cooled on ice, added to fresh *in situ* Hyb and hybridised to embryos O/N at 65°C. The following day, embryos were washed at 65°C for 5 min. sequentially in *in situ* buffer 1 (2.2.5), 70% *in situ* buffer 1:30% 2X SSC, and 30% *in situ* buffer 1:70% 2X SSC. Embryos were then washed twice in 2X SSC/0.1% CHAPS at 65°C for 30 min. and twice in 0.2X SSC/0.1% CHAPS at 65°C for 30 min. Embryos were washed twice in TBST (2.2.5) for 10 min., followed by incubation in 10% FCS/TBST for at least one hour. Preblocked anti-digoxygenin Fab fragments-AP conjugate was prepared as described in section 2.4.8, and incubated with the embryos for 6-24 hours at 4°C with gentle rocking. Embryos were rinsed for 5 min. with TBST before being washed O/N in TBST with gentle rocking. The following day, embryos were washed for 3 x 10 min. in AP buffer (2.2.5). The embryos were gently flushed from the processing baskets into siliconised dishes containing AP buffer. Embryos were subsequently identified and transferred to siliconised welled microscope slides (Crown Scientific) for staining. *In situ* hybridisations were developed in humid chambers by addition of 50-100 µl of *in situ* substrate mix (2.2.5) and incubation in the dark until purple staining appeared (1-6 hours). The staining reaction was terminated by transferring embryos through several rinses of PBT/1 mM EDTA. The stain was fixed by incubation in 4% (w/v) PFA / 0.2% (v/v) gluteraldehyde in PBT for 20 min. and embryos were stored in 0.4% (w/v) PFA/PBT. Embryos were viewed on a Nikon Eclipse TE300 inverted microscope using Hoffmann modulation contrast system optics, and photographed with Ektachrome 100 ASA slide film (Kodak).

2.4.7 Mouse Embryo Powder

14.5 d.p.c. mouse embryos were homogenised in PBS, mixed with 4 volumes ice-cold acetone and incubated on ice for 30 min. The homogenate was spun at 10000 r.p.m. for 10 min. at 4°C (SS-34 rotor, Sorvall RC-5) and the supernatant discarded. The pellet was washed in ice-cold acetone and centrifuged as above. The pellet was ground to a powder on 3MM chromatography paper (Whatman), air dried and stored in an airtight tube at 4°C.

2.4.8 Antibody Preblocking

100 µl embryo powder (2.4.7) was resuspended in 1 ml TBST, heat inactivated for 30 min. at 70°C, spun and the pellet cooled on ice. The pellet was resuspended in 1 ml 1% (v/v) FCS/TBST, and 1 µl anti-digoxygenin Fab fragments-AP conjugate (2.2.8) added for each *in situ* to be performed. The antibody was blocked for at least 3 hours at 4°C with gentle rocking. The antibody was centrifuged for 5 min. at 4°C and the supernatant diluted with 1% (v/v) FCS/TBST to give 2 ml of pre-blocked antibody solution for each *in situ* hybridisation to be performed.

2.4.9 Animal Manipulations

All procedures involving animals were carried out with the approval of the University of Adelaide animal ethics committee.

2.5 POLYCLONAL ANTIBODY GENERATION

2.5.1 Generation of GST-K7 peptide constructs

In order to generate anti-K7 polyclonal antibodies from two distinct regions of the K7 protein, two recombinant K7 protein fragments from amino acid positions 30-170 (K7-Ab1) and 1267-1369 (K7-Ab2) were amplified by PCR (2.3.17) using the K7 cDNA clones as templates (7A pBs-SK for K7-Ab1 and S1 pBs-SK for K7-Ab2; refer Figure 4.1) and 5'Ab1 and 3'Ab1, and 5'Ab2 and 3'Ab2 primers respectively (2.2.7.3). PCR products were digested with *Bam*HI and *Eco*RI and directionally cloned into *Bam*HI/*Eco*RI digested pGex2T. Clones were sequenced to check that no errors were incorporated during the PCR reaction and that the K7 peptide sequence was in frame with the GST sequence.

2.5.2 Preparation of Competent BL21 Cells

50 ml of LB (2.2.9) was inoculated with a single BL21 *E. coli* colony and cultured to an OD_{600} of 0.6. The cells were harvested by centrifugation at 4000 r.p.m. for 10 minutes at 4°C (SS-34 rotor, Sorvall RC-5). Cell pellets were resuspended in 25 ml of cold 0.1 M $MgCl_2$ and pelleted by centrifugation at 4000 r.p.m. for 10 min. at 4°C. Cells were resuspended in 12.5 ml of cold 0.1 $CaCl_2$ and incubated on ice for 20 min. Cells were harvested by centrifugation at 4000 r.p.m. for 10 min. at 4°C and resuspended in 1.6 ml of cold 0.1 M $CaCl_2$ and 15% (v/v) glycerol. 200 µl aliquots were placed into eppendorf tubes, snap frozen in a dry ice/ethanol bath and stored at -80°C.

2.5.3 Small Scale GST-fusion Protein Induction

Following transformation of competent BL21 with the appropriate plasmid DNA (2.3.12), a single transformed colony was picked and grown in 2 ml of LB plus ampicillin (100 µg/ml) overnight. The overnight culture was then sub-cultured 1/100 into 10 ml of fresh LB and ampicillin (100 µg/ml) and grown at 37°C with agitation to an OD_{600} of approximately 0.6. Fusion protein expression was induced by the addition of 0.2 mM IPTG following the removal of a 1 ml pre-induction sample. Incubation continued for 4 hours at 37°C with 1 ml samples being taken every hour. Cells were harvested by centrifugation and resuspended in 100 µl bacterial lysis buffer (2.2.5). For ease of manipulation, the cell lysates were heated at 100°C for several minutes before transferring 50 µl of bacterial cell lysate into 50 µl of 2X SDS load buffer (2.2.5). This solution was again heated at 100°C for 3 min. before loading 20 µl onto a 12.5% SDS-polyacrylamide gel for electrophoresis (2.3.29).

2.5.4 Large Scale GST Fusion Protein Induction

200 ml of LB (2.2.9) plus ampicillin (100 µg/ml) was inoculated with diluted overnight culture, grown to log phase (OD_{600} = 0.6) and fusion protein expression induced as described in section 2.5.3. The cells were harvested by centrifugation at 5000 r.p.m. for 10 min. (GSA rotor, Sorvall RC-5) and the cell pellet resuspended in 20 ml of TBST containing 0.25 mM PMSF. Cells were lysed by 3 rounds of sonication for 30 sec., adding 0.25 mM PMSF between sonications. Cell debris was pelleted by centrifugation at 10000 r.p.m. for 15 min. (SS-34 rotor,

Sorvall RC-5) and the supernatant removed. Protein solubility was determined by resuspending the pellet in 1.5 ml of TBS and analysing 10 µl of both resuspended pellet and supernatant in 10 µl of 2X SDS load buffer by SDS-PAGE (2.3.29).

2.5.5 Preparation of GST-fusion Protein for Immunisation

Insoluble fusion proteins were prepared for immunisation into rabbits using the large scale induction protocol outlined above (2.5.4). Insoluble GST-fusion protein concentration was estimated by analysing several dilutions of the resuspended pellet by SDS-PAGE (2.3.29), along with BSA standard of known concentration. Approximately 100 µg of fusion protein was then loaded onto 10% SDS-polyacrylamide preparative gels, along with prestained molecular weight standards (2.2.12), and a band excised which corresponded to the correct molecular weight of the fusion protein of interest. Gel bands were then processed serially through 18 to 25 gauge needles in readiness for immunisation.

2.5.6 Immunisation Regime

Preimmune samples (5 ml) were taken from each rabbit prior to immunisation. Approximately 100 µg GST-fusion protein was injected subcutaneously into two Semi lop male rabbits. Boosting injections of approximately 100 µg of protein were carried out on 2 subsequent occasions, 3 weeks apart. Serum samples were taken 14 days following the final booster injection. The presence of anti-K7 antibodies was assayed in Cos-1 cells transiently transfected (2.6.7) with *K7* containing expression vectors by immunofluorescence (2.3.31) and western analysis (2.3.30) of Cos-1 cell extracts (2.3.27).

2.6 TISSUE CULTURE MATERIALS AND METHODS

2.6.1 Chemicals and Reagents

Chemicals and reagents for tissue culture were supplied as follows:

Roche:	FuGENE™ 6 Transfection reagent
Commonwealth Serum Laboratories:	FCS
Delta West:	Gentamicin
Gibco/BRL:	DMEM, Trypsin

Pharmacia:	Reduced Glutathione
Sigma:	ATP, β-mercaptoethanol, L-glutamine, Trypan blue, IMDM

2.6.2 Materials

Corning/Falcon:	6 well trays, 24 well trays, 6 cm and 10 cm dishes, 75 and 175 cm^2 flasks, 2 ml and 10 ml plastic pipettes
Biorad:	disposable cuvettes
Millipore:	0.22 μm filters

2.6.3 Cell Lines

D3 ES cells:	Mouse embryonic stem cell line was obtained form Dr. Lindsay Williams, Ludwig Institute, Melbourne, Australia
HepG2:	Human hepatocarcinoma cell line was obtained from ATCC (HB8065)
Cos-1:	SV-40 transformed African green monkey hepatocyte cell line was obtained from ATCC (CRL1650)
Baf-3:	Mouse pre B-cell line was obtained from Dr. Richard D'Andrea, Hanson Centre for Cancer Research, IMVS, Adelaide, Australia
CTLL:	Mouse T-cell line was obtained from Dr. Tom Gonda, Hanson Centre for Cancer Research, IMVS, Adelaide, Australia
FDC-P1:	Mouse myeloblast/early promyelocyte cell line was obtained from Dr. Richard D'Andrea, Hanson Centre for Cancer Research, IMVS, Adelaide, Australia
J2E-1:	Mouse pro-erythrocyte cell line was obtained from Dr Peter Klinken, University of Western Australia, Perth, Australia

J774: Mouse monocyte/macrophage cell line was obtained from Dr. Graham Myerhofer, Department of Microbiology and Immunology, University of Adelaide, Adelaide, Australia

P388D1: Mouse Lymphoid Neoplasm (Lymphoblast-like) cell line was obtained form ATCC (CCL46)

WEHI 3B D: Mouse myelomonocytic leukaemia cell line was obtained from Dr. Tom Gonda, Hanson Centre for Cancer Research, IMVS, Adelaide, Australia

2.6.4 Solutions

β-mercaptoethanol/PBS: 100 mM β-mercaptoethanol in 14 ml PBS. Fresh solution made every 2 weeks

Cytomix: 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4, 25 mM HEPES, 2 mM EGTA, 5 mM $MgCl_2$; pH to 7.6; 2 mM ATP and 5 mM reduced glutathione added shortly before use

Electroporation Buffer: 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 20 mM HEPES, 6 mM Dextrose; pH 7.05

L-Glutamine: 100 mM L-glutamine in MQ H_2O

PBS: 136 mM NaCl, 2.6 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 4.

PBS/gelatin: 0.2% (w/v) gelatin in PBS

Trypan blue: 0.4 g trypan blue, 0.06 g KH_2PO_4, in 100 ml MQ H_2O

Trypsin: 0.1% trypsin (w/v) and 1X Versene buffer

10X Versene: 80 g NaCl, 2 g KCl, 2 g EDTA, 2 g KH_2PO_4, 11.5 g N_2HPO_4

2.6.5 Media

ES DMEM: 67.4 g DMEM, 18.5 g $NaHCO_3$, dissolved in 5 litres sterile water, 6.25 ml gentamicin (40 mg/ml)

Incomplete ES cell medium:	84% ES DMEM, 15% FCS, 1% L-glutamine, 0.1 mM β-mercaptoethanol/PBS
Complete ES cell medium:	Incomplete ES cell medium with 1% Cos-1 cell conditioned medium containing LIF (Smith, 1991)
EPL cell medium:	50% ES cell medium (with or without LIF), 50% HepG2 conditioned medium (medium was isolated from HepG2 cells cultured in incomplete ES cell medium for 4-5 days and supplemented with 0.1 mM β-mercaptoethanol before use)
DMEM:	67.4 g DMEM (high glucose, with L-glutamine, Na-pyruvate, without $NaHCO_3$), 3.7 g $NaHCO_3$, 4.8 g HEPES, dissolved in 5L MQ H_2O pH 7.5, 1.25 ml gentamicin (40 mg/ml)
IMDM:	17.7 g powdered IMDM (with L-glutamine, 25 mM HEPES, without $NaHCO_3$) 3.024 g $NaHCO_3$, dissolved in 1 L MQ H_2O pH 7.4, 1.25 ml gentamicin (40 mg/ml)
Standard Culture Medium:	90% DMEM, 10% FCS
WEHI 3B Medium:	90% IMEM, 10% FCS, 0.1 mM β-mercaptoethanol
WEHI 3B Supplemented Medium:	80% DMEM, 10% FCS, 10% WEHI 3B conditioned medium

2.6.6 Maintenance of Cells Lines

ES Cells:

ES cells were routinely maintained on gelatinised 10 cm petri dishes in complete ES medium (2.6.5) at 37°C in 10% CO_2. Cells were passaged every 3-4 days by the following procedure. Cells were washed twice in PBS, incubated with trypsin (1 ml) for several mins. and transferred into 4 ml complete ES cell medium. The cells were spun at 1200 r.p.m. for 4 min., gently resuspended in 10 ml complete ES medium (2.6.5) and re-seeded at clonal density (20-40 fold dilution). A Heraeus Sepatch megacentrifuge was used for all tissue culture centrifugations.

EPL Cells:

ES cells were differentiated into EPL (early primitive ectoderm-like) by the addition of EPL cell media (2.6.5), in the presence or absence of LIF. EPL cells were grown at 37°C in 10% CO_2 and seeded at clonal density (5-20 fold dilutions) in EPL cell medium (with or without LIF) every 2 days as described above for ES cells.

HepG2 Cells:

HepG2 cells were maintained in incomplete ES medium (without β-Me; 2.6.5) at 37°C in 5% CO_2. Conditioned media (MEDII) was removed every 4-5 days and cells split following 2 washes in PBS and then incubation with 2 ml of trypsin for 5 min. Cells were dislodged from the flask and transferred into 8 ml of media, spun at 1200 r.p.m. for 4 min. and resuspended in 10 ml of media. Cells were generally seeded using 3-4 fold dilutions.

Cos-1 and J774 Cells:

Cells were maintained in standard culture media (2.6.5), grown at 37°C in 5% CO_2 and passaged every 3-4 days when cultures had reached near confluence. Cell passaging involved two washes with PBS and then trypsinisation for 5 min. at 37°C. Cells were dislodged from flask, transferred into standard culture media, and spun at 1200 r.p.m. for 4 min. Cell pellets were resuspended in 10 ml of media and re-seeded using 5-20 fold dilutions.

Suspension Cultures:

Suspension cultures were maintained at 37°C in 5% CO2 and were split using 3-6 fold dilutions when cell densities within the culture media reached near confluence, with the exception of the J2E-1 cell line which was maintained at cell densities between 1 x 10^5 and 1.5 x 10^6 cells. CTLL, J2E-1, and P388D1 cells were grown in standard culture media (2.6.5). WEHI 3B cells were maintained in WEHI 3B media (2.6.5) in 15 cm bacteriological grade petri dishes to prevent adherence. Baf-3 and FDC-P1 cells were grown in WEHI 3B supplemented media (2.6.5).

2.6.7 Transient Transfection of Cos-1 Cells

Two protocols for were used to transiently transfect Cos-1 cells:

i) Electroporation:

Three 175 cm^2 flasks of Cos-1 cells were grown to 90% confluence and trypsinised as described in section 2.4.6. Trypsinised cells were pooled in 30 ml tube and harvested by centrifugation at 1200 r.p.m. for 4 min. The cell pellet was resuspended in 20 mls of cold electroporation buffer (2.4.4). A 10 µl sample was taken and added to 90 µl trypan blue for cell count. Cells were pelleted, washed with 10 ml electroporation buffer, and re-pelleted. Cell pellets were resuspended in an appropriate volume of cytomix (2.6.4) to yield $1x10^7$ cells per ml. To a sterile electroporation cuvette, 50 µl of FCS and 50 µl of 10 mg/ml sonicated ssDNA were added to 10 µg of DNA. $5x10^6$ cells were added to the cuvette and allowed to stand at 4°C for 10 min. before electroporation at 0.27 V with a capacitance of 250 µFD. Cells were allowed to stand for a further 10 min. at RT before being plating into a 10 cm petri dish containing 10 ml of standard culture media (2.6.5) and sterilised cover slips for cell staining. Cells were grown for 24-72 hours.

ii) FuGENE™ 6

In order to gain higher transfection efficiencies, transient transfections were carried out using FuGENE™ 6 transfection reagent according to the manufacturer's instructions (Roche).

CHAPTER 3:

TRANSIENT PLURIPOTENT CELL SUB-POPULATIONS EXIST DURING PERI-IMPLANTATION MOUSE DEVELOPMENT

3.1 INTRODUCTION

Analysis of differential gene expression in the pluripotent cell pool between ES cells, an ICM/early epiblast *in vitro* equivalent, and EPL cells, an early primitive ectoderm-like cell population, using dd-PCR, lead to the identification of three novel potential pluripotent cell marker genes, *Psc1*, *L17* and *K7* (1.6.2). *In situ* hybridisation to monolayer cultures demonstrated that expression of these markers was restricted to pluripotent cell types *in vitro* (Schulz, 1996; S. Sharma, unpublished). Northern blot analysis demonstrated that *Psc1* and *L17* were expressed by ES and early passage EPL cells but down regulated in later passage EPL cells (Schulz, 1996; S. Sharma, unpublished; Rodda, 1998). *K7* was expressed at low levels in ES cells and up-regulated in EPL cell types following 4-6 days culture in MEDII in the presence of LIF (S. Sharma, unpublished). Preliminary *in vivo* expression analysis of *Psc1* by wholemount *in situ* hybridisation indicated that this gene was detectable in the pluripotent cells of the mid-late blastocyst stage (approximately 4.0-4.75 d.p.c.) and perhaps the epiblast bud at 5.0 d.p.c. (Schulz, 1996). *In vivo* expression analysis of *L17* and *K7* had not been undertaken.

The availability of differentially expressed marker genes with the capacity to identify pluripotent cell populations *in vivo* could potentially allow the resolution of important cellular and developmental questions. The existence of transient sub-populations within the pluripotent cell pool has been postulated following analysis of the developmental potency of pluripotent cells (1.4.1), gene expression analysis (1.4.2) and analysis of null mutations (1.4.3). While molecular evidence for their existence is not currently available, expression analysis of marker genes during peri-implantation development may allow the identification of temporally or spatially distinct sub-populations within the pluripotent cell pool of mouse embryogenesis. Correlation of marker gene expression with embryological events could provide markers for signals which influence pluripotent cell populations, perhaps allowing the identification of instructive and responsive cell types and deciphering signal transduction in the early embryo. The relationship between *in vivo* pluripotent cell sub-populations and the demonstrated pluripotent cell "states" identified *in vitro*, namely ES and EPL cells (Rathjen *et al.*, 1999), could be more precisely defined, allowing resolution of the relationship between the ICM to primitive ectoderm transition *in vivo* and the ES to EPL cell transition *in vitro*. Further analysis of the expression of these genes could also

provide markers for other pluripotent cell types of controversial origin, such as mouse EC and EG cells and human ES, EG and EC cells.

The purpose of this chapter therefore was to increase our knowledge of marker gene expression *in vivo* and *in vitro*. The three novel marker genes isolated during the dd-PCR screen of ES and EPL cells were selected for expression analysis; *L17* and *Psc1* as examples of genes which were likely to downregulate expression during peri-implantation development, and *K7* as an example of a gene likely to upregulate expression. Further expression mapping of *Fgf5* and *Rex1* to refine their expression profiles, was carried out as the timing of up and downregulation of these genes had not been mapped closely during peri-implantation development.

In vivo expression patterns of these markers were analysed by wholemount *in situ* hybridisation on mouse embryos at intervals of 0.25 of a day, from 3.5 d.p.c. to 5.5 d.p.c.. This enabled embryos to be timed according to morphologically distinguishable events, such as blastocyst expansion, proamniotic cavitation, and egg cylinder formation. To allow correlation with *in vitro* pluripotent cell progression, refined *in vitro* expression analysis of the marker genes was also undertaken by northern blot on RNA isolated from ES cells and EPL cells grown for 2, 4 and 6 days in MEDII in the presence and absence of LIF.

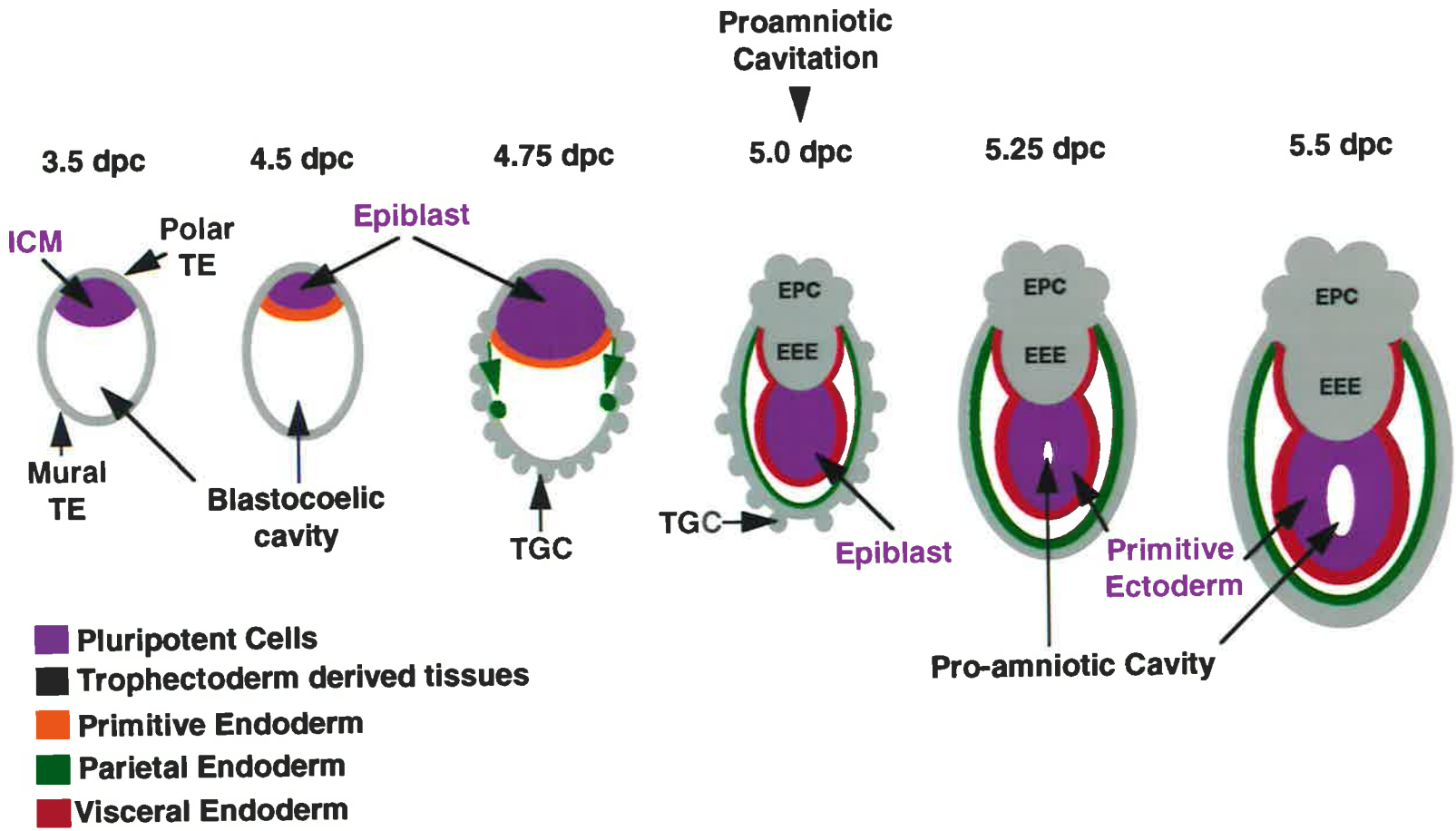
3.2 *IN VIVO* EXPRESSION ANALYSIS OF MARKER GENES DURING EARLY MOUSE EMBRYOGENESIS

3.2.1 Establishment of Morphologically Distinguishable Stages During Peri-implantation Development

The predicted range of marker gene expression, as indicated by the ES to EPL cell model of early development, encompassed the pluripotent cells of the blastocyst inner cell mass (ICM) (3.5 d.p.c.) to the cells of the primitive ectoderm (5.5 d.p.c.), when the egg cylinder has formed (Figure 3.1). Distinct embryo stages were determined on the basis of key cellular changes rather than timing since variation in developmental progression can be seen within and between litters. The ICM of the 3.5 d.p.c. blastocyst is partitioned to one end of the blastocyst following formation of the blastocoelic cavity. Primitive endoderm differentiates from ICM cells which line the blastocoelic cavity at approximately day 4.0 of embryogenesis. Following this differentiation event the blastocyst hatches from the zona pellucida and takes on an enlarged, ovoid morphology,

Figure 3.1: Schematic representation of peri-implantation mouse developmental stages from 3.5 to 5.5 d.p.c. as classified in this thesis

At 3.5 d.p.c. blastocysts consist of pluripotent cells termed the inner cell mass (ICM), which following the formation of the blastocoelic cavity, are displaced to one end of the blastocyst. The trophectoderm (TE) which overlays the ICM is termed polar trophectoderm and the trophectoderm in contact with the blastocoel is termed the mural trophectoderm. ICM cells lining the blastocoel differentiate into primitive endoderm at approximately 4.0 d.p.c. and following primitive endoderm differentiation, the pluripotent cells are termed the epiblast. At approximately 4.75 d.p.c. the mural trophectoderm of the late stage blastocyst has differentiated into trophoblast giant cells and the epiblast cell population has expanded in cell number. The primitive endoderm differentiates into either visceral endoderm, if endoderm cells remain in contact with the epiblast, or parietal endoderm if cells lose contact with the epiblast and migrate along the inner surface of the mural trophectoderm. Parietal endoderm forms the Reichert's membrane. The polar trophectoderm also proliferates and pushes the epiblast into the blastocoel, creating the 5.0 d.p.c. embryo, consisting of a solid mass of epiblast cells (epiblast bud) overlayed by the polar trophectoderm derived extra-embryonic ectoderm (EEE) and ectoplacental cone (EPC), and surrounded by visceral endoderm. Proamniotic cavitation begins around day 5.0 such that by day 5.25 of embryogenesis a small cavity is observable within the pluripotent cells, which in this thesis are termed primitive ectoderm. At 5.5 d.p.c. a larger proamniotic cavity has formed, and the primitive ectoderm cells become polarised and organise into a pseudostratified epithelial sheet.


Proamniotic Cavitation
3.5 dpc
4.5 dpc
4.75 dpc
5.0 dpc
5.25 dpc
5.5 dpc
ICM
Polar TE
Epiblast
Mural TE
Blastocoelic cavity
TGC
EPC
EEE
TGC
Epiblast
Primitive Ectoderm
Pro-amniotic Cavity
Pluripotent Cells
Trophectoderm derived tissues
Primitive Endoderm
Parietal Endoderm
Visceral Endoderm

termed the fully expanded blastocyst at approximately 4.5 d.p.c., in which the pluripotent cells have flattened and are termed the epiblast. Late stage blastocyst present at approximately 4.75 d.p.c. embryos were defined by two parameters; formation of primary trophoblast giant cells which contain polytene chromosomes and increased epiblast cell number. Proliferating polar trophectoderm projects down upon the epiblast cells, pushing the epiblast into the blastocoelic cavity to create an embryo structure at 5.0 d.p.c. consisting of an apolar mass of epiblast cells, termed in this thesis as the "epiblast bud", overlayed by the polar trophectoderm-derived extra-embryonic ectoderm. Programmed cell death results in the formation of a small proamniotic cavity visible by 5.25 d.p.c. within the apolar epiblast population. Following cavitation, the pluripotent cells become polarised and reorganise into a pseudostratified epithelial sheet of primitive ectoderm by day 5.5 of embryogenesis. These morphological changes were used to stage embryos, in conjunction with arbitrary time points, in which 12 noon on the day that the vaginal plug was discovered was designated 0.5 d.p.c.. Post-proamniotic cavitated embryos were also staged based on size and morphology of the proamniotic cavity.

3.2.2 *In situ* Hybridisation Analysis of the ICM to Primitive Ectoderm Transition using Established Marker Genes, *Oct4*, *Rex1* and *Fgf5*

In vivo expression of established marker genes, including the pan-specific marker of pluripotent cells, *Oct4* (Rosner *et al.*, 1990; Schöler *et al.*, 1990a and b; Yeom *et al.*, 1991), *Rex1,* which demonstrates expression in the pluripotent cells of the 4.5 d.p.c. blastocyst, but is not detected in the primitive ectoderm at 6.0 d.p.c. (Rogers *et al.*, 1991), and *Fgf5,* which is not expressed in the 3.5 d.p.c. blastocyst, but upregulated in the primitive ectoderm of the 5.25-5.5 d.p.c. embryo (Haub and Goldfarb, 1991; Hébert *et al.*, 1991), was determined by whole mount *in situ* hybridisation (2.4.6). Elucidation of *in vivo* expression patterns for *Rex1* and *Fgf5* was particularly important as previously published expression patterns during peri-implantation development were not mapped with respect to the precise up and downregulation of these genes, presumably due to experimental difficulties associated with this stage of embryogenesis. This analysis also allowed generation of a robust *in situ* hybridisation procedure for this early and fragile stage of mouse development.

Flushed pre- and early post-implantation embryos (up to approximately 4.75 d.p.c.), or dissected post-implantation embryos from 5.0-5.5 d.p.c., were examined for gene expression using DIG labelled *Oct4, Rex1 and Fgf5* antisense riboprobes (2.4.5). Probe hybridisations were carried out at 65°C for *Oct4* and *Rex1*. The *Fgf5* riboprobe demonstrated non-specific binding at this temperature and the hybridisation temperature was increased to 66-68°C with shaking in order to remove non-specific background staining and visualise *Fgf5* specific expression. Sense riboprobes were utilised as negative controls to demonstrate the specificity of hybridisation and to verify the conditions for whole mount *in situ* hybridisation (data not shown; Figure 3.4 F). The specificity of probes used in the *in situ* hybridisation protocol was further exemplified by the use of multiple antisense riboprobes which displayed unique expression profiles, the presence of positive and negative cells within each embryo and the presence of non-expressing embryological stages for each probe examined.

3.2.2.1 Oct4 Expression In vivo during Pre- and Early Post-implantation Mouse Development

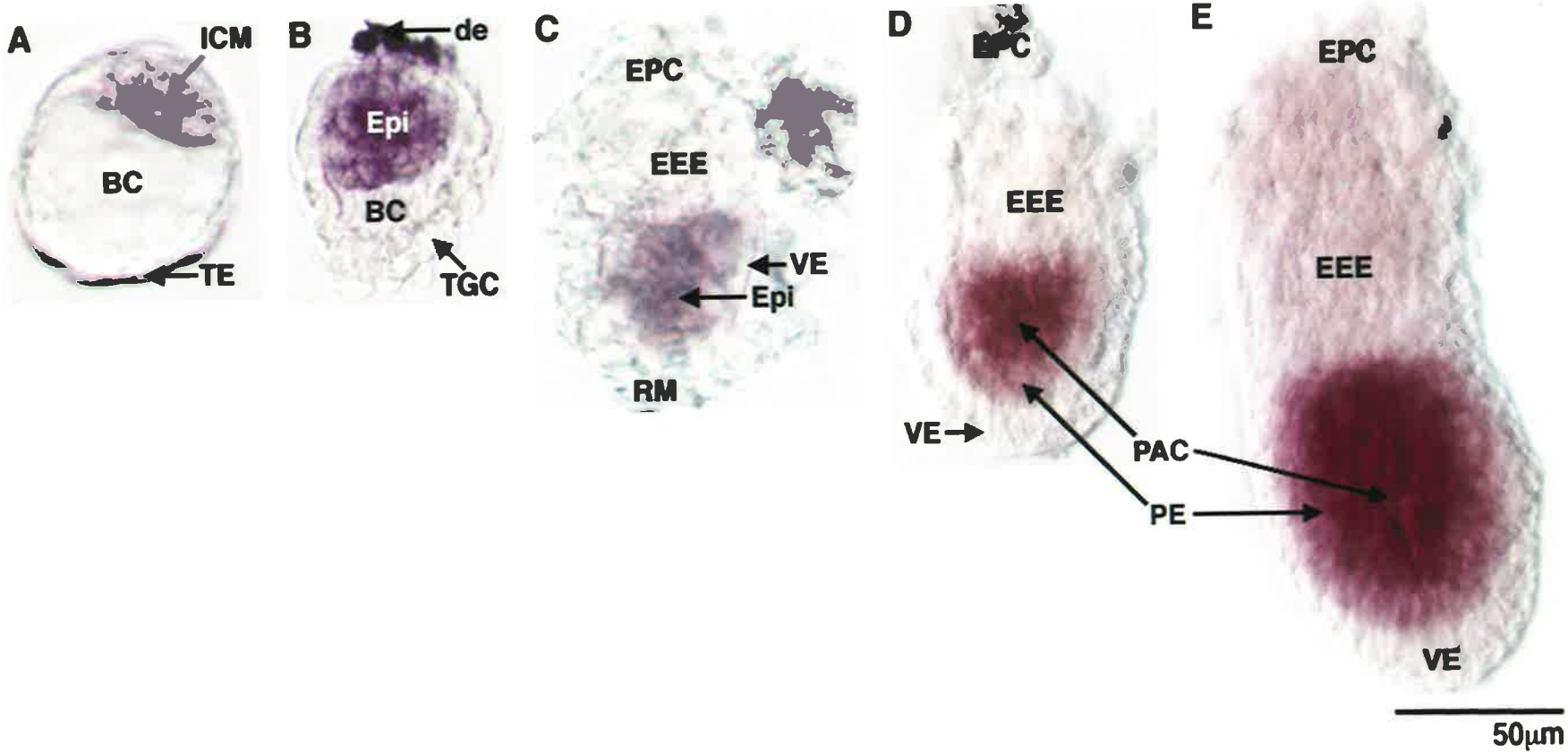
Investigation of *Oct4* expression during the ICM to primitive ectoderm transition allowed visualisation of pluripotent cell populations. As previously reported (Rosner *et al.*, 1990; Schöler *et al.*, 1990a and b; Yeom *et al.*, 1991), *Oct4* transcription was restricted to the pluripotent cell types within the early mouse embryo, with no expression detectable in trophectodermal derivatives (Figure 3.2 A, B and C) nor derivatives of the primitive endoderm (Figure 3.2 A, B and C). *Oct4* was expressed by the ICM of the 3.5-4.0 d.p.c. blastocyst (Figure 3.2 A), in the epiblast of the late stage blastocyst at 4.75 d.p.c. (Figure 3.2 B), in the epiblast bud at 5.0 d.p.c., following epiblast protrusion into the blastocoel (Figure 3.2 C), the early post-proamniotic cavitated epiblast (5.25 d.p.c.) (Figure 3.2 D), and in the pseudostratified epithelial sheet of primitive ectoderm at 5.5 d.p.c (Figure 3.2 E). Therefore, consistent with previous reports, *Oct4* was expressed by all pluripotent embryonic cells, and there was no evidence for spatial heterogeneity.

Figure 3.2: Expression of *Oct4* during early mouse development

Analysis of *Oct4* expression during peri-implantation mouse development by whole mount *in situ* hybridisation using a 462 bp DIG labelled riboprobe (2.4.5/6). *Oct4* expressing cells can be visualised by the purple staining which was developed for 1-2 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system.

A) 3.5-4.0 d.p.c. blastocyst; **B)** 4.75 d.p.c. late stage blastocyst; **C)** 5.0 d.p.c. "epiblast bud" embryo; **D)** 5.25 d.p.c. embryo in which a proamniotic cavity is just discernible; **E)** 5.5 d.p.c. egg cylinder stage embryo.

Abbreviations: ICM = inner cell mass; BC = blastocoelic cavity; TE = trophectoderm; Epi = Epiblast; TGC = trophoblast giant cell; de = non-specific debris; EEE = extra-embryonic ectoderm; EPC = ectoplacental cone; RM = Reichert's membrane; VE = visceral endoderm; PE = primitive ectoderm; PAC = proamniotic cavity.


A
ICM
BC
TE
B
de
Epi
BC
TGC
C
EPC
EEE
VE
Epi
RM
D
EPC
EEE
VE
PAC
PE
E
EPC
EEE
VE
50μm

3.2.2.2 Rex1 Expression In vivo during Pre- and Early Post-implantation Mouse Development

Rex1 was expressed by all cells of the ICM, but not the trophectoderm at 3.5 d.p.c. (Figure 3.3 A), and in the epiblast of the fully expanded blastocyst, but not the trophectoderm or primitive endoderm lineages at 4.5 d.p.c. (Figure 3.3 B). Variable *Rex1* expression was observed at 4.75 d.p.c., with *Rex1* expression detected in the epiblast of some, but not all embryos examined (Figure 3.3 Ci & ii). This was interpreted to mean that *Rex1* expression began to downregulate at this developmental stage. Low level expression was seen in the trophectoderm derived extra-embryonic ectoderm (Figure 3.3 D and E) following proamniotic cavitation at 5.25 and 5.5 d.p.c., but expression was not observed in pluripotent cells or visceral endoderm at these stages. These data are consistent with the expression pattern previously reported by Rogers *et al.* (1991), but refined the downregulation of *Rex1* expression in pluripotent cells to around 4.75 d.p.c., a day and a quarter earlier than that identified by Rogers and colleagues (1991), who did not examine embryos between day 4.5 and 6.0. However, in contrast to their report, *Rex1* expression was not detected in the trophectoderm at the blastocyst stage.

3.2.2.3 Fgf5 Expression In vivo during Pre- and Early Post-implantation Mouse Development

Fgf5 expression in pluripotent cells was not detected between 3.5 and 4.75 d.p.c., during blastocyst formation and maturation (Figure 3.4 A and B). *Fgf5* expression was observed in all pluripotent cells following the initiation of proamniotic cavitation (5.25-5.5 d.p.c.) but not in extra-embryonic tissues at this stage (Figure 3.4 D and E). In addition to the pluripotent cell expression, *Fgf5* transcripts were often detected at low levels in the primitive endoderm at 4.75 d.p.c. (Figure 3.4 C). Expression of *Fgf5* continued in the primitive ectoderm until the onset of gastrulation (data not shown). Expression analysis of *Fgf5* was consistent with previous reports, which have also documented *Fgf5* expression within the primitive streak, paraxial mesoderm and definitive endoderm following the onset of gastrulation (Haub and Goldfaub, 1991; Hébert *et al.*, 1991). However, the analysis undertaken in this thesis refined the upregulation of *Fgf5* expression to between day 5.0 and 5.25, instead of between day 3.5 and 5.25 as previously reported (Haub and Goldfaub, 1991; Hébert *et al.*, 1991).

Figure 3.3: Expression of *Rex1* in peri-implantation mouse development

Analysis of *Rex1* expression during peri-implantation mouse development by whole mount *in situ* hybridisation using an 848 bp DIG labelled riboprobe (2.4.5/6). *Rex1* expressing cells can be visualised by the purple/dark pink staining which was developed for 1-3 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system.

A) 3.5 d.p.c. blastocyst; **B)** 4.5 d.p.c. expanded blastocyst; **Ci) & Cii)** 4.75 d.p.c. late stage blastocysts; **D)** 5.25 d.p.c. embryo in which a proamniotic cavity is just discernible; **E)** 5.5 d.p.c. egg cylinder stage embryo.

Abbreviations: ICM = inner cell mass; BC = blastocoelic cavity; TE = trophectoderm; TGC = trophoblast giant cell; PEn = primitive endoderm; EEE = extra-embryonic ectoderm; EPC = ectoplacental cone; VE = visceral endoderm; PE = primitive ectoderm; PAC = proamniotic cavity.

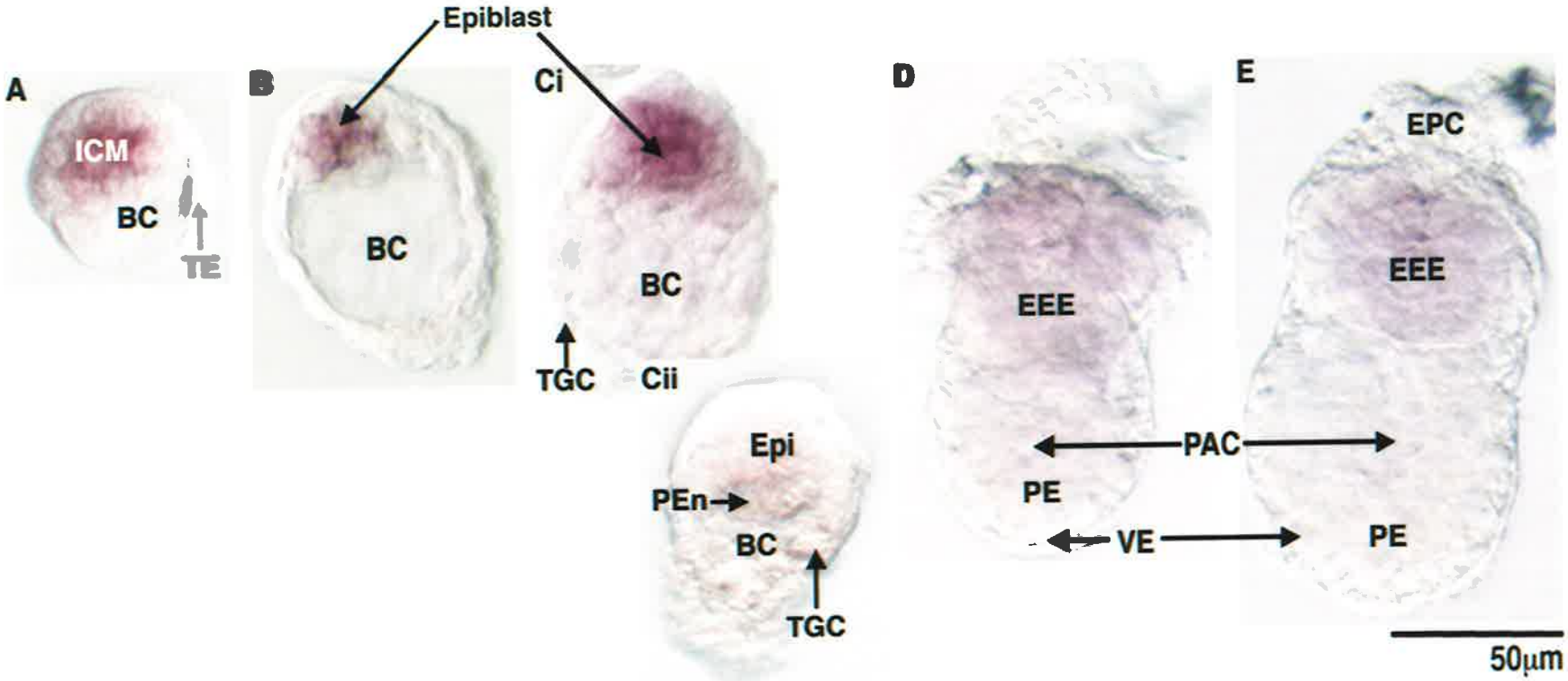
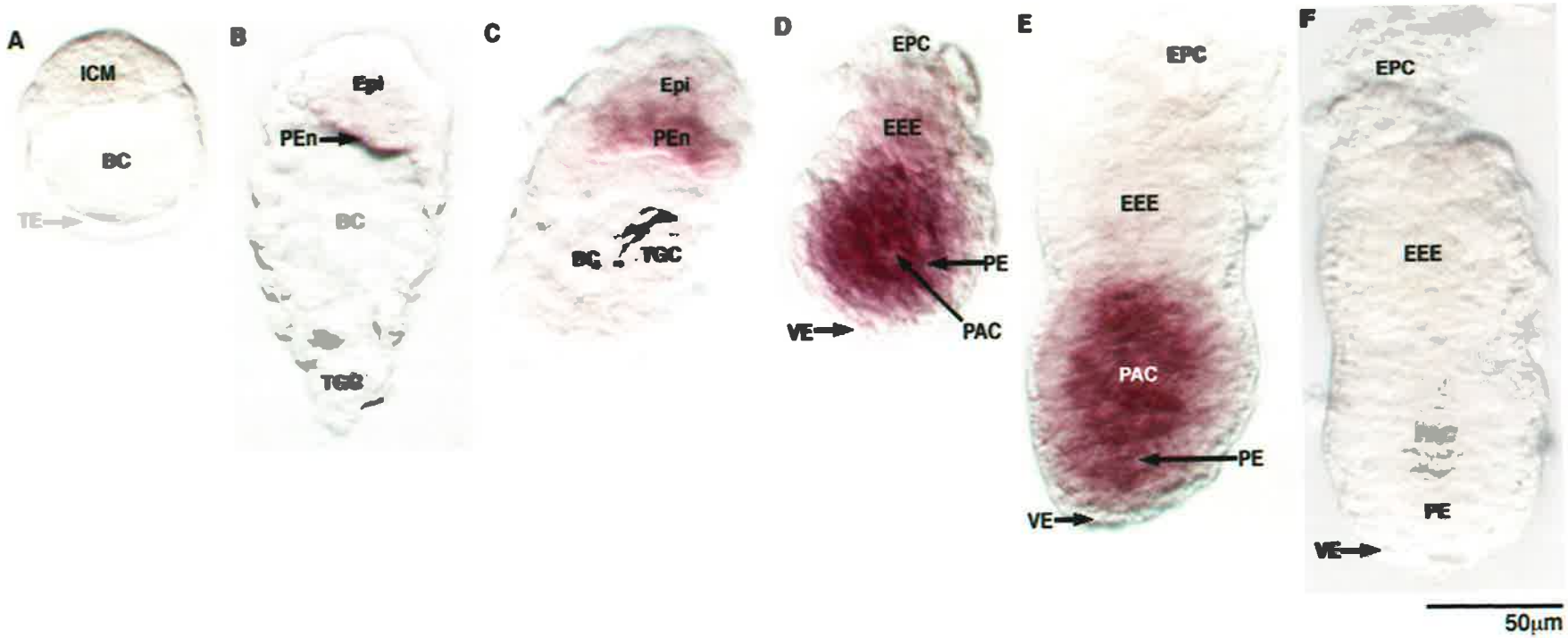

A
ICM
BC
TE
B
Epiblast
BC
Ci
BC
TGC
Cii
Epi
PEn
BC
TGC
D
EEE
PE
E
EPC
EEE
PAC
VE
PE
50μm

Figure 3.4: Expression of *Fgf5* in peri-implantation mouse development

Analysis of *Fgf5* expression during peri-implantation mouse development by whole mount *in situ* hybridisation using an 800 bp DIG labelled antisense riboprobe (**A-E**) or a sense riboprobe (**F**) (2.4.5/6). *Fgf5* expressing cells can be visualised by the purple/dark pink staining which was developed for 1-3 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system.

A) 3.5 d.p.c. blastocyst; **B**) 4.75 d.p.c. stage blastocyst; **C**) 4.75 d.p.c. late stage blastocyst with discernible primitive endoderm staining; **D**) 5.25 d.p.c. embryo containing a small proamniotic cavity; **E**) 5.5 d.p.c. egg cylinder stage embryo; **F**) 5.5 d.p.c. embryo hybridised with a *Fgf5* sense riboprobe

Abbreviations: ICM = inner cell mass; BC = blastocoelic cavity; TE = trophectoderm; TGC = trophoblast giant cell; PEn = primitive endoderm; Epi = epiblast; EEE = extra-embryonic ectoderm; EPC = ectoplacental cone; VE = visceral endoderm; PE = primitive ectoderm; PAC = proamniotic cavity.


A
ICM
BC
TE
B
Epi
PEn
BC
TGC
C
Epi
PEn
BC
TGC
D
EPC
EEE
PE
VE
PAC
E
EPC
EEE
PAC
PE
VE
F
EPC
EEE
PE
VE
50μm

3.2.3 *In situ* hybridisation Analysis of the ICM to Primitive Ectoderm Transition *In vivo* using Marker Genes Identified by dd-PCR Analysis of ES cells and EPL cells

3.2.3.1 L17 Expression In vivo during Pre- and Early Post-implantation Mouse Development

Pre-and early post-implantation embryos were subjected to wholemount *in situ* hybridisation analysis using DIG labelled sense and antisense *L17* riboprobes (2.4.5/6), generated from the *L17* dd-PCR isolated clone (2.2.6.2) which were hybridised to embryos at 66-68°C. Sense riboprobes did not display specific staining above background levels (Figure 3.5 G). *L17* expression was readily detectable in all ICM cells of 3.5 d.p.c. embryos, but not in the trophectoderm lineages (Figure 3.5 A). *L17* expression was maintained in the flattened epiblast of the expanded blastocyst stage at approximately 4.5 d.p.c., however, *L17* transcripts were not found in the primitive endoderm lineages, indicating that the expression of this gene was downregulated upon differentiation of the primitive endoderm from the ICM (Figure 3.5 B). *L17* expression appeared to downregulate within pluripotent cells at around 4.75 d.p.c., as expression of this transcript was variable at this stage, with some but not all 4.75 d.p.c. embryos expressing the transcript (Figure 3.5 Ci & ii). *L17* was not expressed in 5.0, 5.25 or 5.5 d.p.c. embryos (Figure 3.5 D-F).

In situ hybridisation analysis during early embryogenesis demonstrated similarity between expression of *L17* and *Rex1* expression. Expression of both genes was restricted to all pluripotent cells of the pre- and early post-implantation blastocyst, with expression downregulating within the epiblast following trophoblast giant cell formation, just prior to proamniotic cavitation. This indicated the possibility of coordinate regulation of these two genes.

3.2.3.2 Psc1 Expression In vivo during Pre- and Early Post-implantation Mouse Development

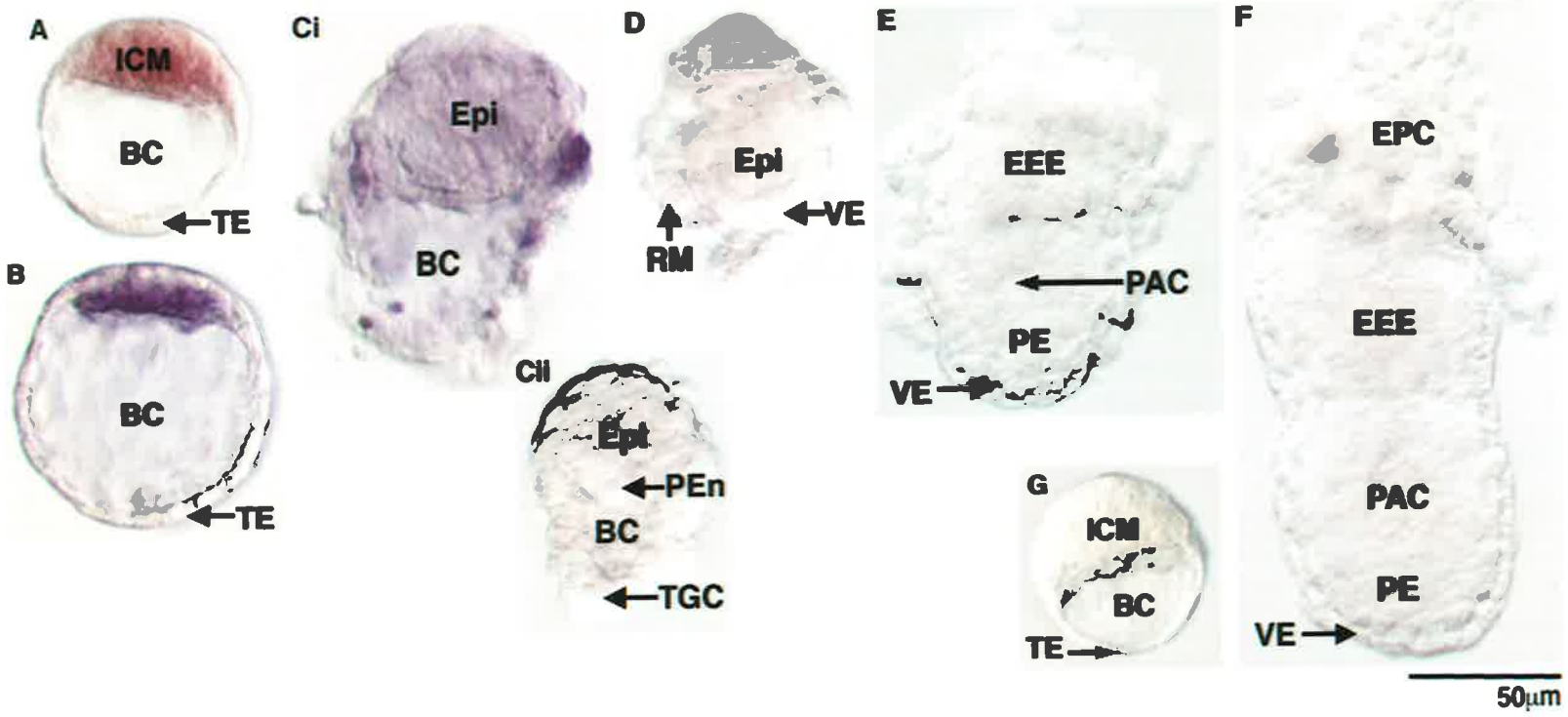
The expression of *Psc1* during early mouse embryogenesis was determined by wholemount *in situ* hybridisation using DIG labelled sense and antisense *Psc1* riboprobes (2.4.5/6), generated from the dd-PCR clone (2.2.6.2). Sense and antisense riboprobes were hybridised to embryos at 63.5°C, and sense riboprobes did not produce staining above background levels (Figure 3.6 G).

Figure 3.5: Expression of *L17* in peri-implantation mouse development

Analysis of *L17* expression during peri-implantation mouse development by whole mount *in situ* hybridisation using a 737 bp DIG labelled antisense riboprobe (**A-F**) or sense riboprobe (**G**) (2.4.5/6). *L17* expressing cells can be visualised by the purple staining which was developed for 1-3 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system.

A) 3.5 d.p.c. blastocyst; **B**) 4. 5 d.p.c. expanded blastocyst; **Ci) & Cii)** 4.75 d.p.c. late stage blastocysts; **D**) 5.0 d.p.c. embryo; **E**) 5.25 d.p.c. embryo containing a small proamniotic cavity; **F**) 5.5 d.p.c. egg cylinder stage embryo; **G**) 3.5 d.p.c. blastocyst probed with a sense riboprobe.

Abbreviations: ICM = inner cell mass; BC = blastocoelic cavity; TE = trophectoderm; TGC = trophoblast giant cell; PEn = primitive endoderm; Epi = epiblast; EEE = extra-embryonic ectoderm; EPC = ectoplacental cone; VE = visceral endoderm; RM = Reichert's membrane; PE = primitive ectoderm; PAC = proamniotic cavity.


A
ICM
BC
TE
B
BC
TE
Ci
Epi
BC
Cii
Epi
PEn
BC
TGC
D
Epi
VE
RM
E
EEE
PAC
PE
VE
G
ICM
BC
TE
F
EPC
EEE
PAC
PE
VE
50µm

Psc1 expression was overlapping, but distinct from that of *L17* within the pluripotent cell pool. *Psc1* expression was detectable in all ICM cells of the developing blastocyst, but not the trophectoderm at 3.5 d.p.c. (Figure 3.6 A), and was maintained in the flattened epiblast of the fully expanded blastocyst at 4.5 d.p.c. (Figure 3.6 B). *Psc1* transcripts were also expressed at high levels throughout the epiblast cells at 4.75 d.p.c. (Figure 3.6 C), and detectable at variable levels in epiblast bud at 5.0 d.p.c., immediately prior to proamniotic cavity formation (Figure 3.6 D), but not detectable in the primitive endoderm or derivatives. Embryos in which the process of proamniotic cavitation had initiated (approximately 5.25 d.p.c.) displayed downregulated *Psc1* expression (Figure 3.6 E). *Psc1* was not detected in pluripotent cells in later stage embryos which contained larger proamniotic cavities (5.5 d.p.c.) (Figure 3.6 F).

3.2.3.3 K7 Expression In vivo during Pre- and Early Post-implantation Mouse Development

The expression of *K7* during early mouse embryogenesis was determined by wholemount *in situ* hybridisation using DIG labelled sense and antisense *K7* riboprobes (2.4.5/6), generated from the dd-PCR clone (2.2.6.2). Sense riboprobes did not produce any positive staining (data not shown). Antisense riboprobes produced a gene expression profile for *K7* which displayed an alternate pattern to of *L17* and *Psc1*, with expression not detected in cells of the ICM at 3.5 d.p.c. (data not shown) or the early epiblast cell population at 4.5 d.p.c. (Figure 3.7 A). Variable *K7* expression was observed at 4.75 d.p.c., with low level expression detected in the epiblast of some late stage blastocysts, but not all (Figure 3.7 Bi and ii). This expression pattern was interpreted to mean upregulation of the *K7* transcript at around 4.75 d.p.c.. Higher level *K7* expression was observed within all cells of the epiblast bud and in the apparent overlying extra-embryonic ectoderm, but not the ectoplacental cone, or visceral endoderm at 5.0 d.p.c. (Figure 3.7 C). Immediately following proamniotic cavitation, when the cavity first becomes visible (5.25 d.p.c.), *K7* transcripts were detectable in all pluripotent cells but not in the visceral endoderm or extra-embryonic ectoderm (Figure 3.7 D). Staining was sometimes observed in the ectoplacental cone, however, this region was often removed from post-implantation embryos during dissection and as a result the importance of this staining, if any, was not determined. *K7* expression within pluripotent cells was downregulated as primitive ectoderm cells become

Figure 3.6: Expression of *Psc1* in peri-implantation mouse development

Analysis of *Psc1* expression during peri-implantation mouse development by whole mount *in situ* hybridisation using a 458 bp DIG labelled antisense riboprobe (**A-F**) or sense riboprobe (**G**). *Psc1* expressing cells can be visualised by the purple/dark pink staining which was developed for 30 min. to 2 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system.

A) 3.5 d.p.c. blastocyst; **B**) 4.5 d.p.c. expanded blastocyst; **C**) 4.75 d.p.c. late stage blastocyst; **D**) 5.0 d.p.c. "epiblast bud" embryo; **E**) 5.25 d.p.c. embryo in which a proamniotic cavity is just discernible; **F**) 5.5 d.p.c. egg cylinder stage embryo; **G**) 4.75 d.p.c. embryo probed with a sense riboprobe.

Abbreviations: ICM = inner cell mass; BC = blastocoelic cavity; TE = trophectoderm; Epi = Epiblast; TGC = trophoblast giant cell; PEn = primitive endoderm; EEE = extra-embryonic ectoderm; EPC = ectoplacental cone; VE = visceral endoderm; PE = primitive ectoderm; PAC = proamniotic cavity.

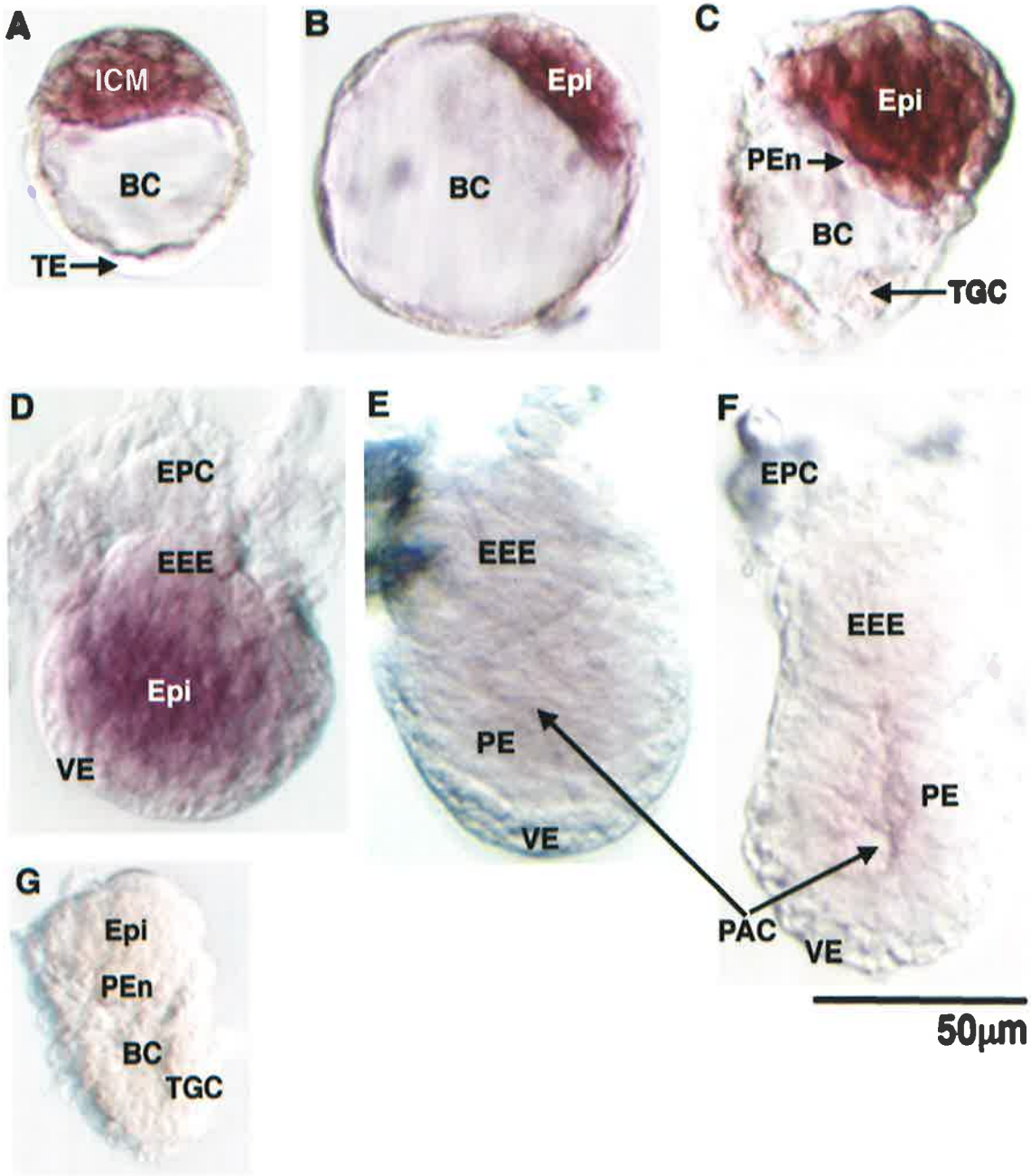
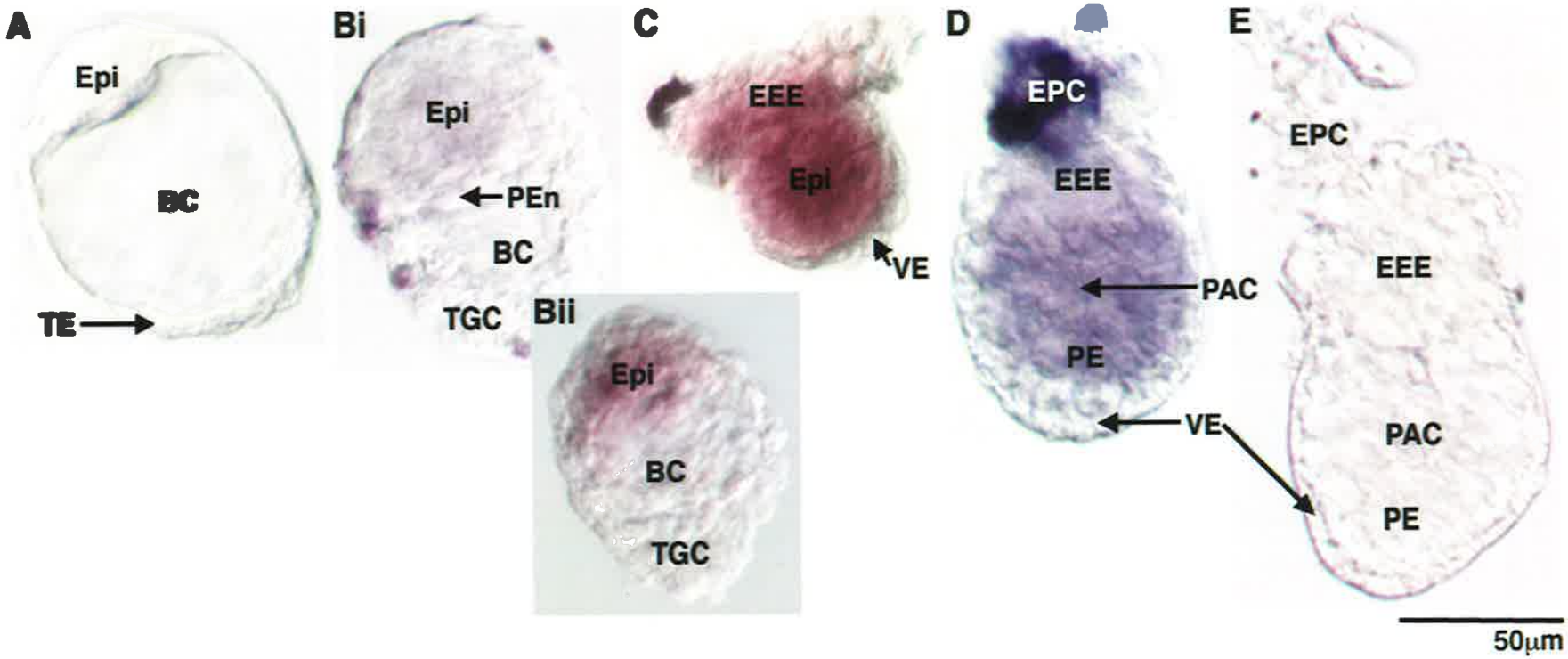

A
ICM
BC
TE
B
Epi
BC
C
Epi
PEn
BC
TGC
D
EPC
EEE
Epi
VE
E
EEE
PE
VE
F
EPC
EEE
PE
PAC
VE
G
Epi
PEn
BC
TGC
50μm

Figure 3.7: Expression of *K7* in peri-implantation mouse development

Analysis of *K7* expression during peri-implantation mouse development by whole mount *in situ* hybridisation using a 232 bp DIG labelled riboprobe (2.4.5/6). *K7* expressing cells can be visualised by the purple/dark pink staining which was developed for 1-4 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system.

A) 4.5 d.p.c. expanded blastocyst; **Bi & Bii)** 4.75 d.p.c. late stage blastocysts; **C)** 5.0 d.p.c. "epiblast bud" embryo; **D)** 5.25 d.p.c. embryo in which a proamniotic cavity is just discernible; **E)** 5.5-5.75 d.p.c. egg cylinder stage embryo.

Abbreviations: Epi = epiblast; BC = blastocoelic cavity; TE = trophectoderm; TGC = trophoblast giant cell; PEn = primitive endoderm; EEE = extra-embryonic ectoderm; EPC = ectoplacental cone; VE = visceral endoderm; PE = primitive ectoderm; PAC = proamniotic cavity.


A
Epi
BC
TE
Bi
Epi
PEn
BC
TGC
Bii
Epi
BC
TGC
C
EEE
Epi
VE
D
EPC
EEE
PAC
PE
VE
E
EPC
EEE
PAC
PE
50μm

polarised and reorganised into the columnar epithelial cell layer within the egg cylinder of 5.5-5.75 d.p.c. (Figure 3.7 E).

3.2.4 Summary

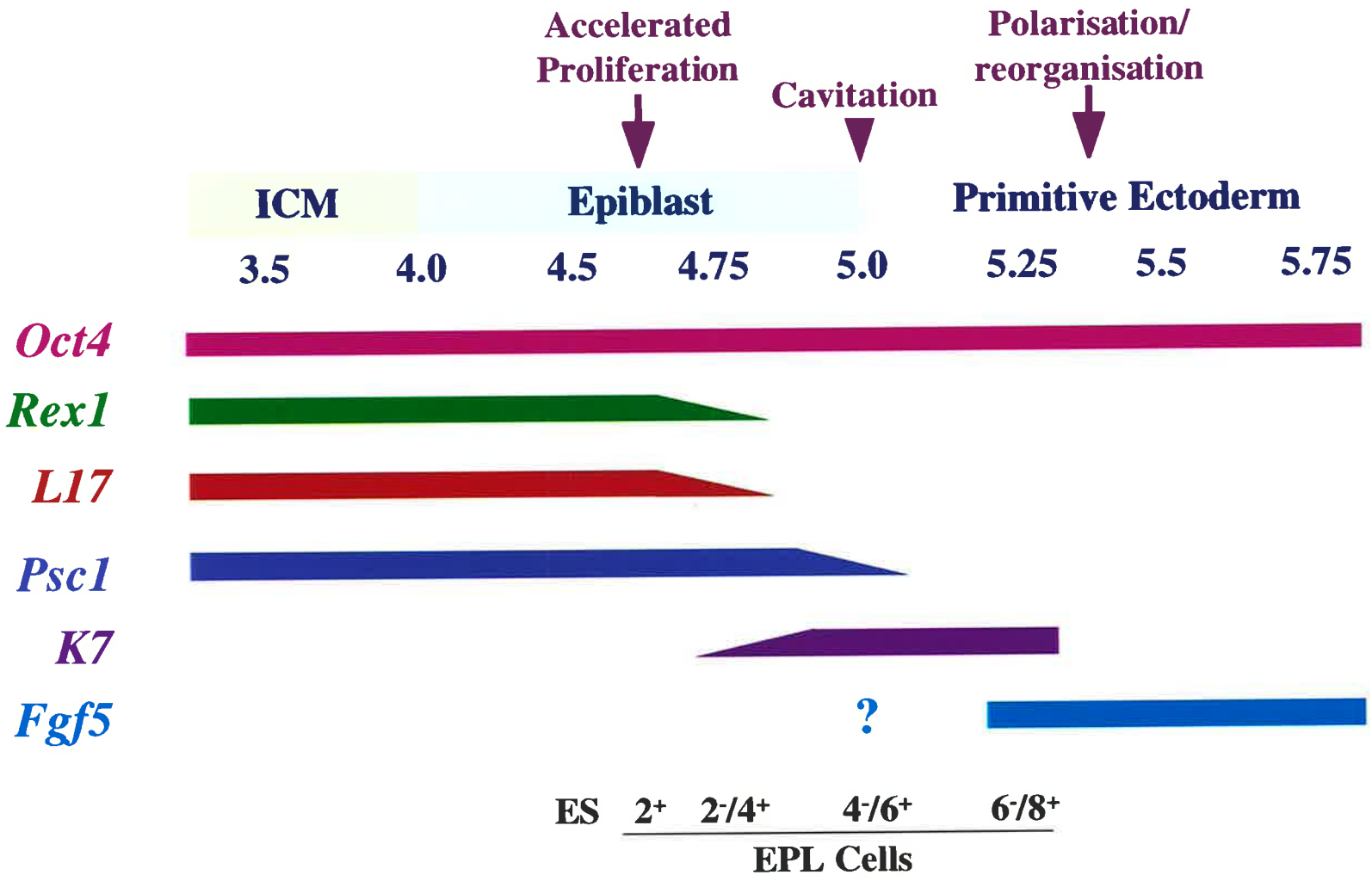
Compilation of *in vivo* expression profiles identified transient sub-populations within the pluripotent cells of the early mouse embryo that could be distinguished on the basis of gene expression (Figure 3.8). Pluripotent cells progressed from a *Rex1*$^+$/*L17*$^+$/*Psc1*$^+$/*K7*$^-$/*Fgf5*$^-$ sub-population during blastocyst formation and maturation, to a sub-population present from approximately 4.75-5.0 d.p.c., which had downregulated *Rex1* and *L17,* continued to express *Psc1* and begun to activate expression of *K7*, but not *Fgf5* (*Rex1*$^-$/*L17*$^-$/*Psc1*$^+$/*K7*$^+$/*Fgf5*$^-$). The epiblast bud population progressed to a later sub-population at 5.25 d.p.c., following downregulation of *Psc1* and upregulation of *Fgf5* (*Rex1*$^-$/*L17*$^-$/*Psc1*$^-$/*K7*$^+$/*Fgf5*$^+$). Following egg cylinder formation at 5.5-5.75 d.p.c. *K7* was downregulated but *Fgf5* expression was maintained, demarcating a further sub-population within the pluripotent cell pool (*Rex1*$^-$/*L17*$^-$/*Psc1*$^-$/*K7*$^-$/*Fgf5*$^+$). Interestingly, the regulation of these genes appeared to correlate with embryological events. *Rex1* and *L17* downregulated expression and *K7* upregulated expression at 4.75 d.p.c., suggesting a change in the developmental program at this time, perhaps involving increased epiblast proliferation (see 3.4.2). *Psc1* expression was downregulated between 5.0-5.25 d.p.c., correlating with proamniotic cavity formation. The downregulation of *K7* expression at 5.5 d.p.c. coincided with organisation of primitive ectoderm into a pseudostratified, polarised epithelium. In contrast to temporal progression, no evidence for spatial heterogeneity was observed within pluripotent cell populations of the early mouse development as marker genes were expressed in an apparently uniform manner within pluripotent cells.

3.3 *IN VITRO* EXPRESSION ANALYSIS OF MARKER GENES DURING THE ES TO EPL CELL TRANSITION

In vivo expression analysis of marker genes demonstrated that each gene displayed a distinct expression profile, and consequently could be used to identify distinct stages of pluripotent cell development *in vivo*. However, the ability to relate these embryonic expression profiles to the *in vitro* expression patterns of Schulz (1996) and S. Sharma (unpublished) was

Figure 3.8: Summary of marker gene expression in the early mouse embryo

Schematic representation of marker gene expression during early mouse embryogenesis. The pluripotent cell populations present within the early mouse embryo are aligned with approximate time points in days *post coitum* (d.p.c.) and the embryonic events which occur during this developmental stage. The likely embryonic equivalents of ES cells and EPL cells grown for 2, 4 and 6 days in culture in the presence (+) or absence (-) of LIF are also indicated at the bottom of the diagram.


Accelerated Proliferation
Cavitation
Polarisation/ reorganisation
ICM
Epiblast
Primitive Ectoderm
3.5
4.0
4.5
4.75
5.0
5.25
5.5
5.75
Oct4
Rex1
L17
Psc1
K7
Fgf5
?
ES
2+
2-/4+
4-/6+
6-/8+
EPL Cells

limited as a consistent EPL cell preparation was not used during these analyses and different EPL cell cultures can develop at variable rates. Direct comparison of marker gene expression *in vivo* and *in vitro* required comparative analysis of gene expression during EPL cell formation *in vitro*.

Direct comparison of the pluripotent cell marker genes *Oct4*, *Rex1*, *Fgf5*, *L17*, and *Psc1* was carried out using northern blot analysis on 4 μg poly-$(A)^+$ RNA isolated from ES cells and EPL cells, grown using the standard culture conditions of Rathjen *et al.* (1999), for 2, 4 and 6 days in the presence or absence of LIF. The filter was sequentially probed for *L17* and *Psc1*, *Fgf5*, *Rex1*, *Oct4* and mouse *glyceraldehyde phosphate dehydrogenase* (*mGap*). Expression patterns are shown in Figure 3.9, and quantitation of expression levels by volume integration (2.3.25), normalised to *Oct4* and *mGap* to account for differences in spontaneous differentiation within cultures and loading differences respectively, is shown in Figure 3.10. *Oct4* expression is indicative of the proportion of pluripotent cells in the culture, and was expressed at a relatively consistent level in pluripotent cells during EPL cell formation and culturing, consistent with previous reports (Rathjen *et al.*, 1999).

Consistent with *in vivo* expression, *Rex1* and *L17* demonstrated a similar expression profile *in vitro*, with highest expression observed in ES cells, and expression downregulating following formation and passaging of EPL cells. *Rex1* and *L17* expression could not be detected after 4 days of EPL cell culture in the absence of LIF, or 6 days in the presence of LIF. *Psc1* also demonstrated high ES cell expression, which was gradually downregulated following EPL cell formation and passaging. Quantitation of *Psc1* expression levels demonstrated 65.8% and 50.1% of ES cell expression by EPL cells grown for 2 days in the presence and absence of LIF respectively, compared with *L17* expression of 49% and 42% respectively. *Psc1* expression following 4 days of culture in the presence and absence of LIF was still detectable in EPL cell cultures at 64% and 22% of ES cell expression respectively, compared with 46% and 9.7% of ES cell expression for *L17*. *Psc1* expression was 12.5% and 14.6% of ES cell expression following 6 days of EPL cell culture in the presence and absence of LIF respectively, compared with 5.9% and 5.3% of ES cell expression for *L17* in the same cultures. Therefore, similar to its *in vivo* expression pattern, *Psc1* expression persisted longer than that of *Rex1* and *L17* (Figure 3.10) during the ES to EPL cell transition. *Fgf5* expression during EPL cell formation and progression *in vitro* was consistent with previous findings (Rathjen *et al.*, 1999). *Fgf5* transcripts could not

Figure 3.9: Analysis of pluripotent cell marker expression in ES cells and EPL cells

Poly-$(A)^+$ RNA was isolated from ES cells grown for 3 days in culture in the presence of LIF, and EPL cells grown for 2, 4 and 6 days in culture in the presence of MEDII, with (+) or without (-) exogenous LIF (2.6.6). 4 µg of poly-$(A)^+$ RNA was analysed by northern blot (2..3.22) and the filter sequentially probed using a DNA probe for *L17* and *Psc1*, *Fgf5*, *Rex1*, *Oct4* and finally *mGap*. The expression of each gene was determined by phosphorimager analysis by exposing filters overnight in phosphorimager cassettes. Transcript sizes were: *Rex1*, 1.9 kb; *L17*, 9.3 kb; *Psc1*, 5.5, 3.7 and 3.5 kb; *Fgf5*, 2.7 kb; *Oct4*, 1.6 kb and *mGap*, 1.5 kb.

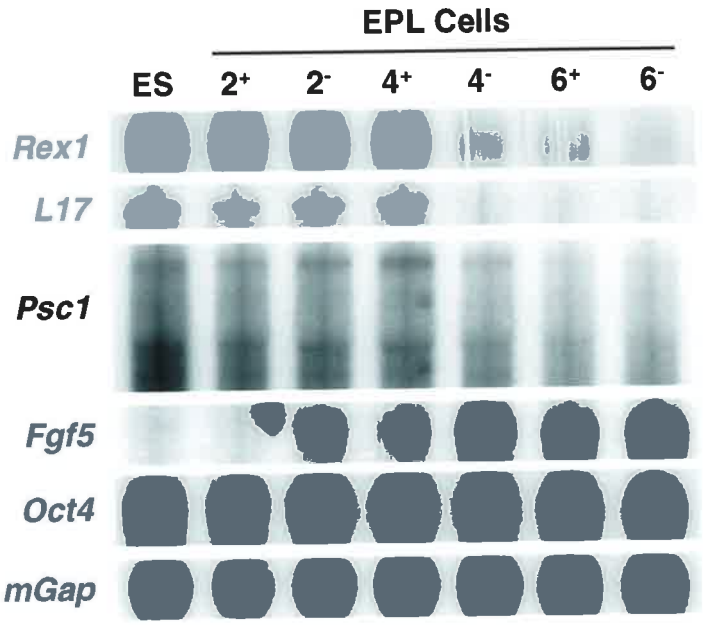
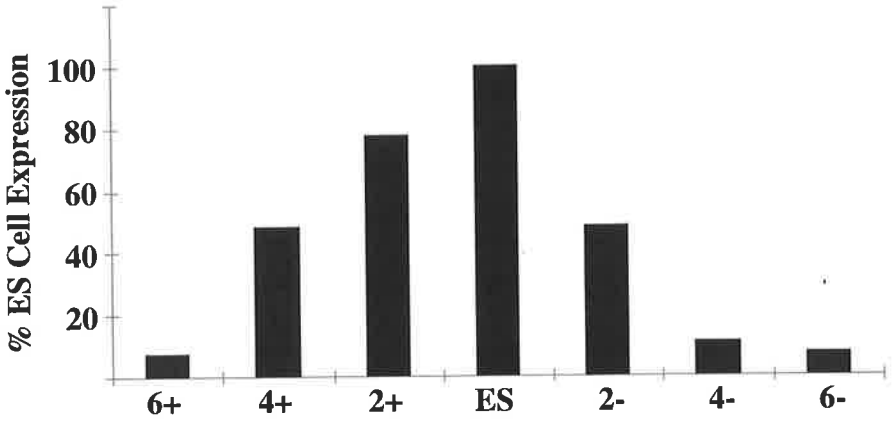

EPL Cells
ES
2+
2-
4+
4-
6+
6-
Rex1
L17
Psc1
Fgf5
Oct4
mGap

Figure 3.10: Comparison of *Rex1*, *L17*, *Psc1* and *Fgf5* expression in ES cells and EPL cell

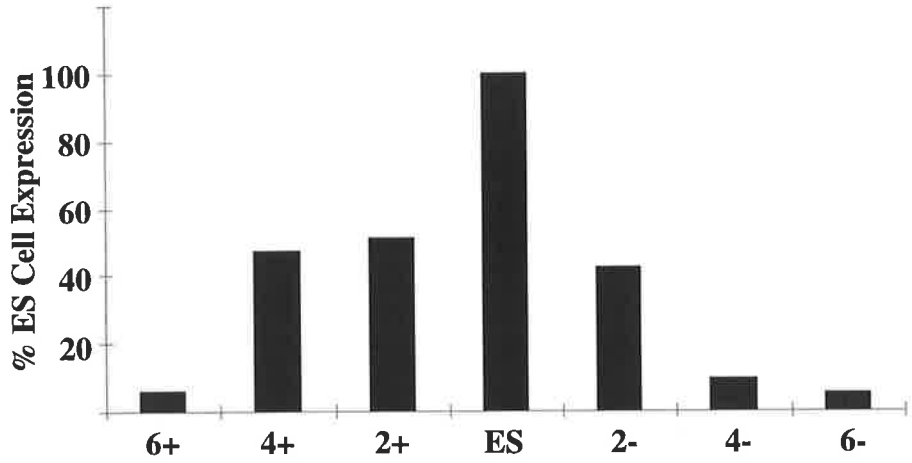
A B C D

Graphical representation of *Rex1*, *L17*, *Psc1* and *Fgf5* expression in ES and EPL cells, as analysed by northern blot (Figure 3.9). The expression level of each gene was quantitated by volume integration (the 5.5 kb transcript was quantitated in the case of *Psc1*), and normalised against *Oct4* and *mGap* to account for any differences in pluripotent cell number within individual cultures or loading differences respectively. *Rex1*, *L17* and *Psc1* expression in ES cells was assigned 100%, and the expression level of each gene in EPL cell cultures depicted as a percentage of ES cell expression. *Fgf5* expression in ES cells was assigned 100%, and *Fgf5* expression in EPL cell cultures depicted as a percentage above ES cell expression.

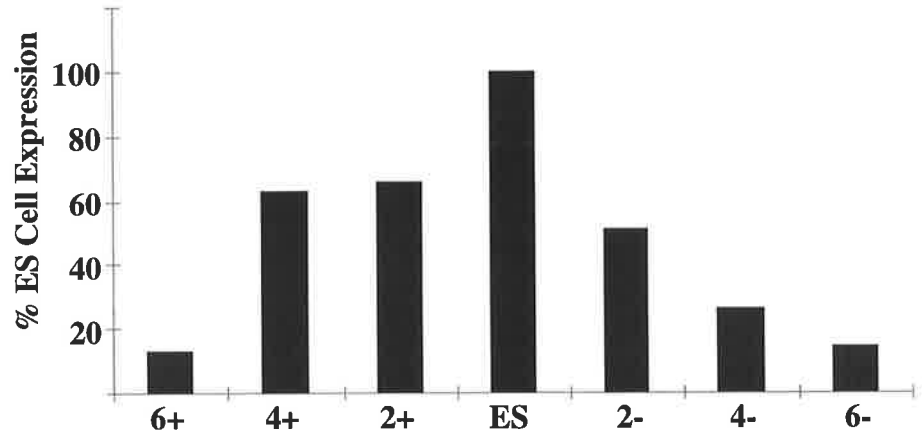
A


% ES Cell Expression
100
80
60
40
20
6+
4+
2+
ES
2-
4-
6-

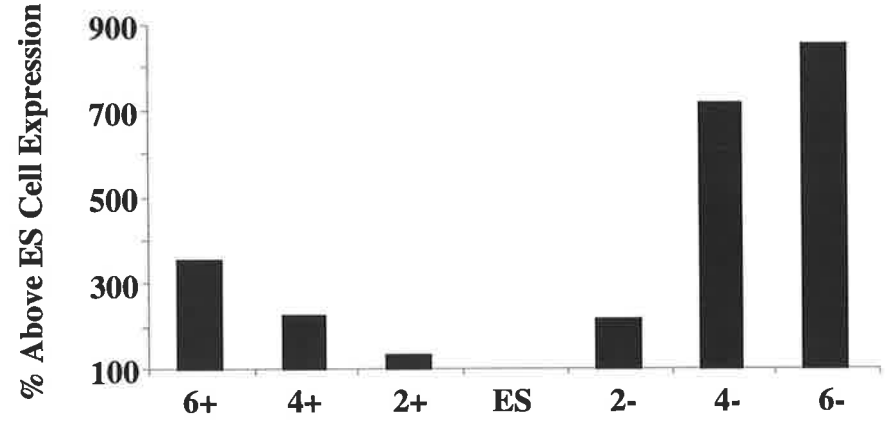
B


% ES Cell Expression
100
80
60
40
20
6+
4+
2+
ES
2-
4-
6-

C


% ES Cell Expression
100
80
60
40
20
6+
4+
2+
ES
2-
4-
6-

D


% Above ES Cell Expression
900
700
500
300
100
6+
4+
2+
ES
2-
4-
6-

be detected in ES cells and EPL cells grown for 2 days in the presence of LIF. *Fgf5* expression was detected following 2 days of culture in the absence of LIF and 4 days of culture in the presence of LIF. *Fgf5* expression continued to increase as EPL cells were cultured with highest expression observed after 6 days of culturing in the absence of LIF.

K7 expression within ES and EPL cells had been analysed using dd-PCR and verified using northern analysis (S. Sharma, unpublished). By dd-PCR, a band corresponding to *K7* was detected at low levels in ES cells and EPL cells grown for 2 days in the presence or absence of LIF or for 4 days in the presence of LIF, and detected at higher levels in EPL cells grown for 4 days in the absence of LIF or 6 days in the presence or absence of LIF (Figure 3.11 A). By northern blot, *K7* was detected at similar levels in ES cells and EPL cells grown in culture for 2 days in the presence of LIF, was upregulated to an intermediate expression level following EPL cell culture for 4 days in the presence of LIF, and upregulated two fold following *in vitro* culture of EPL cells for 6 days in the presence of LIF. Expression was maintained in EPL cells grown for 8 days in the presence of LIF (Figure 3.11 B and C). Therefore, low level *K7* expression in ES and early EPL cells was present and expression upregulated following 4 days of EPL cell culture. Slightly variable rates of EPL cell progression in culture may account for upregulated *K7* expression by northern analysis in EPL cells grown for 4 days in the presence of LIF compared with dd-PCR where upregulation of K7 expression was not observed until day 4 of culture in the absence of LIF. In contrast to the *in vivo* expression of *K7*, the transcript was not downregulated during EPL cell culture.

The downregulation of *Lf7*, *Rex1* and *Psc1* transcripts, and the upregulation of *Fgf5* and *K7* was accelerated in EPL cell cultures grown in the absence of LIF such that EPL cells grown in the absence of LIF demonstrated similar marker gene expression to EPL cells grown for 2 extra days in the presence of LIF. This was consistent with previous findings, and has been attributed to inhibition of the developmental progression of EPL cells in culture by LIF (Rathjen *et al.*, 1999).

3.4 DISCUSSION

Circumstantial evidence implied the existence of unrecognised temporal heterogeneity within the *Oct4*$^{+}$ pool of early mouse embryogenesis. Analysis of developmental potential of early

Figure 3.11: Analysis of *K7* expression in ES and EPL cells *in vitro*

A) Analysis of *K7* expression *in vitro*, during the ES cell to EPL cell transition, by dd-PCR (S. Sharma, unpublished). 2 μg of cytoplasmic RNA isolated from ES cells and EPL cells grown for 2-6 days in the presence (+) or absence (-) of LIF (2.6.6) was reverse transcribed with an internal 3'-oligo d(T) primer, before dilution 1/20 for dd-PCR, using OPA-01 5' primer (Operon technologies inc.) and the 3'-oligo d(T) primer. The PCR reaction was carried out in the presence of (^{33}P) dATP. Reactions were resolved by electrophoresis on a 6% (v/v) denaturing polyacrylamide gel prepared from 40% sequagel. The *K7* band is indicated by the arrow.

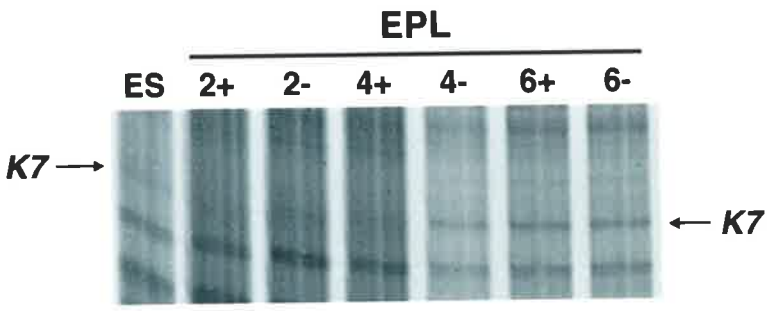
5' OPA-01 primer: CAGGCCCTTC

3' oligo d(T) primer: TTTTTTTTTTTCA

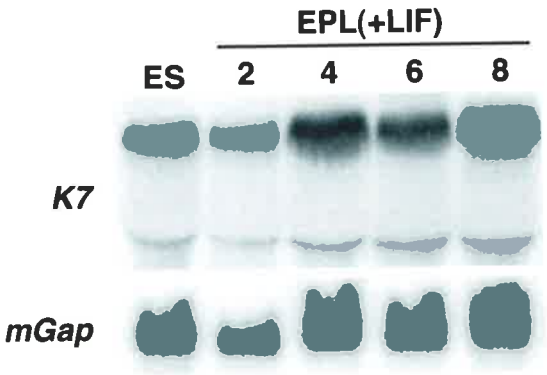
B) Northern analysis of *K7* during the ES cell to EPL cell transition (S, Sharma, unpublished). 25 μg of cytoplasmic RNA isolated from ES cells and EPL cells grown for 2-8 days in culture in the presence of LIF (+LIF), was electrophoresed on a 1% agarose gel and transferred to nylon membrane. The filter was hybridised overnight with a 232 bp *K7* DNA probe (2.3.23/24) prior to washing and exposure for 3 days in a phosphorimager cassette. The filter was subsequently stripped and re-probed with *mGap* as a general loading control. Transcript size for *K7* was 7 kb.

C) Graphical representation *K7* expression in ES and EPL cells, as analysed by northern blot (B). The expression level of *K7* were quantitated by volume integration (2.3.25) and normalised against *mGap* expression. *K7* expression in ES cells was assigned 100%, and the expression level in EPL cell culture depicted as a percentage of ES cell expression.

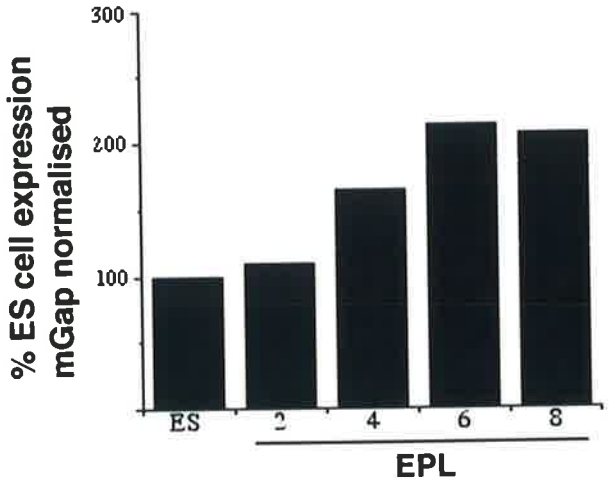
A


EPL
ES
2+
2-
4+
4-
6+
6-
K7
K7

B


EPL(+LIF)
ES
2
4
6
8
K7
mGap

C


% ES cell expression
mGap normalised
300
200
100
ES
2
4
6
8
EPL

pluripotent cell lineages revealed that competence of pluripotent cells to form extra-embryonic lineages was not lost with the initial differentiation of these tissues, rather there was a gradual restriction in the developmental potential for pluripotent cells to form these tissues (see section 1.4.1). Compilation of a large number of gene ablation experiments revealed the existence of pluripotent cell defects at different stages of peri-implantation development (see section 1.4.3). This suggested that pluripotent cell developmental progression following implantation is a dynamic process, in which temporally and spatially localised signals regulate diverse aspects of pluripotent cell behaviour, including proliferation, proamniotic cavitation, cell survival and primitive ectoderm organisation. Gene expression analysis also provided a glimpse into pluripotent cell developmental progression (see section 1.4.2), but genes had not been mapped in sufficient detail to allow molecular description of transient, intermediate pluripotent cell sub-populations which had been postulated. Potential markers were generated by the analysis of differentially expressed transcripts during the ES cell to EPL cell transition *in vitro* (Schulz, 1996; S. Sharma, unpublished) and investigated as candidates for differentially expressed markers *in vivo*.

3.4.1 Gene Expression Analysis Demonstrates the Existence of Transient Pluripotent Cell Sub-populations *In Vivo*

Detailed gene expression analysis of *Rex1*, *L17*, *Psc1*, *K7* and *Fgf5* in this study has provided the first molecular evidence for transient pluripotent cell sub-populations in the peri-implantation mouse embryo, which differ subtly at the molecular level from both preceding and following pluripotent cell sub-populations.

Rex1 and *L17* were expressed in the pluripotent cells of the early blastocyst and the fully expanded blastocyst until 4.5 d.p.c.. Expression began to downregulate during the late blastocyst stage at approximately 4.75 d.p.c.. Thus, *L17* and *Rex1* defined a pluripotent cell sub-population within the *Oct4* expressing pool, from the ICM population at 3.5 d.p.c. to the epiblast cells of the late stage blastocyst. Expression of these two genes in the developing ICM from 3.5-4.5 d.p.c., but not the trophectoderm or primitive endoderm, may indicate a role for these genes in specifying the ICM/early epiblast lineage or maintaining the undifferentiated state of these cells. *L17* and *Rex1* also appeared to be regulated in a similar manner during ES cell differentiation to

EPL cells, raising the possibility that expression of these two genes are coordinately regulated *in vitro* and *in vivo*. *Rex1* is a zinc finger-containing transcription factor which is regulated by Oct4 and another protein complex termed Rox1 (Hosler *et al.*, 1993, Ben-Shushan *et al.*, 1998). However, the role of *Rex1* in pluripotent cell development during early embryogenesis is not known. Full length *L17* cDNAs have been isolated and homology searches indicated that *L17* shared 79% similarity with the transcription factor, CP-2 (Rodda, 1998; Scherer, 1999). However, *L17* promoter sequences have not yet been obtained and therefore potential regulation of *L17* by Oct4 is yet to be analysed.

Expression of *Psc1 in vivo* was overlapping, yet distinct from that of *L17* and *Rex1*. *Psc1* was expressed in the ICM of the early blastocyst, epiblast cells of the fully expanded blastocyst, and began to downregulate expression in peri-implantation embryos at day 5.0 of embryogenesis. *Psc1* expression persisted in pluripotent cells beyond that observed for *Rex1* and *L17*, both *in vivo* and *in vitro*, enabling identification of a second distinct pluripotent cell sub-population at 4.75-5.0 d.p.c., which was *Rex1*$^-$/*L17*$^-$/*Oct4*$^+$/*Psc1*$^+$. Expression of *Psc1* in the developing blastocyst and epiblast bud population, but not the primitive endoderm and derivatives may suggest a role for this gene in maintenance of the pluripotent cells during this time. The *Psc1* cDNA has been isolated (Schulz, 1996), and encoded a protein of 1005 amino acids which contained an RS domain, a series of serine-arginine repeats. This protein motif is found in splicing factors and proteins which are implicated in the regulation of pre-mRNA splicing. *Psc1* was shown to co-localise with one of these splicing factors, SC-35 (Kavanagh, 1998), suggesting a role for *Psc1* in controlling pre-mRNA splicing during development and in the adult. Mutation of a putative RNA binding domain has also implicated *Psc1* in nucleocytoplasmic trafficking (Russell, 1999).

In contrast to *Rex1*, *L17* and *Psc1*, *K7* expression was not detected in the epiblast of the fully expanded blastocyst at 4.5 d.p.c.. The *K7* gene was first detectable at low levels in the epiblast of 4.75 d.p.c. embryos and was expressed at higher levels in the epiblast bud at 5.0 d.p.c.. *K7* expression persisted within pluripotent cells following the initiation of proamniotic cavitation and was downregulated following formation of a pseudostratified epithelial sheet of primitive ectoderm around the proamniotic cavity at 5.5-5.75 d.p.c.. *K7* expression over a 0.5 d.p.c. period refined the identification of pluripotent sub-populations within the *Oct4*$^+$ pool. The

population present from 4.75-5.0 d.p.c. which had downregulated *Rex1* and *L17* but expressed *Oct4* and *Psc1*, also expressed *K7*. Furthermore, downregulation of *Psc1* and maintenance of *K7* expression at 5.25 d.p.c. defined a third pluripotent cell sub-population present following proamniotic cavitation.

Analysis of *Fgf5* expression was in accordance with that previously documented by Haub and Goldfarb (1991) with expression not detected in pluripotent cells at 3.5 d.p.c., but observable following proamniotic cavitation at 5.25 d.p.c.. However, the expression pattern for this gene was refined by the demonstration that *Fgf5* was not expressed in mature blastocysts in which TGC differentiation had occurred (4.75 d.p.c.). Onset of *Fgf5* expression was not mapped between 5.0 and 5.25 d.p.c., but extrapolation from the *Fgf5 in vitro* expression pattern which sees substantially increased expression in EPL cells grown for 4 days in culture in the absence of LIF indicates that *Fgf5* is likely to upregulate expression *in vivo* around day 5.0 (Figure 3.8). *Fgf5* expression at 5.25 d.p.c. further refined the expression characteristics of the third transient pluripotent cell sub-population to *Rex1*, *L17* and *Psc1* negative and *Oct4*, *K7* and *Fgf5* positive.

Fgf5 expression persists in the primitive ectoderm at 6.5 d.p.c. (data not shown; Haub and Goldfarb, 1991; Hébert *et al.*, 1991) and during gastrulation *Fgf5* expression can be detected within the primitive ectoderm, primitive streak, paraxial mesoderm and definitive endoderm, but was downregulated in the embryo once the formation of the three primary germ layers was complete (Hébert *et al.*, 1991). *Fgf5* expression from 5.5-6.5 d.p.c. in the primitive ectoderm therefore identified a fourth distinct pluripotent cell sub-population within the *Oct4*$^{+}$ pool during the ICM to primitive ectoderm transition, and was characterised by a lack of *Rex1*, *L17*, *Psc1* and *K7* expression and expression of *Oct4* and *Fgf5*. The expression pattern of *Fgf5* indicates that this gene may specify the primitive ectoderm lineage, following cavitation and polarisation, or function during gastrulation to either stimulate mobility of cells or impart competency to respond to the inductive signals of gastrulation (Hébert *et al.*, 1991). However, null mutation of this gene does not result in primitive ectoderm or gastrulation defects (Hébert *et al.*, 1994). Redundancy in FGF signaling within the primitive ectoderm may compensate for the loss of *Fgf5* given that *Fgf4* is also expressed specifically within the primitive ectoderm prior to gastrulation and within the primitive steak during gastrulation (Niswander and Martin, 1992).

Taken together, these data provided evidence for the existence of at least four pluripotent cell sub-populations which can be discriminated on the basis of gene expression (Figure 3.8). The blastocyst sub-population between 3.5-4.5 d.p.c. (*Oct4/Rex1/L17/Psc1* positive, *K7/Fgf5* negative); a second sub-population at 4.75-5.0 d.p.c. (*Rex1/L17/Fgf5* negative, *Oct4/Psc1/K7* positive); a third sub-population at approximately 5.25 d.p.c. (*Rex1/L17/Psc1* negative, *Oct4/K7/Fgf5* positive); and a fourth sub-population from approximately 5.5-5.75 d.p.c. (*Rex1/L17/Psc1/K7* negative, *Oct4/Fgf5* positive). Gene expression did not provide any evidence for spatial heterogeneity within the transient pluripotent cell sub-populations at any of these stages.

3.4.2 Pluripotent Cell Sub-populations Correlate with Embryological Events

The pluripotent cell sub-populations identified by changes in marker gene expression correlated with embryological events of peri-implantation mouse development. Regulation of three genes, *Rex1*, *L17* and *K7*, at 4.75 d.p.c. suggested a shift in the developmental program occurred at this time of mouse embryogenesis. This key developmental event may involve switching from a cleavage based developmental system, which sees no overall increase in embryo size, to a proliferation based system, which couples cell division with embryo growth. The cleavage rate of mouse embryos is initially long, around 20 hours, but progressively shortens to around 10 hours by the blastocyst stage (3.5-4.0 d.p.c.) (Snow, 1976). However, from implantation until the late primitive streak stage at 7.5 d.p.c. of gastrulation, the embryo must undergo more than a 500-fold increase in tissue volume. This is not brought about by increases in cell size, but is achieved by extremely rapid epiblast cell proliferation, with cell cycle times of 5-6 hours required to account for such growth (Snow, 1976). Approximately 25-30 epiblast cells present at 4.5 d.p.c. (Snow, 1976) increases to 120 cells by 5.5 d.p.c. (Snow, 1976; Snow, 1977), indicating that a cell cycle time of 6 hours is required to account for the increase in cell number, and suggests increased proliferation in the pluripotent cells from 4.5 d.p.c.. It is interesting to note that a number of null mutants, including *Fgf4, Fgfr2, Brca1, β1 integrin*, and *EGFR*, begin to degenerate just following implantation at 4.5 d.p.c. as a result of impaired pluripotent cell proliferation (1.4.3.3) (Fässler and Meyer, 1995; Feldman *et al.*, 1995; Stephens *et al.*, 1995; Threadgill *et al.*, 1995; Hakem *et al.*, 1996; Liu *et al.*, 1996; Ludwig *et al.*, 1997;

Arman *et al.*, 1998). Therefore, *Rex1*, *L17* and *K7* may potentially be regulated in a coordinate manner by a pathway involved in controlling pluripotent cell proliferation/maintenance, and may provide reporter genes for the molecular description of this pathway.

Psc1 expression may also be responsive to embryological signals as the downregulation of this gene correlated with the onset of proamniotic cavitation in a manner expected for a gene controlled by the "death" signal postulated by Coucouvanis and Martin (1995). Furthermore, downregulation of *K7* expression at 5.25 d.p.c. is coincident with the onset of primitive ectoderm polarisation and organisation into a pseudostratified epithelial sheet. This may suggest that signals responsible for primitive ectoderm polarisation and/or reorganisation may influence *K7* expression. However, as these genes were identified *in vitro*, the changes in gene expression observed may reflect acquisition of competence to respond to a proliferation signal, such as FGF4, the proposed death signal, or reorganisational cues, rather than production of these embryonic signals in culture.

3.4.3 Expression of Marker Genes in Non-Pluripotent Cells

Rex1, *Fgf5* and *K7* were expressed within extra-embryonic lineages in addition to their pluripotent cell expression. *Rex1* was expressed by the extra-embryonic ectoderm following proamniotic cavitation at day 5.25. In contrast to a previous report (Rogers *et al.*, 1991), *Rex1* could not detected at low levels in the trophectoderm of blastocyst stage embryos. This divergence from the published expression pattern may be attributed to a decreased sensitivity of the whole mount *in situ* hybridisation protocol compared with a radiolabelled sectioned *in situ* hybridisation procedure, as both *in situ* hybridisation methods detect *Rex1* expression in the extra-embryonic ectoderm, a trophectoderm derivative. The importance of *Rex1* expression in the extra-embryonic ectoderm was not investigated further. *Fgf5* appeared to be expressed by primitive endoderm at day 4.75, but was not detected in visceral endoderm, a primitive endoderm derivative, at later stages of embryogenesis from 5.25 d.p.c.. Interestingly, a minority of fully expanded blastocysts display positive staining for FGFR2 protein by immunohistochemistry in morphologically distinct cells which line the blastocoelic cavity, and are assumed to be primitive endoderm cells (Haffner-Krausz *et al.*, 1999). This suggests that *Fgf5* signals may play a role in primitive endoderm differentiation. *K7* expression was detected in the extra-embryonic ectoderm

at 5.0 d.p.c. and often in the ectoplacental cone at 5.25 d.p.c. However, these regions were often damaged or removed from post-implantation embryos during dissection and as a result the importance of this extra-embryonic expression was not investigated further.

3.4.4 Gene Expression Analysis Demonstrates the Existence of Transient Pluripotent Cell Sub-populations *In Vitro* During the ES to EPL Cell Transition

More detailed examination of *Rex1*, *L17*, *Psc1* and *Fgf5* expression during the differentiation of ES cells to EPL cells provided insight into the nature of pluripotent cell maintenance and gene expression *in vitro*. All markers examined were regulated in a precise manner, being restricted to sub-populations of *Oct4* expressing cells. The expression of *L17* was indistinguishable from that of the previously documented ES cell marker, *Rex1* (Rathjen *et al.*, 1999), with high level expression detected in ES cells. Transcription of both genes continued, although at decreasing levels in EPL cells grown for 2-4 days in the presence of LIF and 2 days in the absence of LIF. The loss of *L17* and *Rex1* transcripts from later passage EPL cells (grown for 4 days in the absence of LIF or 6 days in the presence or absence of LIF) indicated the progression of EPL cells to a "later" EPL cell state. The resemblance between the expression profiles of *Rex1* and *L17 in vitro* provided further evidence for coordinate regulation of these two transcription factors.

In a manner similar to that observed *in vivo*, *Psc1* expression *in vitro* demonstrated an overlapping, yet distinct expression pattern for *L17* and *Rex1*, with expression downregulating slower than that observed for *Rex1* and *L17*. *Fgf5* transcripts were not detected in ES cells, but became progressively upregulated following culture of EPL cell, with highest level expression seen following 4 and 6 days culture in the absence of LIF. *K7* expression began to upregulate in EPL cells grown for 4 days in the presence LIF and reached maximal levels in EPL cells grown for 6 days in the presence of LIF.

Expression analysis of these markers therefore demonstrated the existence of transient pluripotent cell sub-populations within the *in vitro* culture system which correlated with those observed *in vivo* (Figure 3.8). High *Oct4*, *Rex1*, *L17* and *Psc1* expression, and low or undetectable *K7* and *Fgf5* expression demarcated the ES cell sub-population *in vitro,* which displayed similar expression characteristics to the 3.5-4.5 d.p.c. ICM/early epiblast population *in*

vivo. Morphological transition from three dimensional ES cell colonies to monolayer cultures containing specific nuclear characteristics (Rathjen *et al.*, 1999) distinguished EPL cells cultured in MEDII for 2 days in the presence of LIF. This population of EPL cells demonstrated lower *Rex1*, *L17* and *Psc1* expression levels, with *K7* and *Fgf5* expression comparable to that of ES cells. EPL cells grown for 2 days in the absence of LIF or 4 days in the presence of LIF continued to express *Rex1*, *L17* and *Psc1*, although at lower levels, and began to upregulate *K7* expression. This expression profile was consistent with gene expression *in vivo* at day 4.75. EPL cells grown for 6 days in the presence of LIF and 4 days in the absence of LIF displayed high *K7* and *Fgf5* expression, low *Psc1* expression, and almost undetectable *Rex1* and *L17* expression, similar to the marker gene expression profile of the 5.0 d.p.c. epiblast. *Fgf5* expression by EPL cells grown for 4 days in the absence of LIF or 6 days in the presence of LIF suggested that *Fgf5* expression may also be detected in the 5.0 d.p.c. embryo. EPL cells grown for 6 days in the absence of LIF, in the case of *Fgf5*, or 8 days in the presence of LIF, in the case of *K7*, still expressed these transcripts at maximal levels, consistent with expression of these genes by the 5.25 d.p.c. embryo.

In the embryo, K7 expression downregulates at 5.5-5.75 d.p.c., however, EPL cells were not grown for a longer period in culture to determine if the transcript also undergoes downregulation with further passage of EPL cells. Alternatively, differences between the ES to EPL cell model and *in vivo* mouse development may exist such that K7 expression may not downregulate in EPL cells. EPL cells grow as monolayers, without the three dimensional architecture of the embryo. Crucial cell-cell contacts or signals may be lacking within the cell culture system such that the organisational cue which normally results in the formation of a pseudostratified epithelial sheet around the proamniotic cavities is not expressed in monolayer culture. Therefore, EPL cells in monolayer culture may only progress to the early primitive ectoderm stage equivalent to approximately 5.25 d.p.c. *in vivo*, resulting in maintenance of *K7* expression.

The rate at which the changes in marker gene expression were observed was delayed when EPL cells were grown in the presence of LIF. This was consistent with previous findings which report inhibited developmental progression of EPL cells in culture when grown in exogenous LIF (Rathjen *et al.*, 1999). Several other lines of data have also implicated LIF in inhibiting the

progression of pluripotent cell development both *in vitro* and *in vivo*. Primitive ectoderm formation during embryoid body differentiation of ES cells is inhibited during culture in the presence of LIF (Shen and Leder, 1992). Furthermore, abnormal proliferation and delayed gastrulation is observed in transgenic mice overexpressing the extracellular matrix localised form of LIF (Conquet *et al.*, 1992).

3.4.5 ES and EPL Cell Embryonic Equivalents

In vitro expression profiles of marker genes was consistent with the expression patterns observed for these genes *in vivo*. This validated the use of the ES to EPL cell transition *in vitro* as a model of the ICM to primitive ectoderm transition *in vivo,* and enabled alignment of embryonic populations with ES cells and EPL cells (Figure 3.8).

Molecular analysis of multiple genes expressed by ES cells and their cognate expression patterns *in vivo* during this study suggested that ES cells could represent either ICM cells at 3.5 d.p.c. or epiblast cells prior to 4.75 d.p.c., that express *Rex1*, *L17* and *Psc1* at high levels. ES cells were originally isolated from pluripotent cells of the pre-implantation stage blastocyst and following reintroduction into host blastocysts were shown to contribute to embryonic and extraembryonic lineages (Beddington and Robertson, 1989), suggesting that ES cells were similar to the ICM of pre-implantation blastocysts. Early passage ES cells isolated from pre-implantation blastocysts were also capable of forming both trophoblast giant cells and extra-embryonic endoderm cell types *in vitro* (Evans and Kaufman, 1983; Doetschman *et al.*, 1985). However, isolation of ES cell lines from epiblast cells of 4.5 d.p.c. blastocysts can be achieved at double the frequency when compared with ES cell lines isolated from 3.5 d.p.c. blastocysts (Brook and Gardner, 1997). Furthermore, these ES cell lines were devoid of trophoblast giant cells, suggesting that the presence of trophoblast giant cells in ES cell lines isolated from pre-implantation blastocysts can be explained by contaminating trophectoderm stem cells with a limited proliferative potential (Brook and Gardner, 1997). Contaminating trophectoderm stem cells within ES cell lines may also account for the low frequency of contribution to trophoblast giant cells following blastocyst injection of multiple cells from ES cell cultures, but not from the injection of single ES cells into host blastocysts (Beddington and Robertson, 1989). Therefore,

taken together, ES cells most likely represent epiblast cells prior to 4.75 d.p.c., and not ICM cells.

EPL cells cultured for 4 days in the presence of LIF or 2 days in the absence of LIF are likely to represent pluripotent cell subtypes equivalent to 4.75 d.p.c. epiblast tissue, as these respective *in vitro* and *in vivo* pluripotent cell populations both begin to downregulate expression of *Rex1* and *L17* and begin to upregulate *K7* expression. EPL cells cultured for 4 days in the absence of LIF or 6 days in the presence of LIF are most closely related to pluripotent cells of approximately 5.0 d.p.c. as both populations still express *Psc1*, although at lower levels, and have high *K7* expression. EPL cells grown for 6 days in the absence of LIF or 8 days in the presence of LIF are most closely related to primitive ectoderm cells at 5.25 d.p.c. as both populations have downregulated *Psc1*, but express *K7* and *Fgf5*.

3.4.6 Markers for the Identification of Other Pluripotent Cell Populations

The availability of these marker genes which have facilitated the detection of ES and EPL cell embryonic equivalents may improve ES cell isolation methods from mice other than the 129/sv strain or from other commercially important species. The ability to isolate germ line competent pluripotent cells from mice appears to be strain dependent as variations in the timing of developmental progression of pluripotent cells and/or specific genetic components within a particular strain appear critical for ES cell isolation. Similar expression mapping of marker genes in other strains of mice could allow *Rex1/L17/Psc1* positive pluripotent cells and *Fgf5* negative to be determined and targeted for isolation and culture *in vitro*. The identification of a marker which demarcates the ICM lineage from the early epiblast lineage facilitate the isolation of ES cell-like cells from other species, as in the 129 mouse a 2 fold higher isolation efficiency is achieved when this population is isolated (Brook and Gardner, 1997). Methods may require the isolation of later, peri-implantation embryos instead of pre-implantation blastocysts if pluripotent cell progression is slower in other strains. Moreover, the characterisation of related pluripotent cell sub-populations in other mammals via expression analysis using species specific *Rex1*, *L17* or *Psc1* homologues, may also improve the chances of isolating pluripotent cells from these species.

Mouse ES, EG and EC cells and human ES, EG and EC cells undergo indefinite self-renewal in culture, however, these cells have diverse phenotypes and developmental potentials.

Mouse EG cells are derived from cultured primordial germ cells and convert to an ES cell phenotype in culture (Matsui *et al.*, 1992; Resnick *et al.*, 1992). Like ES cells, EG cells injected into blastocysts can contribute to chimera formation and pass through the germline. However, ES and EG cells differ in their imprinting of the *Igf2r* gene (Labosky *et al.*, 1994). EC cell lines derived from spontaneous gonadal tumours or from tumours produced by ectopic engraftment of ES cells or embryonic cells, are more phenotypically diverse than ES and EG cells. Although capable of contributing to many tissues in chimeric mice, they are generally unable to contribute to the germline, indicating that they have undergone genetic transformations incompatible with gametogenesis (Papaioannou and Rossant, 1983). Some EC cell lines, for example F9 EC cells, give rise to primitive endoderm *in vitro*, suggesting an equivalence with ICM cells. Furthermore, many EC cell lines are capable of differentiation as embryoid bodies, that resembles the differentiation behaviour of ICM cells *in vitro* (Martin, 1975a). However, protein expression profiles of EC cells and early embryonic cells suggest that EC cells may be more similar to the primitive ectoderm than the ICM (Martin *et al.*, 1978). Human EC cell lines derived from germ cell tumours, the most common tumour in young males, are morphologically distinct from mouse EC cell lines (Pera *et al.*, 1987) and more similar to human ES and EG cell lines, recently isolated from human blastocysts (Thomson *et al.*, 1998) and primordial germ cells (Shamblott *et al.*, 1998) respectively. Human ES and EG cell lines however do not appear to show the robust proliferation or clonability of mouse ES clones, and their developmental relationship is unknown. Furthermore, morphology, marker expression and growth requirements of human EG cells also differ in many ways from those of human ES cells.

Although there are apparent differences between *in vitro* pluripotent cell lines of mouse and humans, expression of molecular markers may be common to mammalian pluripotent cells. Therefore, gene expression analysis of markers derived from the ES to EPL cell transition *in vitro* within these phenotypically diverse pluripotent cell populations may provide valuable insight into the inter-relationships between these cell lines and to pluripotent cells within the embryo.

3.4.7 Nomenclature for Pluripotent Cell Sub-populations

Molecular subdivision of pluripotent cell populations of the early mouse embryo allows the establishment of a more precise pluripotent cell nomenclature. At present the term epiblast is used

to encompass diverse pluripotent cell populations which exist from the formation of primitive endoderm until the onset of gastrulation. The term does not delineate the transient, pluripotent cell sub-populations which exist during this time which can be recognised by distinct marker gene expression. Results presented in this thesis suggests that the term ICM should be used to describe pluripotent cells until 4.5 d.p.c. which are *Oct4*, *Rex1*, *L17* and *Psc1* positive. A new term "periblast" could be introduced to describe the pluripotent cell populations of the late stage blastocyst at 4.75 d.p.c. until proamniotic cavitation (*Rex1/L17* negative, *Oct4/Psc1/K7* positive). The term primitive ectoderm could then be used to describe pluripotent cell populations following proamniotic cavitation.

CHAPTER 4:

MOLECULAR ANALYSIS OF THE MARKER GENE, *K7*

4.1 INTRODUCTION

Psc1, *L17* and *K7* were three novel marker genes expressed in overlapping, but distinct pluripotent cell sub-populations within the *Oct4*$^{+}$ pool of early mouse embryogenesis. Their expression profiles were tightly regulated and correlated with important developmental events. *K7* demonstrated a particularly interesting expression pattern *in vivo*, with expression detected in a short window of early mouse development from approximately day 4.75 to day 5.25. This time period was coincident with a number of cellular events, such as the proposed commencement of more rapid pluripotent cell proliferation, proamniotic cavitation and formation of a polarised, pseudostratified epithelium of primitive ectoderm. The regulated expression of *K7* during this period of embryogenesis suggested a potential role for this gene in pluripotent cell progression *in vivo*. This chapter investigated a possible function for the *K7* gene via three approaches; (i) isolation and sequence analysis of a complete *K7* cDNA, (ii) physical mapping of the *K7* locus within the mouse genome, and (iii) analysis of *K7* expression at later stages of mouse development and within adult tissues.

(i) Information regarding possible gene function can be obtained from sequence analysis, via the identification of conserved motifs and relationships to other genes of known function. Little information about a possible function for *K7* was available from sequence data of the existing 232 bp dd-PCR product (see 1.6.3). Therefore, further isolation and characterisation of the *K7* sequence was required, and was carried out using molecular cloning techniques, sequencing and homology searches using the cDNA sequence and/or protein sequence of the encoded gene product.

(ii) Possible biological roles for *K7* during development and/or in the adult were also investigated by determining the position of the *K7* gene within the mouse genome. Identification of developmental mutation(s) which co-localise with the *K7* locus and investigation of the *K7* gene within the mutant strain may allow causative association, and provide further functional information.

(iii) Genes which demonstrate early, developmentally regulated, embryonic expression are often re-expressed later in development and within adult tissues. Investigation of *K7* expression during later embryonic stages and within adult tissues was carried out to determine

whether *K7* function was specific to early peri-implantation development, or if additional, later sites of *K7* function could be detected.

4.2 ISOLATION AND SEQUENCE ANALYSIS OF THE *K7* cDNA

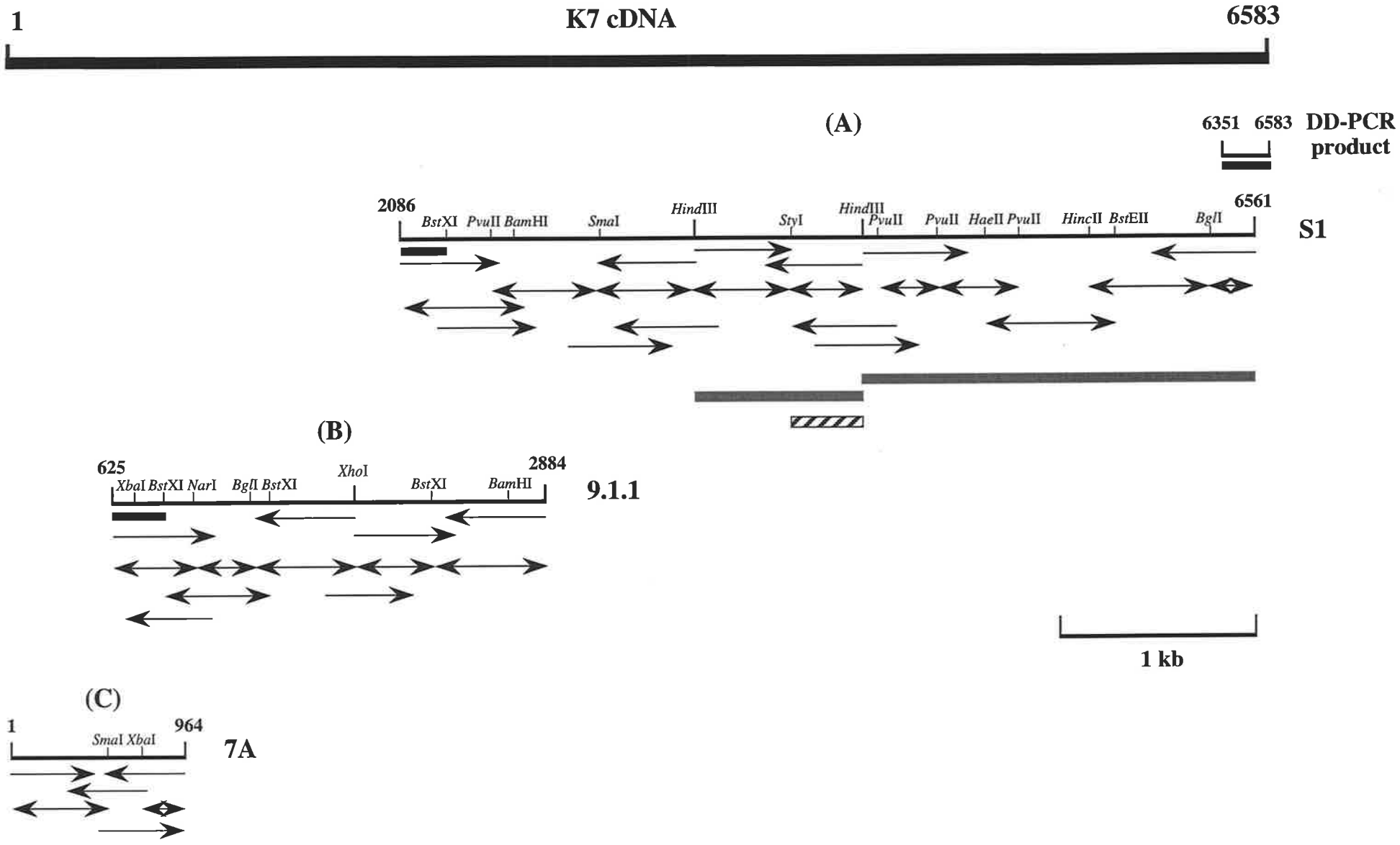
4.2.1 Isolation of *K7* cDNA Clones

K7 transcripts identified by northern analysis were estimated to be approximately 7 kb and 4.5 kb (see Figure 1.2 C) (S. Sharma, unpublished). As they were expressed by D3 embryonic stem (ES) cells, *K7* cDNA clones were isolated from a λ Zap II D3 ES cDNA library (Clontech). This library was constructed by reverse transcription of undifferentiated D3 ES cell RNA, using both poly-d(T) and random primers to provide good coverage of long transcripts. The library was initially screened by Dr. Shiwani Sharma using the 232 bp *K7* dd-PCR product, which contained a polyadenylation element and a poly-$(A)^+$ sequence, indicating that this clone was from the 3' end of the gene. An approximately 4.5 kb *K7* cDNA clone, termed S1, was isolated (Figure 4.1 A). Sequence analysis of the 5' and 3' ends of this clone using USP and RSP sequencing primers (2.2.7.1), demonstrated that the 3' end of S1 contained the consensus polyadenylation element, but did not contain a poly-$(A)^+$ tail. Therefore, S1 contained the majority of the 3' end of the *K7* transcript.

Since the larger *K7* transcript was predicted to be approximately 7 kb in size, a second library screen was carried out in order to obtain the 5' end of the sequence. An *Eco*RI/*Bst*XI fragment, which constituted the most 5' 207 bp of the S1 cDNA clone, was used as a probe to screen 500,000 plaques (Figure 4.1). Five duplicate positive plaques were identified, isolated and purified through two subsequent rounds of library screening (2.3.14). cDNA clones were excised using the λ ZAPII excision procedure into pBluescript SK (2.3.15). DNA was isolated from the bacterial colonies and digested with *Eco*RI to excise the cloned cDNAs. Digest reactions were run on a 1% TAE agarose gel, stained with ethidium bromide, and viewed under UV light to determine approximate sizes of the inserts. Southern analysis indicated that four of the purified plaques isolated from the library screen contained *K7* cDNA clones (data not shown). The cDNA inserts ranged in size from approximately 800 bp to 2 kb. Sequence analysis with USP and RSP primers (2.2.7.1) indicated that the 2 kb *K7* cDNA clone, termed 9.1.1, overlapped with the 5' end of the S1 clone and extended further upstream into the *K7* sequence (Figure 4.1 B).

Figure 4.1: *K7* cDNA cloning and sequencing strategy

Schematic representation of the relative positions and sizes of the three *K7* cDNA clones, termed S1 (**A**), 9.1.1 (**B**) and 7A (**C**), isolated from the λ ZAP II D3 ES cell library. Arrow heads depict the direction of sequencing reactions, with arrow heads at both ends representing sequencing of a particular clone in both directions. Bold lines represent fragments used as probes for library screening, the crosshatched line represents the probe used in RNAse protection assays and the grey lines represent the *K7* fragments used in chromosome localisation studies.


1
K7 cDNA
6583
(A)
6351 6583 DD-PCR product
2086
BstXI PvuII BamHI SmaI HindIII StyI HindIII PvuII PvuII HaeII PvuII HincII BstEII BglI
6561
S1
(B)
625
XbaI BstXI NarI BglI BstXI XhoI BstXI BamHI
2884
9.1.1
1 kb
(C)
1
SmaI XbaI
964
7A

Restriction mapping using six base pair cutting restriction enzymes was carried out and appropriate DNA fragments were subcloned into pBluescript II KS to produce shorter subclones of both S1 and 9.1.1 *K7* cDNA clones (Figure 4.1 A and B). Subcloned fragments were sequenced in both directions using USP and RSP sequencing primers (2.2.7.1), and the resulting sequence was used to design further subcloning strategies in order to sequence S1 and 9.1.1 completely in both directions. Alignment of a continuous sequence indicated that S1 and 9.1.1 were 4475 bp and 2259 bp in length respectively, and together represented 5936 bp of the *K7* sequence. Conceptual translation of the compiled *K7* sequence demonstrated the presence of an extensive open reading frame, extending from the most 5' end of the sequence. Since northern analysis indicated a transcript size of approximately 7 kb, the *K7* cDNA sequence was still incomplete at the 5' end.

A third library screen was carried out using an *Eco*RI/*Bst*XI fragment constituting the most 5' 268 bp of the compiled *K7* sequence as a probe (Figure 4.1). Three duplicate plaques were identified, isolated and purified through a further round of library screening (2.3.14), and "zapped" into pBluescript II SK for DNA isolation and analysis. Sequencing of termini using USP and RSP primers (2.2.7.1) demonstrated that all three clones overlapped with the known *K7* sequence. Clone 7A, of approximately 1 kb, overlapped with the known sequence by 339 bp and extended a further 625 bp upstream (Figure 4.1 C). This clone was chosen for further restriction mapping and subcloning into pBluescript II KS. Following sequencing of the overlapping subclones in both directions, the amassed sequence of this clone was 964 bp, extending the *K7* sequence to 6561 bp in length.

Assembly of a contiguous *K7* sequence from the three library clones indicated a number of sequencing discrepancies, which were resolved following generation of sequencing primers (2.2.7.2) directed against the *K7* sequence. Following this analysis, one base pair discrepancy between clone 7A and 9.1.1 remained at nucleotide position 941 (star/bold; Figure 4.2), indicating that either a transcription error occurred during library construction, or the presence of a polymorphism at this position within the *K7* sequence. RT-PCR was carried out on D3 ES cell RNA (2.3.16/17), using the RT-2 primer as a gene specific primer in the reverse transcriptase reaction, and primers RT-1 and 7A2R in the PCR reaction (2.2.7.2/3). A PCR product of the predicted size was obtained (Figure 4.3 A) and was sequenced directly (2.3.13) using primers

Figure 4.2: The compiled *K7* cDNA sequence and deduced protein sequence

The compiled *K7* sequence was generated from the dd-PCR product and the three overlapping cDNA library clones isolated from the λ ZAP II D3 ES cell library: Nucleotide position 1-703 from *K7* cDNA clone 7A, position 704-2713 from *K7* cDNA clone 9.1.1, nucleotide position 2714-6561 from the S1 *K7* cDNA clone, and position 6562-6583 was obtained from the dd-PCR product sequence (S. Sharma, unpublished). The nucleotide discrepancy at position 941 bp is highlighted by bold type and a star (*). Conceptual translation of the K7 ORF (2118 amino acids) is shown below the nucleotide sequence and spanned nucleotides 71 bp to 6524 bp. An in frame termination codon (TGA; bold/double underline) was located at nucleotide positions 41-43 bp, 20 bp upstream of the translation initiation site (ATG; bold/boxed). The polyadenylation element was located at position 6548-6563 bp (bold/shaded). Two potential bipartite nuclear localisation signals (Robbins *et al.*, 1991) were located one amino acid apart, at amino acid positions 1411-1427 and 1412-1428 (boxed). The region containing similarity to bimB, Cut1 and Esp1 was found between amino acids 1886-2118 (bold/underline). Hydrophobic stretches from amino acid 354-374 and 847-867 (boxed/shaded) were identified using the simple modular architecture research tool (Schultz *et al.*, 1998) at the PROSITE database (2.3.33).

```
   1  gcccggcagcttcgagggaattaaggcacggagggtctggtgagcgcgcgctcaagcc

  59 gcgacttttgccATGAGGAACTTCAAAGGAGTCAACTTCGCCACTCTGCTCTGCAGCAAG
   1              M  R  N  F  K  G  V  N  F  A  T  L  L  C  S  K

 119 GAGGAGACCCAGCAGCTGCTGCCTGACTTGAAGGAGTTCCTGTCCAGATCCCGGACTGAT
  17  E  E  T  Q  Q  L  L  P  D  L  K  E  F  L  S  R  S  R  T  D

 179 TTCCCCAGCAGCCGAACAGATGCCGAAAGAAGACAAATTTGCGATACCATCCTGAGGGCT
  37  F  P  S  S  R  T  D  A  E  R  R  Q  I  C  D  T  I  L  R  A

 239 TGCACCCAGCAGTTGACTGCCAAGCTCGATTGCCCTGGGCACCTGAGGAGTATTTTGGAC
  57  C  T  Q  Q  L  T  A  K  L  D  C  P  G  H  L  R  S  I  L  D

 299 CTGGCAGAGCTGGCCTGTGATGGCTATTTGTTGTCTACTCCACAGCGCCCTCCCCTCTAC
  77  L  A  E  L  A  C  D  G  Y  L  L  S  T  P  Q  R  P  P  L  Y

 359 CTGGAACGCATTCTCTTCATCCTACTTCGCAATGGTTCCACTCAGGGAAGCCCAGACACT
  97  L  E  R  I  L  F  I  L  L  R  N  G  S  T  Q  G  S  P  D  T

 419 GTGCTGCGCCTTGCCCAGCCTCTCCATGCCTGCTTGGTTCAGAACTCTGGAGAGGCTGCC
 117  V  L  R  L  A  Q  P  L  H  A  C  L  V  Q  N  S  G  E  A  A

 479 CCTCAGGACTATGAGGCCGTGACCCGGGGCAGCTTCTCTCTGTTTTGGAAGGGAGCAGAG
 137  P  Q  D  Y  E  A  V  T  R  G  S  F  S  L  F  W  K  G  A  E

 539 GCACTGTTGGAGCGGCGAGCTGCGTTCTCAACTCGGCTGAATGCCTTGAGCTTCTTGGTG
 157  A  L  L  E  R  R  A  A  F  S  T  R  L  N  A  L  S  F  L  V

 599 CTCTTGGAGGATGGCAGTGTCCCCTGTGAGGTCCCTCACTTTGCTTCTCCAACCGCCTGT
 177  L  L  E  D  G  S  V  P  C  E  V  P  H  F  A  S  P  T  A  C

 659 CGTTTAGTTGCCGCCTATCAGCTGTACGATGCTACAGGACAAGGTCTAGATGAAGCTGAT
 197  R  L  V  A  A  Y  Q  L  Y  D  A  T  G  Q  G  L  D  E  A  D

 719 GCCGACTTCCTGTATGAAGTGCTTTCTAGGCACCTGATCAGAGTCTTAGTAGGTGAGGGT
 217  A  D  F  L  Y  E  V  L  S  R  H  L  I  R  V  L  V  G  E  G

 779 GGTAGCTCTCCCGGTCCTCTCTCGCCTCAGAGGGCCCTCTGCCTCTTGGAAATCACCCTG
 237  G  S  S  P  G  P  L  S  P  Q  R  A  L  C  L  L  E  I  T  L

 839 GAACACTGCCGTCGGCTCTGCTGGAACCACCACCACCGCCAGGCTGCCAGAGCAGTGGAG
 257  E  H  C  R  R  L  C  W  N  H  H  H  R  Q  A  A  R  A  V  E

 899 AGGGCTCGTAATCACCTAGAGAAGACCAGTGTAGCTCCCAGCCCTCCAGTTATGCCAGATG
 277  R  A  R  N  H  L  E  K  T  S  V  A  P  S  L  Q  L  C  Q  M

 959 GGGGTTGAGCTGTTGGAAGCTGTAGAAGAAAGACCAGGGGCAGTGGCCCAGCTTCTGAGG
 297  G  V  E  L  L  E  A  V  E  E  R  P  G  A  V  A  Q  L  L  R

1019 AAGGCAGCAGCCGTTCTTATCAACAGTATCAAGGCGCCGTCGCCCCCGCTCCGGGCATTG
 317  K  A  A  A  V  L  I  N  S  I  K  A  P  S  P  P  L  R  A  L

1079 TACGACAGCTGCCAGTTCTTCCTTTCAGGCCTGGAGAGAGGCATCAGGAGGCACTGTGGA
 337  Y  D  S  C  Q  F  F  L  S  G  L  E  R  G  I  R  R  H  C  G

1139 CTTGATGCCATATTGAGTCTCTTTGCTTTTCTTGGGGGATACAGCTCCCTTGTCCGGCAC
 357  L  D  A  I  L  S  L  F  A  F  L  G  G  Y  S  S  L  V  R  H

1199 CTGAGAGAAGTCTCTGAAGCATCTTCCAAGCAGCAGCAGTGCTTGCTTCAGATGCACTTC
 377  L  R  E  V  S  E  A  S  S  K  Q  Q  Q  C  L  L  Q  M  H  F

1259 CAGGGTTTCCACCTCTTCACCGGGATTGTGTATGACTTTGCTCAAGGCTGTCAGGCAACT
 397  Q  G  F  H  L  F  T  G  I  V  Y  D  F  A  Q  G  C  Q  A  T

1319 GAGCTTGCTCAGCTAGTGGACGGTTGCAGATCTGCTGCTGTCTGGATGTTGGAGGCCTTA
 417  E  L  A  Q  L  V  D  G  C  R  S  A  A  V  W  M  L  E  A  L

1379 GAGGGCCTGTCGGGTGGAGAGCTGGCTGACTACCTGAGCATGACTGCCTCTTACACCAGC
 437  E  G  L  S  G  G  E  L  A  D  Y  L  S  M  T  A  S  Y  T  S
```

```
1439 AACTTGGCCTACAGCTTCTTTAGTCAGAAGCTCTATGAGGAGGCCTGCGTCATCTCAGAG
 457  N  L  A  Y  S  F  F  S  Q  K  L  Y  E  E  A  C  V  I  S  E

1499 CCAGTCTGTCAGCACTTGGGCTCGGCAACGTTGGGTGCTTGTCCTGAAGTGCCTCCTGAG
 477  P  V  C  Q  H  L  G  S  A  T  L  G  A  C  P  E  V  P  P  E

1559 AAGCTACATAGGTGCTTCAGGCTGCATGTGGAGAGTTTGAAGAAACTGGGTAAACAGGCC
 497  K  L  H  R  C  F  R  L  H  V  E  S  L  K  K  L  G  K  Q  A

1619 CAGGGCTGCAAGATGGTGACCTTATGGCTGGCAGCTCTGAAACCCTACAGTCTGGAACAC
 517  Q  G  C  K  M  V  T  L  W  L  A  A  L  K  P  Y  S  L  E  H

1679 ATGGTCGAGCCAGTCACTTTCTGGGTTCGAGTCAAAATGGATGCATCCAGGGCAGGAGAC
 537  M  V  E  P  V  T  F  W  V  R  V  K  M  D  A  S  R  A  G  D

1739 AAGGAGCTACAACTTCAGACTTTGCGAGACAGCCTAAGCTGCTGGGACCCAGAGACACAG
 557  K  E  L  Q  L  Q  T  L  R  D  S  L  S  C  W  D  P  E  T  Q

1799 TCTCTGCTTCTCAGGGAGGAGCTGCGGGCCTACAAGGCAGTACGGGCTGACACTGGGCAG
 577  S  L  L  L  R  E  E  L  R  A  Y  K  A  V  R  A  D  T  G  Q

1859 GAGCGCTTCAACATCATCTGTGATCTGCTCGAGCTGAGCCCTGAAGAGACAGCAGCTGGA
 597  E  R  F  N  I  I  C  D  L  L  E  L  S  P  E  E  T  A  A  G

1919 GCCTGGGCACGAGCTACCTACCTGGTGGAACTGGCCCAAGTGCTGTGCTACCACAACTTT
 617  A  W  A  R  A  T  Y  L  V  E  L  A  Q  V  L  C  Y  H  N  F

1979 ACCCAGCAGACCAACTGCTCTGCCTTGGATGCTGTCCAGGAAGCCCTGCAGCTTCTGGAG
 637  T  Q  Q  T  N  C  S  A  L  D  A  V  Q  E  A  L  Q  L  L  E

2039 TCCGTGAGTCCTGAGGCCCAGGAGCAGGATCGACTTCTGGATGATAAAGCACAGGCCTTG
 657  S  V  S  P  E  A  Q  E  Q  D  R  L  L  D  D  K  A  Q  A  L

2099 TTATGGCTTTACATCTGTACCTTAGAGGCCAAGATGCAGGAAGGTATTGAGCGGGATCGG
 677  L  W  L  Y  I  C  T  L  E  A  K  M  Q  E  G  I  E  R  D  R

2159 AGAGCCCAGGCCCCTAGTAACTTGGAGGAATTTGAAGTCAATGACCTGAACTATGAAGAT
 697  R  A  Q  A  P  S  N  L  E  E  F  E  V  N  D  L  N  Y  E  D

2219 AAACTGCAGGAAGATCGTTTTCTGTATAGTAGCATTGCCTTTAATCTGGCTGCTGATGCA
 717  K  L  Q  E  D  R  F  L  Y  S  S  I  A  F  N  L  A  A  D  A

2279 GCTCAGTCCAAATGCCTGGACCAGGCCCTGACTCTGTGGAAAGAAGTGCTCACCAAGGGG
 737  A  Q  S  K  C  L  D  Q  A  L  T  L  W  K  E  V  L  T  K  G

2339 CGAGCCCCAGCTGTGCGGTGTCTCCAGCAGACTGCAGCCTCCCTGCAAATCCTAGCGGCC
 757  R  A  P  A  V  R  C  L  Q  Q  T  A  A  S  L  Q  I  L  A  A

2399 GTCTATCAGTTGGTGGCAAAGCCCTTGCAGGCTCTGGAGACCCTCCTGCTCCTGCAGATT
 777  V  Y  Q  L  V  A  K  P  L  Q  A  L  E  T  L  L  L  L  Q  I

2459 GTATCTAAGAGACTGCAGGACCACGCAAAGGCAGCCAGCTCTTCCTGCCAGCTTACCCAG
 797  V  S  K  R  L  Q  D  H  A  K  A  A  S  S  S  C  Q  L  T  Q

2519 CTACTCTTGAACCTTGGCTGTCCCAGCTATGCGCAGTTGTATCTGGAAGAGGCGGAGTCA
 817  L  L  L  N  L  G  C  P  S  Y  A  Q  L  Y  L  E  E  A  E  S

2579 AGCCTGAGGAGCCTGGACCAGACAAGTGACGCGTGCCAGCTGCTGTCCCTGACCTGTGCT
 837  S  L  R  S  L  D  Q  T  S  D  A  C  Q  L  L  S  L  T  C  A

2639 CTGCTTGGAAGTCAGCTCTGCTGGGCTTGCCAGAAGGTAACTGCAGGCGTTTCTTTGTTG
 857  L  L  G  S  Q  L  C  W  A  C  Q  K  V  T  A  G  V  S  L  L

2699 CTTTCTGTGCTTCGGGATCCTGCCCTCCAGAAGTCATCTAAGGCTTGGTACTTGCTACGT
 877  L  S  V  L  R  D  P  A  L  Q  K  S  S  K  A  W  Y  L  L  R

2759 GTCCAGGCCCTGCAAGTTCTGGCTTTTTACCTCAGTCTCTCGTCCAACCTCCTCTCAAGT
 897  V  Q  A  L  Q  V  L  A  F  Y  L  S  L  S  S  N  L  L  S  S
```

```
2819 GCCCTGCGTGAGCAGCTCTGGGATCAAGGCTGGCAGACCCCGGAGACGGCGCTGATTGAT
 917  A  L  R  E  Q  L  W  D  Q  G  W  Q  T  P  E  T  A  L  I  D

2879 GCCCACAAGCTCCTCAGAAGCATCATCATTCTGCTCATGGGAAGTGATGTGCTCTCCATT
 937  A  H  K  L  L  R  S  I  I  I  L  L  M  G  S  D  V  L  S  I

2939 CAGAAAGCAGCCACCGAGTCGCCCTTTCTGGACTACGGTGAGAATCTGGTGCAAAAATGG
 957  Q  K  A  A  T  E  S  P  F  L  D  Y  G  E  N  L  V  Q  K  W

2999 CAGGTTCTCACAGAGGTGCTGACCTGCTCGGAGAGGCTGGTCGGCCGCCTGGGTCGCCTG
 977  Q  V  L  T  E  V  L  T  C  S  E  R  L  V  G  R  L  G  R  L

3059 GGGAACGTGAGTGAGGCCAAGGCCTTTTGCTTGGAGGCCCTAAAACTTACGACTAAGCTG
 997  G  N  V  S  E  A  K  A  F  C  L  E  A  L  K  L  T  T  K  L

3119 CAGATACCTCGACAGTGTGCCCTGTTCCTTGTACTGAAGGGTGAGCTGGAGCTGGCCCGG
1017  Q  I  P  R  Q  C  A  L  F  L  V  L  K  G  E  L  E  L  A  R

3179 GGTGACATTGACCTCTGTCAGTCAGACTTGCAGCAGGTTCTCTTCTTGCTCGAGTCTTCC
1037  G  D  I  D  L  C  Q  S  D  L  Q  Q  V  L  F  L  L  E  S  S

3239 ACAGAGTTTGGTGTTGTGACCCAGCATCCAGACTCAGTGAAGAAGGTGCACACGCAGAAA
1057  T  E  F  G  V  V  T  Q  H  P  D  S  V  K  K  V  H  T  Q  K

3299 GGGAAGCATAAGGCTCAAGGCCCATGTTTCCCACCCCTGTCAGAGGAGGAGCCCTTCCTG
1077  G  K  H  K  A  Q  G  P  C  F  P  P  L  S  E  E  E  P  F  L

3359 AAAGGTCCTGCCCTTGAGCTGGTGGACACTGTGCTCAACGAACCTGGTCCCATCCAGTCC
1097  K  G  P  A  L  E  L  V  D  T  V  L  N  E  P  G  P  I  Q  S

3419 TCTGTAAACTCCTCCCCAGTCCTGAAAACCAAGCCACCACCCAACCCTGGCTTCCTGTCT
1117  S  V  N  S  S  P  V  L  K  T  K  P  P  P  N  P  G  F  L  S

3479 CACTTGCCCAGCTGTGACTGCTTACTGTGTGCCAGCCCTGCCCTCTCAGCAGTCTGTCTG
1137  H  L  P  S  C  D  C  L  L  C  A  S  P  A  L  S  A  V  C  L

3539 CGCTGGGTGCTGGTCACTGCCGGAGTGAGGCTGGCCACCGGTCACAAAGCTCAGGGTCTG
1157  R  W  V  L  V  T  A  G  V  R  L  A  T  G  H  K  A  Q  G  L

3599 GATCTGCTGCAAGCTGTGCTGACGCGTTGCCCTGCAGCCACCAAGCGCTTCACACAGAGT
1177  D  L  L  Q  A  V  L  T  R  C  P  A  A  T  K  R  F  T  Q  S

3659 CTCCAAGCTTCTCTGAATCACAGAACAACCCCCTCTTGTGTTCCCAGCCTCTTTGATGAG
1197  L  Q  A  S  L  N  H  R  T  T  P  S  C  V  P  S  L  F  D  E

3719 ATCATGGCTCAAGTGTACACGCACTTGGCATTGGAGTTCCTGAATCAGACATCTGAAAAG
1217  I  M  A  Q  V  Y  T  H  L  A  L  E  F  L  N  Q  T  S  E  K

3779 AGCCTGGGGAAGGTCCTGGCGTCGGGGCTGAAGTTTGTGGCAACACGGATACAGAGCTTG
1237  S  L  G  K  V  L  A  S  G  L  K  F  V  A  T  R  I  Q  S  L

3839 GAGATCTGGAGAGCCCACCTGCTCTTGGTTCAGGCCCTTGCAAAGCTGGCCCACTTCAGC
1257  E  I  W  R  A  H  L  L  L  V  Q  A  L  A  K  L  A  H  F  S

3899 TGCTGTACTTCCGAACTCTTTGCAAGCTCCTGGGGCTGGCACCCACCATTAGTAAAAAGT
1277  C  C  T  S  E  L  F  A  S  S  W  G  W  H  P  P  L  V  K  S

3959 CTGCCAGTCTTAGAGCCTGCTAAAATTCGGAGACAAAAATGCTCTGGACGGGGCCGCCGG
1297  L  P  V  L  E  P  A  K  I  R  R  Q  K  C  S  G  R  G  R  R

4019 CGGATAGCCTCTGTTCCCCCACCCCTCCATAACAGCTCTCAGAAAGGTCTGGAAGAGGAA
1317  R  I  A  S  V  P  P  P  L  H  N  S  S  Q  K  G  L  E  E  E

4079 GGGCCACCTTGTACCCCTAAGCCTCCAGGTCGGGCTAGGCAAGCTGGCCCTCGTGTCCCT
1337  G  P  P  C  T  P  K  P  P  G  R  A  R  Q  A  G  P  R  V  P

4139 TTCACCATATTTGAAGAAGTCCACCCCACCAAGAGCAAGCTCCAGGTTCCCTTGGCTCCC
1357  F  T  I  F  E  E  V  H  P  T  K  S  K  L  Q  V  P  L  A  P
```

```
4199 AGGGTCCACAGGCGGGCCCAGACTCGCCTCAAGGTGATCTTCAGTGATGACAGTGACCTG
1377  R  V  H  R  R  A  Q  T  R  L  K  V  I  F  S  D  D  S  D  L

4259 GAAGACCTTGTCTCAGCTGACACACAACTTGTAGAGGAGCCTAAGAGGCGAGGGACTGCT
1397  E  D  L  V  S  A  D  T  Q  L  V  E  E  P  K  R  R  G  T  A

4319 TCCCGTACCCGGGGCCAAACTAGGAAGGGACGTAGCTTAAAGACAGATGCGGTGGTTGCT
1417  S  R  T  R  G  Q  T  R  K  G  R  S  L  K  T  D  A  V  V  A

4379 ATAGAAAGTACCCCTGGACATTCCAGTGTCAGTGGGAGGACTCGGAGGGCCAGGAAGGTG
1437  I  E  S  T  P  G  H  S  S  V  S  G  R  T  R  R  A  R  K  V

4439 GCTTCAAGAAATTGTGAAGAAGAGAGTCCCAAAGCACCCCTCTGTGTCTGGGCCAGCCAG
1457  A  S  R  N  C  E  E  E  S  P  K  A  P  L  C  V  W  A  S  Q

4499 GGACCCGAGATCATGAGAAGTATCCCTGAAGAGGAGCCAGTGGACAACCATCTGGAAAAA
1479  G  P  E  I  M  R  S  I  P  E  E  E  P  V  D  N  H  L  E  K

4559 AGCTTCGAGATTCTCAGGGGTTCTGATGGGGAAGACTCGGCCTCAGGTGAGAAGGCGGCA
1499  S  F  E  I  L  R  G  S  D  G  E  D  S  A  S  G  E  K  A  A

4619 GCTGCAGACACTGGTCTACCCGTAGGAGAGTGTGAGGTGTTGAGACGGGACTCCAGCAAG
1517  A  A  D  T  G  L  P  V  G  E  C  E  V  L  R  R  D  S  S  K

4679 GCAGAGAGACCTGTCTTGTACTCAGACACGGAGGCCAATAGTGACCCTAGTCCTTGGCTC
1537  A  E  R  P  V  L  Y  S  D  T  E  A  N  S  D  P  S  P  W  L

4739 CCACCCTTCTCAGTTCCTGCACCCATTGATCTTTCTACCCTGGATTCCATTTCTGACTCA
1557  P  P  F  S  V  P  A  P  I  D  L  S  T  L  D  S  I  S  D  S

4799 TTGAGCATTGCTTTCCGTGGTGTTAGTCACTGCCCTCCAAGTGGGCTCTATGCTCACCTC
1577  L  S  I  A  F  R  G  V  S  H  C  P  P  S  G  L  Y  A  H  L

4859 TGCCGCTTCTTGGCCCTGTGTCTGGGCCACCGAGATCCTTATGCCACTGCGTTCCTTGTT
1597  C  R  F  L  A  L  C  L  G  H  R  D  P  Y  A  T  A  F  L  V

4919 GCTGAGTCTATCTCCATCACCTGCCGTCACCAGCTGCTCACCCACTTGCACCGACAGCTC
1617  A  E  S  I  S  I  T  C  R  H  Q  L  L  T  H  L  H  R  Q  L

4979 AGCAAGGCCCAGAAGCAGCAAGAATCACCTGAACTGGCAGAGCATCTGCAGAGGCTGGAC
1637  S  K  A  Q  K  Q  Q  E  S  P  E  L  A  E  H  L  Q  R  L  D

5039 CTGAAGGAGAGGCCCAGAGGTGTCCCACTGGCCCGTATCCAGCGCCTCTTTTCCTTTAAG
1657  L  K  E  R  P  R  G  V  P  L  A  R  I  Q  R  L  F  S  F  K

5099 GCTTTGGGATCTGGCTGCTTCCCCCAGGCAGAGAAGGAGAGTTTCCAGGAGCGCCTGGCT
1677  A  L  G  S  G  C  F  P  Q  A  E  K  E  S  F  Q  E  R  L  A

5159 CTGATCCCTAGTGGGGTGACTGTCTGTGTGTTGGCTCTGGCCACCCTCCAGCCTGGAACC
1697  L  I  P  S  G  V  T  V  C  V  L  A  L  A  T  L  Q  P  G  T

5219 TTGAGCAACACTCTTTTACTGACCCGCCTGGAAAAGGACAACCCCCCAATTACTGTGAAG
1717  L  S  N  T  L  L  L  T  R  L  E  K  D  N  P  P  I  T  V  K

5279 ATCCCCACTGCCCAAAATAAGCTCCCTCTAAGCGCAGTGCTGAAAGAGTTTGATGCCATC
1737  I  P  T  A  Q  N  K  L  P  L  S  A  V  L  K  E  F  D  A  I

5339 CAGAAGGACCAAAAAGAAAACAGCAGCTGTACTGAGAAACGAGTATGGTGGACGGGCCGG
1757  Q  K  D  Q  K  E  N  S  S  C  T  E  K  R  V  W  W  T  G  R

5399 CTGGCACTGGACCAAAGGATGGAGGCACTCATCACCGCCCTGGAGGAACAGGTGCTGGGC
1777  L  A  L  D  Q  R  M  E  A  L  I  T  A  L  E  E  Q  V  L  G

5459 TGCTGGAGGGGGCTGCTGCTGCCATGCAGTGCTGACCCCAGCCTTGCTCAGGAGGCCTCC
1797  C  W  R  G  L  L  L  P  C  S  A  D  P  S  L  A  Q  E  A  S

5519 AAACTACAAGAGCTGCTGCGGGAATGTGGCTGGGAATATCCTGACTCCACTCTCCTGAAA
1817  K  L  Q  E  L  L  R  E  C  G  W  E  Y  P  D  S  T  L  L  K
```

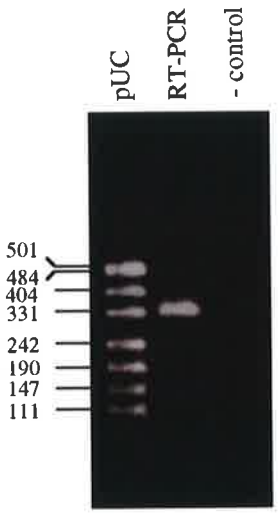
```
5579 GTCATCCTTAGTGGGGCCAGAATCCTCACCAGCCAGGATGTTCAGGCCTTGGCGTGTGGG
1837  V  I  L  S  G  A  R  I  L  T  S  Q  D  V  Q  A  L  A  C  G

5639 CTGTGCCCAGCCCAGCCTGACCGAGCCCAAGTGCTCCTGAGCGAGGCAGTAGGACGGGTT
1857  L  C  P  A  Q  P  D  R  A  Q  V  L  L  S  E  A  V  G  R  V

5699 CAGAGCCAGGAGGCCCCTCGAAGTCAACATCTTGTATTGGTGTTAGACAAGGACCTGCAA
1877  Q  S  Q  E  A  P  R  S  Q  H  L  V  L  V  L  D  K  D  L  Q

5759 AAGCTGCCTTGGGAAAGCACGCCCATCCTCCGAGCACAGCCGGTCACCCGGCTGCCCTCC
1897  K  L  P  W  E  S  T  P  I  L  R  A  Q  P  V  T  R  L  P  S

5819 TTCCGCTTCCTGCTCAGCTACACAGTCACTAAAGAGGCTGGAGCCTCATCAGTGCTGAGC
1917  F  R  F  L  L  S  Y  T  V  T  K  E  A  G  A  S  S  V  L  S

5879 CAAGGTGTTGATCCACAGAATACCTTCTATGTGCTAAACCCTCACAGTAACCTGTCGAGC
1937  Q  G  V  D  P  Q  N  T  F  Y  V  L  N  P  H  S  N  L  S  S

5939 ACAGAGGAGAGATTTCGAGCCAGTTTTAGCAGTGAAACTGGCTGGAAAGGAGTGATTGGG
1957  T  E  E  R  F  R  A  S  F  S  S  E  T  G  W  K  G  V  I  G

5999 GAAGTGCCAAGCCTCGATCAGGTGCAGGCTGCCCTGACAGAGCGGGATTTATATATCTAC
1977  E  V  P  S  L  D  Q  V  Q  A  A  L  T  E  R  D  L  Y  I  Y

6059 GCAGGGCACGGAGCCGGTGCCCGCTTCCTTGATGGGCAGGCTGTCCTGCGGCTGAGCTGC
1997  A  G  H  G  A  G  A  R  F  L  D  G  Q  A  V  L  R  L  S  C

6119 CGGGCGGTGGCCCTGCTGTTCGGCTGCAGCAGTGCAGCCCTTGCCGTGCATGGAAACCTG
2017  R  A  V  A  L  L  F  G  C  S  S  A  A  L  A  V  H  G  N  L

6179 GAGGGAGCTGGTATCGTGCTCAAGTACATCATGGCTGGCTGCCCCTTGTTTCTAGGTAAT
2037  E  G  A  G  I  V  L  K  Y  I  M  A  G  C  P  L  F  L  G  N

6239 CTCTGGGATGTGACTGACCGTGACATTGACCGTTACACAGAGGCTTTGCTGCAGGGCTGG
2057  L  W  D  V  T  D  R  D  I  D  R  Y  T  E  A  L  L  Q  G  W

6299 CTGGGAGCAGGCCCAGGGGCTCCCTTTCTCTACTATGCCAGCCAAGCCCGCCAGGCCCCT
2077  L  G  A  G  P  G  A  P  F  L  Y  Y  A  S  Q  A  R  Q  A  P

6359 CGACTCAAGTACCTCATTGGTGCTGCACCTGTAGCCTACGGCCTACCTATCTCTCTGCAG
2097  R  L  K  Y  L  I  G  A  A  P  V  A  Y  G  L  P  I  S  L  Q

6419 ACACCCtgaagctgtcttgtctatggtagaagctgtgtgactacctccaaacttggtttt
2117  T  P  *

6479 gattatttagaatttttaaagtgattttccctggtgttttgtatgaaatattttctttaa
6539 tgtgaccttXXXXXgatattccaggcccttctaaaaaaaaaaaa
```

Figure 4.3: Confirmation of the *K7* cDNA sequence using RT-PCR

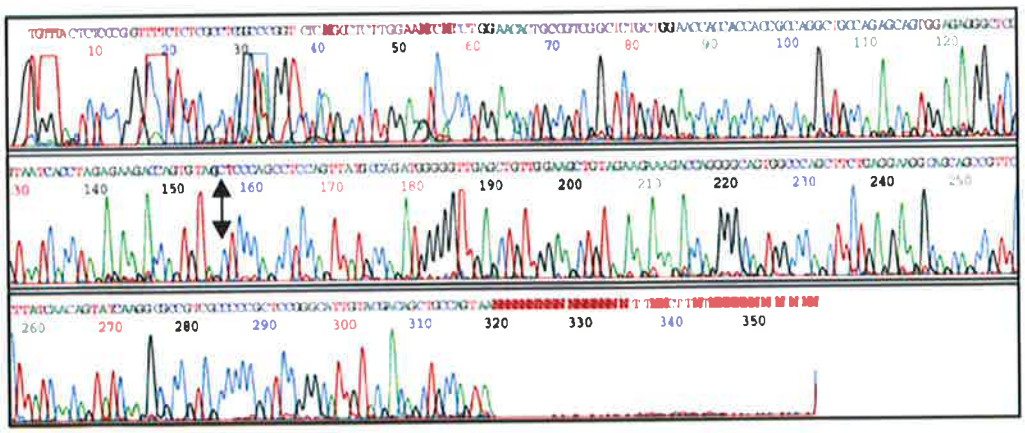
A) Reverse transcription of ES cell RNA (2.3.16) was carried out using the RT-2 primer (2.2.7.3) prior to PCR amplification of the 338 bp region of interest (2.3.17) using primers RT-1 and 7A2R (2.2.7.2/3). An identical PCR reaction was also set up with the exception of MQ H_2O in place of ES cell cDNA to provide a negative control for the RT-PCR experiment. One quarter of each PCR reaction, and 500 ng of *Hpa*II digested pUC19 DNA markers (2.2.11), were electrophoresed on a 2% agarose gel and stained with ethidium bromide to visualise the size of the PCR product. Approximate sizes of the *Hpa*II digested pUC19 markers are indicated.

B) PCR product from **(A)** was gel purified (2.3.7) and the fragment sequenced directly using RT-1 and 7A2R primers (2.3.13). A print out of the sequencing reaction using the RT-1 primer is illustrated. A single "C" peak (arrow) at the appropriate position within the sequence confirmed the identity of the *K7* sequence at this position.

A



B



RT-1 and 7A2R (2.2.7.2/3), following gel purification. The presence of a single "C" peak at the appropriate nucleotide position in the PCR product sequences indicated that a polymorphism at this position was not present (arrow; Figure 4.3 B), and that clone 9.1.1 contained the correct *K7* sequence.

4.2.2 Sequence Analysis of the *K7* cDNA and Identification of a K7 Open Reading Frame

The compiled *K7* cDNA sequence, sequenced in both orientations, and the deduced amino acid sequence is presented in Figure 4.2. The *K7* cDNA was 6583 bp with an open reading frame (ORF) of 2118 amino acids from position 71 bp to 6424 bp of the nucleotide sequence. An in frame stop codon at base pair position 41 (bold/double underline; Figure 4.2) within the 5' untranslated region (UTR) established the identity of the presumptive ATG translation initiation site (bold/boxed, Figure 4.2), 20 bp downstream of this stop codon. While the sequence contained within the S1 cDNA clone was not complete at the 3' end, as a poly-$(A)^+$ tail was not present, a consensus polyadenylation element was found within the 3' UTR (bold/shaded; Figure 4.2). The identity of an additional 10 bp of 3' UTR sequence and the poly-$(A)^+$ tail was deduced from the sequence of the original dd-PCR clone (S. Sharma, unpublished).

The *K7* translation product was predicted to be slightly acidic with an isoelectric point of 6.63 and have an apparent molecular weight of 233,185. Hydrophobic residues were the most prevalent amino acids in the K7 protein, with leucine contributing 15.38% and alanine contributing 9.95% of all residues. Analysis of the putative K7 protein, using the PROSITE data base and the protein structure prediction server of Fischer and Eisenberg (1996) (2.3.33) did not indicate the presence of homologies to known protein motifs or amino acid repeats which could be indicative of potential biological function, although two potential bipartite nuclear localisation signals (NLS) (Robbins *et al.*, 1991) were identified one amino acid apart in the K7 ORF (boxed, Figure 4.2; Figure 4.4), indicating that the protein may be nuclear localised. Two hydrophobic stretches, from amino acid position 354-374 and 847-867 were also identified (shaded/boxed, Figure 4.2) using the simple modular architecture research tool (SMART) (Shultz *et al.*, 1998), and may represent putative transmembrane domains.

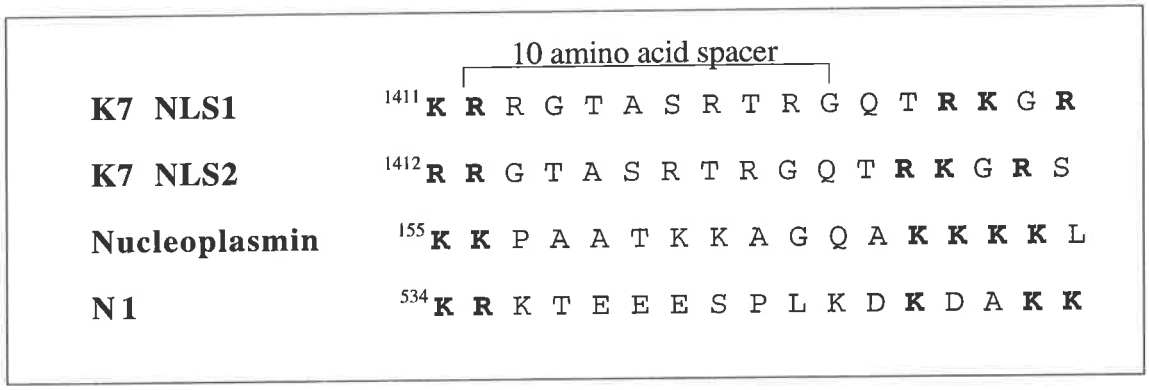


Figure 4.4: The two putative bipartite nuclear localisation sequences (NLSs) of K7, from amino acid 1411 to 1427 and 1412 to 1428, are aligned with the bipartite NLSs of known function (nucleoplasmin and N1 from *Xenopus*) (Robbins *et al.*, 1991). The consensus is loosely defined as 2 basic residues (bold), followed by a spacer of 10 amino acids and a second basic cluster in which at least 3 of 5 residues are basic.

4.2.3 *K7* cDNA and Protein Sequence Comparisons

4.2.3.1 DNA sequence alignments

Database comparison at the DNA level (2.3.33) indicated that *K7* displayed significant identity to a human mRNA for the *KIAA0165* gene (Accession number D79987), isolated from a human immature myeloid cell line (Nagase *et al.*, 1996). Alignment of the sequences using the ALIGN program indicated that these two sequences were 79.3% identical, with identity extending across the entire DNA sequence and the 5' and 3' UTR's (Figure 4.5), suggesting that this gene was the likely human homologue.

The presumptive ATG initiation codon of *K7* was conserved in *KIAA0165* (bold/shaded, Figure 4.5), with high sequence homology observed downstream of this region, but more divergent homology observed upstream of the presumptive ATG codon. However, the *KIAA0165* cDNA included 81 bp of additional sequence from nucleotide position 140 to 220 which was not present in the *K7* cDNA (Figure 4.5). Within this additional sequence in the human gene, two in frame termination codons (TGA at position 205 bp; TAA at position 216 bp; bold, Figure 4.5) were present, with the latter lying downstream of a second potential ATG initiation codon. Therefore, translation of *KIAA0165* was predicted by Nagase *et al.* (1996) to initiate at the next in frame ATG codon at nucleotide position 1114 (bold/shaded, Figure 4.5), resulting in a protein which was truncated at its N-terminus by 325 amino acids relative to the predicted *K7* amino acid sequence. Analysis of the 5' and 3' regions of the additional *KIAA0165* sequence indicated the presence of a consensus 5' splice donor site and near consensus 3' splice acceptor site (Figure 4.5 inset). The additional sequence found in the *KIAA0165* cDNA may therefore represent an unspliced intron. Whether such a splicing event was present in *K7* was not investigated.

4.2.3.2 Protein sequence alignments

Database comparisons (2.3.33) demonstrated that the putative K7 protein also contained significant homology to the deduced KIAA0165 amino acid sequence (Figure 4.6). Comparison of the predicted KIAA0165 amino acid sequence, following removal of the 81 bp insertion which was predicted to be intronic sequence, with the predicted K7 amino acid sequence indicated that these two proteins utilised a conserved translational start site and were 77% identical and 85%

Figure 4.5: Alignment and comparison of *K7* and *KIAA0165* nucleotide sequences

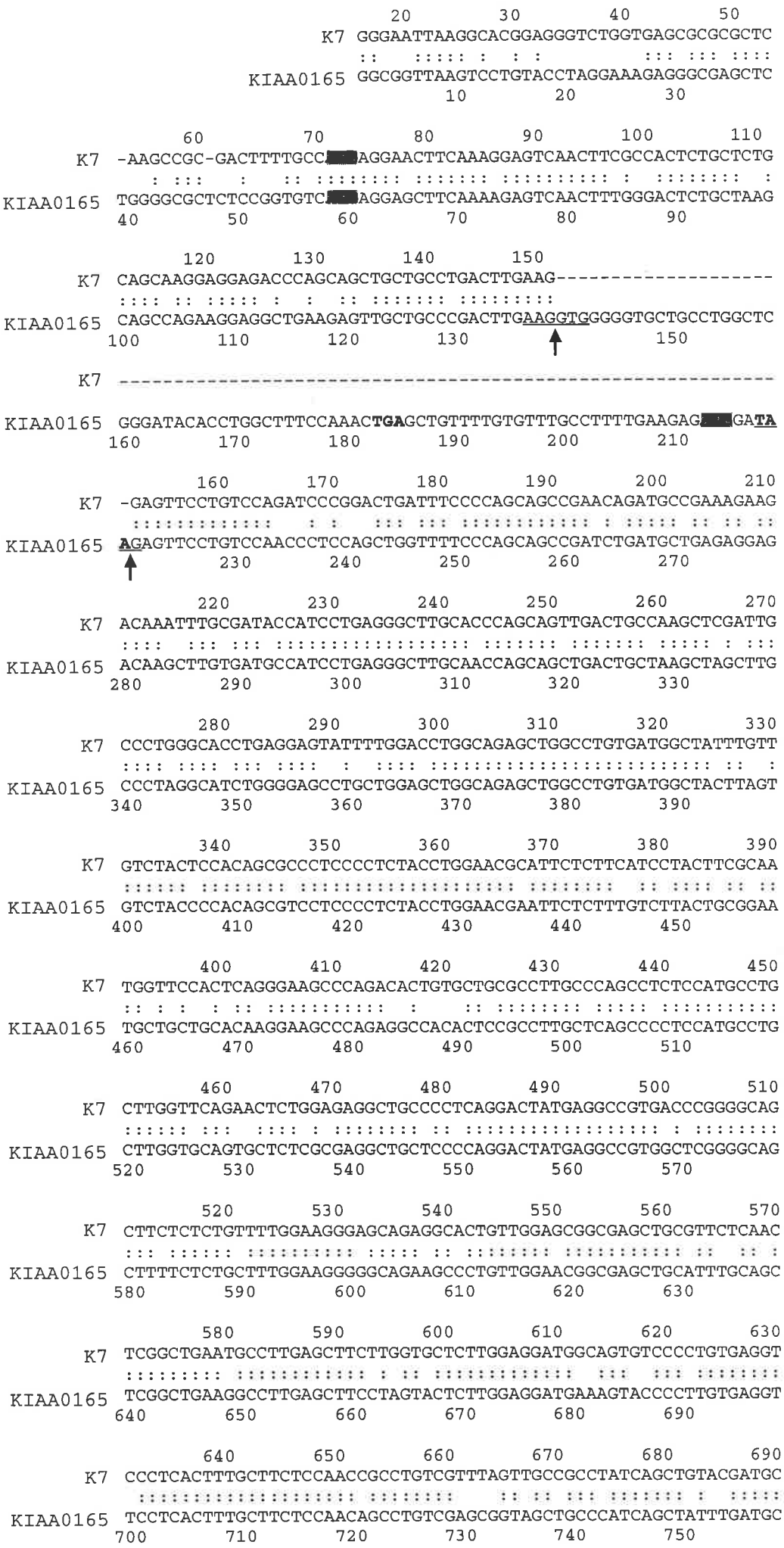
Schematic alignment of the *K7* DNA sequence with the *KIAA0165* DNA sequence. The ATG initiation codons are indicated by bold type and shading, while in frame stop codons are indicated by bold type alone. The predicted *K7* translational start site (bold/shaded) is conserved in *KIAA0165*, but is separated from the remaining ORF by an 81 bp insertion, which consists of two in frame termination codons (bold), one downstream of a second predicted ATG initiation codon. *KIAA0165* translation has thus been predicted to start from the next in frame ATG at nucleotide position 1112 (Nagase *et al.*, 1996) (bold/shaded), some 1052 bp downstream of the predicted *K7* translational start site. The 5' and 3' ends of the 81 bp insertion sequence are aligned with the 5' and 3' splice site consensus sequence (Alberts *et al.*, 1994) in the inset, and underlined in the figure, with the putative splice sites of the 81 bp insertion indicated in the figure using arrows.

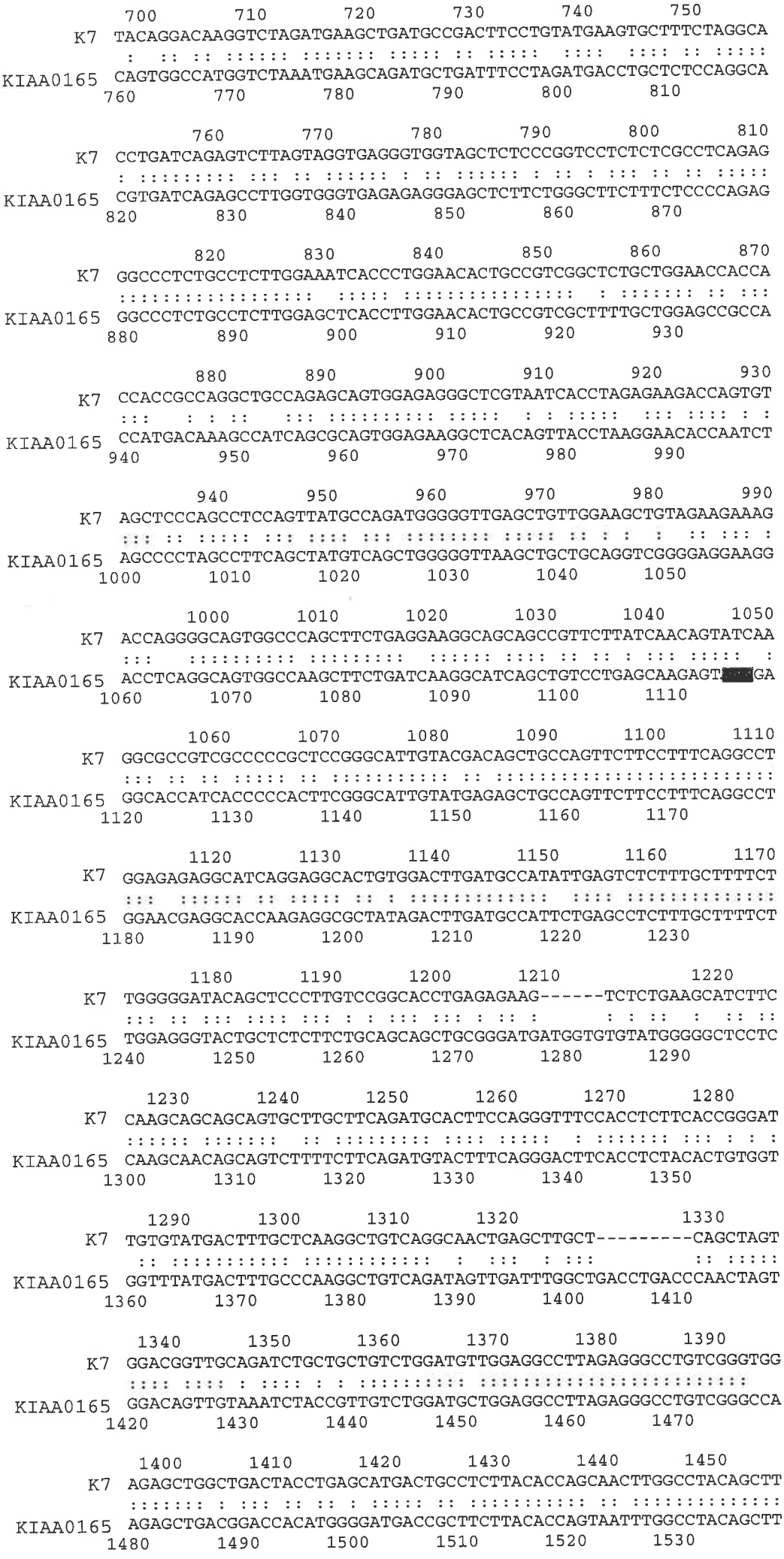
Inset to Figure 4.5: Alignment of the consensus 5' and 3' splice site sequences with the 5' and 3' ends of the *KIAA0165* insertion sequence. Bold nucleotides represent nearly invariant dinucleotides at either end of the intron sequence.

Consensus 5' splice donor site and 3' acceptor site:

5' exon sequence		intron sequence		3' exon sequence
C 5' -- or A G A	**G T** or	A G	--------N--------	or **A G** C T	G or -- 3' A

KIAA0165: A A G **G T** G --------N-------- T **A A** G


K7 GGGAATTAAGGCACGGAGGGTCTGGTGAGCGCGCGCTC
KIAA0165 GGCGGTTAAGTCCTGTACCTAGGAAAGAGGGCGAGCTC
K7 -AAGCCGC-GACTTTTGCC[redacted]AGGAACTTCAAAGGAGTCAACTTCGCCACTCTGCTCTG
KIAA0165 TGGGGCGCTCTCCGGTGTC[redacted]AGGAGCTTCAAAAGAGTCAACTTTGGGACTCTGCTAAG
K7 CAGCAAGGAGGAGACCCAGCAGCTGCTGCCTGACTTGAAG----------
KIAA0165 CAGCCAGAAGGAGGCTGAAGAGTTGCTGCCCGACTTGAAGGTGGGGGTGCTGCCTGGCTC
K7 ----------
KIAA0165 GGGATACACCTGGCTTTCCAAACTGAGCTGTTTTGTGTTTGCCTTTTGAAGAG[redacted]GATA
K7 -GAGTTCCTGTCCAGATCCCGGACTGATTTCCCCAGCAGCCGAACAGATGCCGAAAGAAG
KIAA0165 AGAGTTCCTGTCCAACCCTCCAGCTGGTTTTCCCAGCAGCCGATCTGATGCTGAGAGGAG
K7 ACAAATTTGCGATACCATCCTGAGGGCTTGCACCCAGCAGTTGACTGCCAAGCTCGATTG
KIAA0165 ACAAGCTTGTGATGCCATCCTGAGGGCTTGCAACCAGCAGCTGACTGCTAAGCTAGCTTG
K7 CCCTGGGCACCTGAGGAGTATTTTGGACCTGGCAGAGCTGGCCTGTGATGGCTATTTGTT
KIAA0165 CCCTAGGCATCTGGGGAGCCTGCTGGAGCTGGCAGAGCTGGCCTGTGATGGCTACTTAGT
K7 GTCTACTCCACAGCGCCCTCCCCTCTACCTGGAACGCATTCTCTTCATCCTACTTCGCAA
KIAA0165 GTCTACCCCACAGCGTCCTCCCCTCTACCTGGAACGAATTCTCTTTGTCTTACTGCGGAA
K7 TGGTTCCACTCAGGGAAGCCCAGACACTGTGCTGCGCCTTGCCCAGCCTCTCCATGCCTG
KIAA0165 TGCTGCTGCACAAGGAAGCCCAGAGGCCACACTCCGCCTTGCTCAGCCCCTCCATGCCTG
K7 CTTGGTTCAGAACTCTGGAGAGGCTGCCCCTCAGGACTATGAGGCCGTGACCCGGGGCAG
KIAA0165 CTTGGTGCAGTGCTCTCGCGAGGCTGCTCCCCAGGACTATGAGGCCGTGGCTCGGGGCAG
K7 CTTCTCTCTGTTTTGGAAGGGAGCAGAGGCACTGTTGGAGCGGCGAGCTGCGTTCTCAAC
KIAA0165 CTTTTCTCTGCTTTGGAAGGGGGCAGAAGCCCTGTTGGAACGGCGAGCTGCATTTGCAGC
K7 TCGGCTGAATGCCTTGAGCTTCTTGGTGCTCTTGGAGGATGGCAGTGTCCCCTGTGAGGT
KIAA0165 TCGGCTGAAGGCCTTGAGCTTCCTAGTACTCTTGGAGGATGAAAGTACCCCTTGTGAGGT
K7 CCCTCACTTTGCTTCTCCAACCGCCTGTCGTTTAGTTGCCGCCTATCAGCTGTACGATGC
KIAA0165 TCCTCACTTTGCTTCTCCAACAGCCTGTCGAGCGGTAGCTGCCCATCAGCTATTTGATGC


700 710 720 730 740 750
K7 TACAGGACAAGGTCTAGATGAAGCTGATGCCGACTTCCTGTATGAAGTGCTTTCTAGGCA
KIAA0165 CAGTGGCCATGGTCTAAATGAAGCAGATGCTGATTTCCTAGATGACCTGCTCTCCAGGCA
760 770 780 790 800 810
760 770 780 790 800 810
K7 CCTGATCAGAGTCTTAGTAGGTGAGGGTGGTAGCTCTCCCGGTCCTCTCTCGCCTCAGAG
KIAA0165 CGTGATCAGAGCCTTGGTGGGTGAGAGAGGGAGCTCTTCTGGGCTTCTTTCTCCCCAGAG
820 830 840 850 860 870
820 830 840 850 860 870
K7 GGCCCTCTGCCTCTTGGAAATCACCCTGGAACACTGCCGTCGGCTCTGCTGGAACCACCA
KIAA0165 GGCCCTCTGCCTCTTGGAGCTCACCTTGGAACACTGCCGTCGCTTTTGCTGGAGCCGCCA
880 890 900 910 920 930
880 890 900 910 920 930
K7 CCACCGCCAGGCTGCCAGAGCAGTGGAGAGGGCTCGTAATCACCTAGAGAAGACCAGTGT
KIAA0165 CCATGACAAAGCCATCAGCGCAGTGGAGAAGGCTCACAGTTACCTAAGGAACACCAATCT
940 950 960 970 980 990
940 950 960 970 980 990
K7 AGCTCCCAGCCTCCAGTTATGCCAGATGGGGGTTGAGCTGTTGGAAGCTGTAGAAGAAAG
KIAA0165 AGCCCCTAGCCTTCAGCTATGTCAGCTGGGGGTTAAGCTGCTGCAGGTCGGGGAGGAAGG
1000 1010 1020 1030 1040 1050
1000 1010 1020 1030 1040 1050
K7 ACCAGGGGCAGTGGCCCAGCTTCTGAGGAAGGCAGCAGCCGTTCTTATCAACAGTATCAA
KIAA0165 ACCTCAGGCAGTGGCCAAGCTTCTGATCAAGGCATCAGCTGTCCTGAGCAAGAGT[redacted]GA
1060 1070 1080 1090 1100 1110
1060 1070 1080 1090 1100 1110
K7 GGCGCCGTCGCCCCCGCTCCGGGCATTGTACGACAGCTGCCAGTTCTTCCTTTCAGGCCT
KIAA0165 GGCACCATCACCCCCACTTCGGGCATTGTATGAGAGCTGCCAGTTCTTCCTTTCAGGCCT
1120 1130 1140 1150 1160 1170
1120 1130 1140 1150 1160 1170
K7 GGAGAGAGGCATCAGGAGGCACTGTGGACTTGATGCCATATTGAGTCTCTTTGCTTTTCT
KIAA0165 GGAACGAGGCACCAAGAGGCGCTATAGACTTGATGCCATTCTGAGCCTCTTTGCTTTTCT
1180 1190 1200 1210 1220 1230
1180 1190 1200 1210 1220
K7 TGGGGGATACAGCTCCCTTGTCCGGCACCTGAGAGAAG------TCTCTGAAGCATCTTC
KIAA0165 TGGAGGGTACTGCTCTCTTCTGCAGCAGCTGCGGGATGATGGTGTGTATGGGGGCTCCTC
1240 1250 1260 1270 1280 1290
1230 1240 1250 1260 1270 1280
K7 CAAGCAGCAGCAGTGCTTGCTTCAGATGCACTTCCAGGGTTTCCACCTCTTCACCGGGAT
KIAA0165 CAAGCAACAGCAGTCTTTTCTTCAGATGTACTTTCAGGGACTTCACCTCTACACTGTGGT
1300 1310 1320 1330 1340 1350
1290 1300 1310 1320 1330
K7 TGTGTATGACTTTGCTCAAGGCTGTCAGGCAACTGAGCTTGCT---------CAGCTAGT
KIAA0165 GGTTTATGACTTTGCCCAAGGCTGTCAGATAGTTGATTTGGCTGACCTGACCCAACTAGT
1360 1370 1380 1390 1400 1410
1340 1350 1360 1370 1380 1390
K7 GGACGGTTGCAGATCTGCTGCTGTCTGGATGTTGGAGGCCTTAGAGGGCCTGTCGGGTGG
KIAA0165 GGACAGTTGTAAATCTACCGTTGTCTGGATGCTGGAGGCCTTAGAGGGCCTGTCGGGCCA
1420 1430 1440 1450 1460 1470
1400 1410 1420 1430 1440 1450
K7 AGAGCTGGCTGACTACCTGAGCATGACTGCCTCTTACACCAGCAACTTGGCCTACAGCTT
KIAA0165 AGAGCTGACGGACCACATGGGGATGACCGCTTCTTACACCAGTAATTTGGCCTACAGCTT
1480 1490 1500 1510 1520 1530


1460 1470 1480 1490 1500 1510
K7 CTTTAGTCAGAAGCTCTATGAGGAGGCCTGCGTCATCTCAGAGCCAGTCTGTCAGCACTT
KIAA0165 CTATAGTCACAAGCTCTATGCCGAGGCCTGTGCCATCTCTGAGCCGCTCTGTCAGCACCT
1540 1550 1560 1570 1580 1590
1520 1530 1540 1550 1560 1570
K7 GGGCTCGGCAACGTTGGGTGCTTGTCCTGAAGTGCCTCCTGAGAAGCTACATAGGTGCTT
KIAA0165 GGGTTTGGTGAAGCCAGGCACTTATCCCGAGGTGCCTCCTGAGAAGTTGCACAGGTGCTT
1600 1610 1620 1630 1640 1650
1580 1590 1600 1610 1620 1630
K7 CAGGCTGCATGTGGAGAGTTTGAAGAAACTGGGTAAACAGGCCCAGGGCTGCAAGATGGT
KIAA0165 CCGGCTACAAGTAGAGAGTTTGAAGAAACTGGGTAAACAGGCCCAGGGCTGCAAGATGGT
1660 1670 1680 1690 1700 1710
1640 1650 1660 1670 1680 1690
K7 GACCTTATGGCTGGCAGCTCTGAAACCCTACAGTCTGGAACACATGGTCGAGCCAGTCAC
KIAA0165 GATTTTGTGGCTGGCAGCCCTGCAACCCTGTAGCCCTGAACACATGGCTGAGCCAGTCAC
1720 1730 1740 1750 1760 1770
1700 1710 1720 1730 1740 1750
K7 TTTCTGGGTTCGAGTCAAAATGGATGCATCCAGGGCAGGAGACAAGGAGCTACAACTTCA
KIAA0165 TTTCTGGGTTCGGGTCAAGATGGATGCGGCCAGGGCTGGAGACAAGGAGCTACAGCTAAA
1780 1790 1800 1810 1820 1830
1760 1770 1780 1790 1800 1810
K7 GACTTTGCGAGACAGCCTAAGCTGCTGGGACCCAGAGACACAGTCTCTGCTTCTCAGGGA
KIAA0165 GACTCTGCGAGACAGCCTCAGTGGCTGGGACCCGGAGACCCTGGCCCTCCTGCTGAGGGA
1840 1850 1860 1870 1880 1890
1820 1830 1840 1850 1860 1870
K7 GGAGCTGCGGGCCTACAAGGCAGTACGGGCTGACACTGGGCAGGAGCGCTTCAACATCAT
KIAA0165 GGAGCTGCAGGCCTACAAGGCGGTGCGGGCCGACACTGGACAGGAACGCTTCAACATCAT
1900 1910 1920 1930 1940 1950
1880 1890 1900 1910 1920 1930
K7 CTGTGATCTGCTCGAGCTGAGCCCTGAAGAGACAGCAGCTGGAGCCTGGGCACGAGCTAC
KIAA0165 CTGTGACCTCCTGGAGCTGAGCCCCGAGGAGACACCAGCCGGGGCCTGGGCACGAGCCAC
1960 1970 1980 1990 2000 2010
1940 1950 1960 1970 1980 1990
K7 CTACCTGGTGGAACTGGCCCAAGTGCTGTGCTACCACAACTTTACCCAGCAGACCAACTG
KIAA0165 CCACCTGGTAGAACTGGCTCAGGTGCTCTGCTACCACGACTTTACGCAGCAGACCAACTG
2020 2030 2040 2050 2060 2070
2000 2010 2020 2030 2040 2050
K7 CTCTGCCTTGGATGCTGTCCAGGAAGCCCTGCAGCTTCTGGAGTCCGTGAGTCCTGAGGC
KIAA0165 CTCTGCTCTGGATGCTATCCGGGAAGCCCTGCAGCTTCTGGACTCTGTGAGGCCTGAGGC
2080 2090 2100 2110 2120 2130
2060 2070 2080 2090 2100 2110
K7 CCAGGAGCAG-GATCGACTTCTGGATGATAAAGCACAGGCCTTGTTATGGCTTTACATCT
KIAA0165 CCAGGC-CAGAGATCAGCTTCTGGACGATAAAGCACAGGCCTTGCTGTGGCTTTACATCT
2140 2150 2160 2170 2180 2190
2120 2130 2140 2150 2160 2170
K7 GTACCTTAGAGGCCAAGATGCAGGAAGGTATTGAGCGGGATCGGAGAGCCCAGGCCCCTA
KIAA0165 GTACTCTGGAAGCCAAAATACAGGAAGGTATCGAGCGGGATCGGAGAGCCCAGGCCCCTG
2200 2210 2220 2230 2240 2250
2180 2190 2200 2210 2220 2230
K7 GTAACTTGGAGGAATTTGAAGTCAATGACCTGAACTATGAAGATAAACTGCAGGAAGATC
KIAA0165 GTAACTTGGAGGAATTTGAAGTCAATGACCTGAACTATGAAGATAAACTCCAGGAAGATC
2260 2270 2280 2290 2300 2310

```
            2240       2250       2260       2270       2280       2290
      K7  GTTTTCTGTATAGTAGCATTGCCTTTAATCTGGCTGCTGATGCAGCTCAGTCCAAATGCC
          :::: :: :: :::: ::::::::: :: :::::::: ::::: ::::::::::::::::
KIAA0165  GTTTCCTATACAGTAACATTGCCTTCAACCTGGCTGCAGATGCTGCTCAGTCCAAATGCC
         2320       2330       2340       2350       2360       2370

            2300       2310       2320       2330       2340       2350
      K7  TGGACCAGGCCCTGACTCTGTGGAAAGAAGTGCTCACCAAGGGGCGAGCCCCAGCTGTGC
          ::::::: :::::: : :::::::: ::  :::: :: :::::::  :::::::::::
KIAA0165  TGGACCAAGCCCTGGCCCTGTGGAAGGAGCTGCTTACAAAGGGGCAGGCCCCAGCTGTAC
         2380       2390       2400       2410       2420       2430

            2360       2370       2380       2390       2400       2410
      K7  GGTGTCTCCAGCAGACTGCAGCCTCCCTGCAAATCCTAGCGGCCGTCTATCAGTTGGTGG
          ::::::::::::::::: :::::::: ::::: :::::::: ::: :::: ::: ::::::
KIAA0165  GGTGTCTCCAGCAGACAGCAGCCTCACTGCAGATCCTAGCAGCCCTCTACCAGCTGGTGG
         2440       2450       2460       2470       2480       2490

            2420       2430       2440       2450       2460       2470
      K7  CAAAGCCCTTGCAGGCTCTGGAGACCCTCCTGCTCCTGCAGATTGTATCTAAGAGACTGC
          :::::::: :::::::::::::  ::::::::: :: : :::::: ::: ::::::::
KIAA0165  CAAAGCCCATGCAGGCTCTGGAGGTCCTCCTGCTGCTACGGATTGTCTCTGAGAGACTGA
         2500       2510       2520       2530       2540       2550

            2480       2490       2500       2510       2520       2530
      K7  AGGACCACGCAAAGGCAGCCAGCTCTTCCTGCCAGCTTACCCAGCTACTCTTGAACCTTG
          :::::::: : ::::::::  :::: ::::::::  : :::::::: ::: ::: ::: :
KIAA0165  AGGACCACTCGAAGGCAGCTGGCTCCTCCTGCCACATCACCCAGCTCCTCCTGACCCTCG
         2560       2570       2580       2590       2600       2610

            2540       2550       2560       2570       2580       2590
      K7  GCTGTCCCAGCTATGCGCAGTTGTATCTGGAAGAGGCGGAGTCAAGCCTGAGGAGCCTGG
          ::::::::::::::::: :::::  : :::::::::::: :  :: ::::::: :   :: :
KIAA0165  GCTGTCCCAGCTATGCCCAGTTACACCTGGAAGAGGCAGCATCGAGCCTGAAGCATCTCG
         2620       2630       2640       2650       2660       2670

            2600       2610       2620       2630       2640       2650
      K7  ACCAGACAAGTGACGCGTGCCAGCTGCTGTCCCTGACCTGTGCTCTGCTTGGAAGTCAGC
          : ::::: : :::: : : :: ::: :: :::::::::::::: ::::::: ::::::: :
KIAA0165  ATCAGACTACTGACACATACCTGCTCCTTTCCCTGACCTGTGATCTGCTTCGAAGTCAAC
         2680       2690       2700       2710       2720       2730

            2660       2670       2680       2690       2700       2710
      K7  TCTGCTGGGCTTGCCAGAAGGTAACTGCAGGCGTTTCTTTGTTGCTTTCTGTGCTTCGGG
          ::: :::: ::  ::::::::: ::    :: :: ::: :: :::: :::::::::::::
KIAA0165  TCTACTGGACTCACCAGAAGGTGACCAAGGGTGTCTCTCTGCTGCTGTCTGTGCTTCGGG
         2740       2750       2760       2770       2780       2790

            2720       2730       2740       2750       2760       2770
      K7  ATCCTGCCCTCCAGAAGTCATCTAAGGCTTGGTACTTGCTACGTGTCCAGGCCCTGCAAG
          :::::::::::::::::: :: ::::::::::::::::: :::::::::: ::::::
KIAA0165  ATCCTGCCCTCCAGAAGTCCTCCAAGGCTTGGTACTTGCTGCGTGTCCAGGTCCTGCAGC
         2800       2810       2820       2830       2840       2850

            2780       2790       2800       2810       2820       2830
      K7  TTCTGGCTTTTTACCTCAGTCTCTCGTCCAACCTCCTCTCAAGTGCCCTGCGTGAGCAGC
          :  ::::   :::::: :: ::: :::: ::: :::::::    ::::: : :::::::
KIAA0165  TGGTGGCAGCTTACCTTAGCCTCCCGTCAAACAACCTCTCACACTCCCTGTGGGAGCAGC
         2860       2870       2880       2890       2900       2910

            2840       2850       2860       2870       2880       2890
      K7  TCTGGGATCAAGGCTGGCAGACCCCGGAGACGGCGCTGATTGATGCCCACAAGCTCCTCA
          :::: :  :::::::::::::: :: ::::  :: :: :: ::  :::: :::::::::
KIAA0165  TCTGTGCCCAAGGCTGGCAGACACCTGAGATAGCTCTCATAGACTCCCATAAGCTCCTCC
         2920       2930       2940       2950       2960       2970

            2900       2910       2920       2930       2940       2950
      K7  GAAGCATCATCATTCTGCTCATGGGAAGTGATGTGCTCTCCATTCAGAAAGCAGCCACCG
          :::::::::: : ::::: ::::: :::::  : ::::: : ::::::::::::    :
KIAA0165  GAAGCATCATCCTCCTGCTGATGGGCAGTGACATTCTCTCAACTCAGAAAGCAGCTGTGG
         2980       2990       3000       3010       3020       3030

            2960       2970       2980       2990       3000       3010
      K7  AGTCGCCCTTTCTGGACTACGGTGAGAATCTGGTGCAAAAATGGCAGGTTCTCACAGAGG
          :: :  : ::: :::::::: ::::: :::::::: ::::::::::::::::  ::::::
KIAA0165  AGACATCGTTTTTGGACTATGGTGAAAATCTGGTACAAAAATGGCAGGTTCTTTCAGAGG
         3040       3050       3060       3070       3080       3090
```

```
          3020       3030       3040       3050       3060       3070
K7  TGCTGACCTGCTCGGAGAGGCTGGTCGGCCGCCTGGGTCGCCTGGGGAACGTGAGTGAGG
    :::::: ::::::: :::: :::::::: ::: :::::: :::::::: :  :::::::: :
KIAA0165 TGCTGAGCTGCTCAGAGAAGCTGGTCTGCCACCTGGGCCGCCTGGGTAGTGTGAGTGAAG
        3100       3110       3120       3130       3140       3150

          3080       3090       3100       3110       3120       3130
K7  CCAAGGCCTTTTGCTTGGAGGCCCTAAAACTTACGACTAAGCTGCAGATACCTCGACAGT
    :::::::::::::::::::::::::::::::::: :: :::::::::::::: :: ::::
KIAA0165 CCAAGGCCTTTTGCTTGGAGGCCCTAAAACTTACAACAAAGCTGCAGATACCACGCCAGT
        3160       3170       3180       3190       3200       3210

          3140       3150       3160       3170       3180       3190
K7  GTGCCCTGTTCCTTGTACTGAAGGGTGAGCTGGAGCTGGCCCGGGGTGACATTGACCTCT
    :::::::::::::: :: :::::::: ::::::::::::::::::   ::::::::: ::::
KIAA0165 GTGCCCTGTTCCTGGTGCTGAAGGGCGAGCTGGAGCTGGCCCGCAATGACATTGATCTCT
        3220       3230       3240       3250       3260       3270

          3200       3210       3220       3230       3240       3250
K7  GTCAGTCAGACTTGCAGCAGGTTCTCTTCTTGCTCGAGTCTTCCACAGAGTTTGGTGTTG
    ::::::: ::: :::::::::::::: :::::::: ::::::: ::::::::::::::  :
KIAA0165 GTCAGTCGGACCTGCAGCAGGTTCTGTTCTTGCTTGAGTCTTGCACAGAGTTTGGTGGGG
        3280       3290       3300       3310       3320       3330

          3260       3270       3280       3290       3300       3310
K7  TGACCCAGCATCCAGACTCAGTGAAGAAGGTGCACACGCAGAAAGGGAAGCATAAGGCTC
    :::: ::::: :  ::::: :::::::::::: :::  :::::: ::::::::  :::: :
KIAA0165 TGACTCAGCACCTGGACTCTGTGAAGAAGGTCCACCTGCAGAAGGGGAAGCAGCAGGCCC
        3340       3350       3360       3370       3380       3390

          3320       3330       3340       3350       3360       3370
K7  AAGGCCCATGTTTCCCACCCCTGTCAGAGGAGGAGCCCTTCCTGAAAGGTCCTGCCCTTG
    : : ::: :::   ::::  ::  :::::::::::: :::::: : ::: ::::: :: :
KIAA0165 AGGTCCCCTGTCCTCCACAGCTCCCAGAGGAGGAGCTCTTCCTAAGAGGCCCTGCTCTAG
        3400       3410       3420       3430       3440       3450

          3380       3390       3400       3410       3420       3430
K7  AGCTGGTGGACACTGTGCTCAACGAACCTGGTCCCATCCAGTCCTCTGTAAACTCCTCCC
    ::::::::: ::::::::  ::: :: ::::: :::::     : :::  ::::::::::
KIAA0165 AGCTGGTGGCCACTGTGGCCAAGGAGCCTGGCCCCATAGCACCTTCTACAAACTCCTCCC
        3460       3470       3480       3490       3500       3510

          3440       3450       3460       3470       3480       3490
K7  CAGTCCTGAAAACCAAGCCACCACCCAACCCTGGCTTCCTGTCTCACTTGCCCAGCTGTG
    ::::: :::::::::::::  :  ::::  ::   ::::::::: :: :  :::: :::::
KIAA0165 CAGTCTTGAAAACCAAGCCCCAGCCCATACCCAACTTCCTGTCCCATTCACCCACCTGTG
        3520       3530       3540       3550       3560       3570

          3500       3510       3520       3530       3540       3550
K7  ACTGCTTACTGTGTGCCAGCCCTGCCCTCTCAGCAGTCTGTCTGCGCTGGGTGCTGGTCA
    ::::::  :: :: :::::::::: :::: ::::::::::::::::::::::  ::::::
KIAA0165 ACTGCTCGCTCTGCGCCAGCCCTGTCCTCACAGCAGTCTGTCTGCGCTGGGTATTGGTCA
        3580       3590       3600       3610       3620       3630

          3560       3570       3580       3590       3600       3610
K7  CTGCCGGAGTGAGGCTGGCCACCGGTCACAAAGCTCAGGGTCTGGATCTGCTGCAAGCTG
    : :: :: :::::::::::::  :: ::: :::: :::::::::::::::::::: :  :
KIAA0165 CGGCAGGGGTGAGGCTGGCCATGGGCCACCAAGCCCAGGGTCTGGATCTGCTGCAGGTCG
        3640       3650       3660       3670       3680       3690

          3620       3630       3640       3650       3660       3670
K7  TGCTGACGCGTTGCCCTGCAGCCACCAAGCGCTTCACACAGAGTCTCCAAGCTTCTCTGA
    :::::: : : :: :::: :::: :  ::::: :::: ::   :::::::::::: ::::
KIAA0165 TGCTGAAGGGCTGTCCTGAAGCCGCTGAGCGCCTCACCCAAGCTCTCCAAGCTTCCCTGA
        3700       3710       3720       3730       3740       3750

          3680       3690       3700       3710       3720       3730
K7  ATCACAGAACAACCCCCTCTTGTGTTCCCAGCCTCTTTGATGAGATCATGGCTCAAGTGT
    :::: : :::: :::::::: :  ::::: :::::::: ::::::::: :::::::::  :
KIAA0165 ATCATAAAACACCCCCCTCCTTGGTTCCAAGCCTCTTGGATGAGATCTTGGCTCAAGCAT
        3760       3770       3780       3790       3800       3810

          3740       3750       3760       3770       3780       3790
K7  ACACGCACTTGGCATTGGAGTTCCTGAATCAGACATCTGAAAAGAGCCTGGGGAAGGTCC
    :::: :  :::::: :::::  :::::: ::: ::::  :  ::::::::  :::::: :
KIAA0165 ACACACTGTTGGCACTGGAGGGCCTGAACCAGCCATCAAACGAGAGCCTGCAGAAGGTTC
        3820       3830       3840       3850       3860       3870
```

```
            3800       3810       3820       3830       3840       3850
      K7  TGGCGTCGGGGCTGAAGTTTGTGGCAACACGGATACAGAGCTTGGAGATCTGGAGAGCCC
          :   ::: :::::::::::::: ::: :::::::::    : : :::  :::: :::::
KIAA0165  TACAGTCAGGGCTGAAGTTTGTAGCAGCACGGATACCCCACCTAGAGCCCTGGCGAGCCA
         3880       3890       3900       3910       3920       3930

            3860       3870       3880       3890       3900       3910
      K7  ACCTGCTCTTGGTTCAGGCCCTTGCAAAGCTGGCCCACTTCAGCTGCTGTACTTCCGAAC
          :::::::::: ::  ::::::  :::: :: :    : ::::::::::::::: :: :::
KIAA0165  GCCTGCTCTTGATTTGGGCCCTCACAAAACTAGGTGGCCTCAGCTGCTGTACTACCCAAC
         3940       3950       3960       3970       3980       3990

            3920       3930       3940       3950       3960       3970
      K7  TCTTTGCAAGCTCCTGGGGCTGGCACCCACCATTAGTAAAAAGTCTGCCAGTCTTAGAGC
          : :::::::::::::::::::::: ::::::::: :::::::: : :: : :: :::::
KIAA0165  TTTTTGCAAGCTCCTGGGGCTGGCAGCCACCATTAATAAAAAGTGTCCCTGGCTCAGAGC
         4000       4010       4020       4030       4040       4050

            3980       3990       4000       4010       4020       4030
      K7  CTGCTAAAATTCGGAGACAAAAATGCTCTGGACGGGGCCGCCGGCGGATAGCCTCTGTTC
          :  :::: : :: : : :::::: : :::::::: :: ::::    : ::::::::: ::
KIAA0165  CCTCTAAGACTCAGGGCCAAAAACGTTCTGGACGAGGGCGCCAAAAGTTAGCCTCTGCTC
         4060       4070       4080       4090       4100       4110

            4040       4050       4060       4070       4080       4090
      K7  CCCCACCCCTCCATAACAGCTCTCAGAAAGGTCTGGAAGAGGAAGGGCCACCTTGTACCC
          :::    :::: :::: : ::::::::::::::::::    ::: :  :: :: :: : :
KIAA0165  CCCTGAGCCTCAATAATACCTCTCAGAAAGGTCTGGAAGGTAGAGGACTGCCCTGCACAC
         4120       4130       4140       4150       4160       4170

            4100       4110       4120       4130       4140       4150
      K7  CTAAGCCTCCAGGTCGGGCTAGGCAAGCTGGCCCTCGTGTCCCTTTCACCATATTTGAAG
          :::: :: ::::  :::   ::::::::::::::: :::::: :::::  : ::::: :
KIAA0165  CTAAACCCCCAGACCGGATCAGGCAAGCTGGCCCTCATGTCCCCTTCACGGTGTTTGAGG
         4180       4190       4200       4210       4220       4230

            4160       4170       4180       4190       4200       4210
      K7  AAGTCCACCCCACCAAGAGCAAGCTCCAGGTTCCCTTGGCTCCCAGGGTCCACAGGCGGG
          :::::  ::: ::  :::::::::   : :: :::  ::: :::::::: ::   : : :
KIAA0165  AAGTCTGCCCTACAGAGAGCAAGCCTGAAGTACCCCAGGCCCCCAGGGTACAACAGAGAG
         4240       4250       4260       4270       4280       4290

            4220       4230       4240       4250       4260       4270
      K7  CCCAGACTCGCCTCAAGGTGATCTTCAGTGATGACAGTGACCTGGAAGACCTTGTCTCAG
          ::::::: :::::::::::: :::::::::::::::::::: ::::::::: ::::::::
KIAA0165  TCCAGACGCGCCTCAAGGTGAACTTCAGTGATGACAGTGACTTGGAAGACCCTGTCTCAG
         4300       4310       4320       4330       4340       4350

            4280       4290       4300       4310       4320       4330
      K7  CTGACACACAACTTGTAGAGGAGCCTAAGAGGCGAGGGACTGCTTCCCGTACCCGGGGCC
          ::::  :    :: : :::::::::::::::: :: :: :::::::::::   :::::: :
KIAA0165  CTGAGGCCTGGCTGGCAGAGGAGCCTAAGAGACGGGGCACTGCTTCCCGGGGCCGGGGGC
         4360       4370       4380       4390       4400       4410

            4340       4350       4360       4370       4380       4390
      K7  AAACTAGGAAGGGACGTAGCTTAAAGACAGATGCGGTGGTTGCTATAGAAAGTACCCCTG
          : : :::::::: :  ::: :::::::: ::::: ::::::::   ::  ::: ::::::
KIAA0165  GAGCAAGGAAGGGCCTGAGCCTAAAGACGGATGCCGTGGTTGCCCCAGGTAGTGCCCCTG
         4420       4430       4440       4450       4460       4470

            4400       4410       4420       4430       4440       4450
      K7  GACATTCCAGTGTCAGTGGGAGGACTCGGAGGGCCAGGAAGGTGGCTTCAAGAAATTGTG
          :  :  :  :  : : ::: ::::  :::::::::: ::::::::: :::::: ::::::
KIAA0165  GGAACCCTGGCCTGAATGGCAGGAGCCGGAGGGCCAAGAAGGTGGCATCAAGACATTGTG
         4480       4490       4500       4510       4520       4530

            4460       4470       4480       4490       4500       4510
      K7  AAGAAGAGAGTCCCAA-AGCACCCCTCTGTGTCTGGGCCAGCCAGGGACCCGAGATCATG
          : ::   : ::::: : ::  ::     ::: :  :::::: :  :: :: :::::::::
KIAA0165  AGGAGCGGCGTCCCCAGAGGGCCA----GTGACCAGGCCAGGCCTGGCCCTGAGATCATG
         4540       4550       4560          4570       4580       4590

            4520       4530       4540       4550       4560       4570
      K7  AGAAGTATCCCTGAAGAGGAGCCAGTGGACAACCATCTGGAAAAA--AGCTTCGAGATTC
          :: :  :::::::: :: :: :     ::::::  :   ::::::  ::::: :::::: :
KIAA0165  AGGACCATCCCTGAGGAAGAACTGACTGACAAC--TGGAGAAAAATGAGCTTTGAGATCC
             4600       4610       4620         4630       4640       4650
```

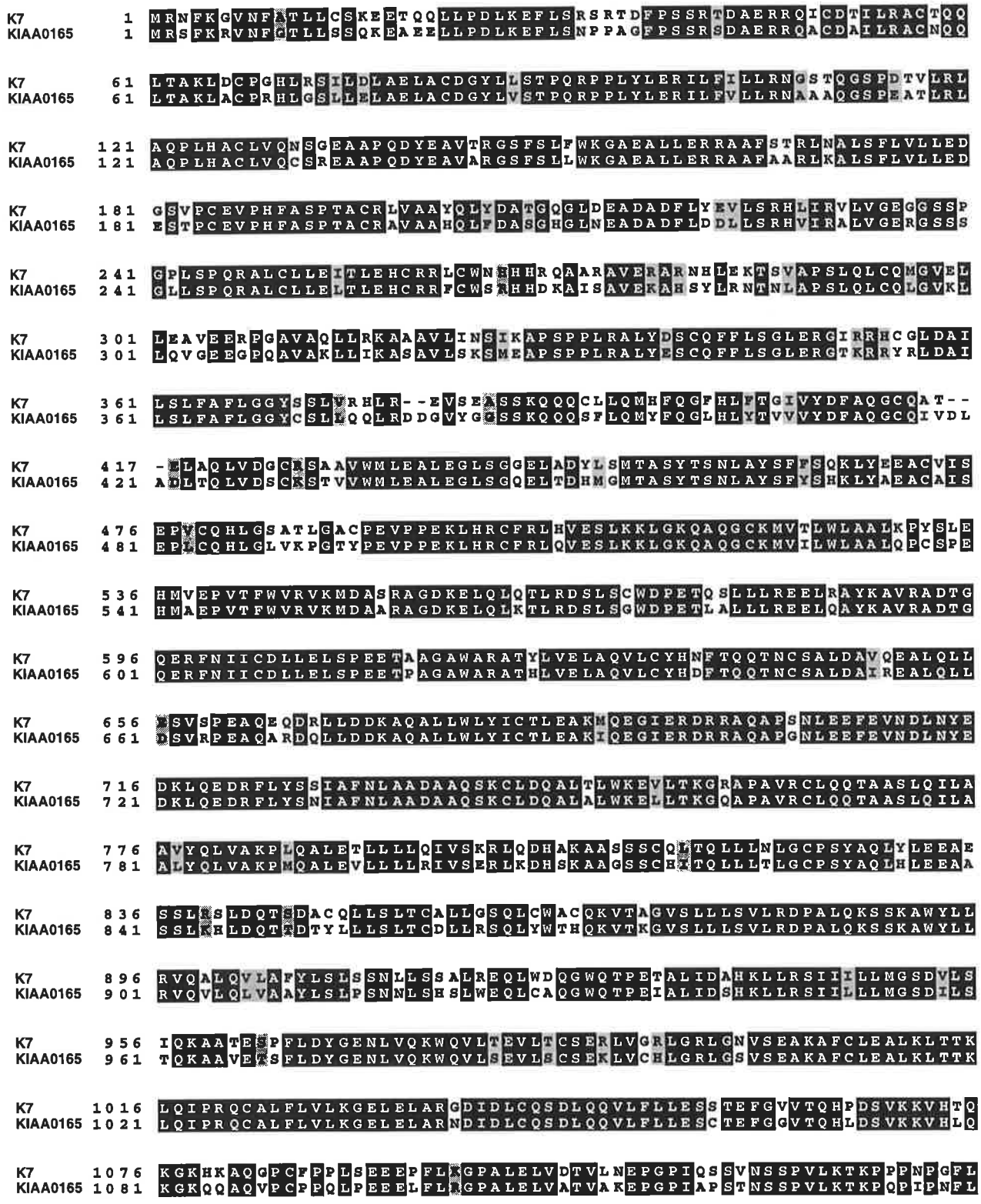
```
                4580       4590       4600       4610       4620       4630
      K7  TCAGGGGTTCTGATGGGGAAGACTCGGCCTCAGGTGAGAAGGCGGCAGCTGCAGACACTG
          ::::::: ::::: :::::::::: ::::::::: :::: :  ::::: : : : :::
KIAA0165  TCAGGGGCTCTGACGGGGAAGACTCAGCCTCAGGTGGGAAGACTCCAGCTCCGGGCCCTG
                4660       4670       4680       4690       4700       4710

                4640       4650       4660       4670       4680       4690
      K7  GTCTACCCGTAGGAGAGTGTGAGGTGTTGAGACGGGACTCCAGCAAGGCAGAGAGACCTG
            : :    ::::: :: ::: :: :::: : ::: :::::::::    ::   ::
KIAA0165  AGGCAGCTTCTGGAGAATGGGAGCTGCTGAGGCTGGATTCCAGCAAGAAGAAGCTGCCCA
                4720       4730       4740       4750       4760       4770

                4700       4710       4720       4730       4740       4750
      K7  TCTTGTACTCAGACACGGAGGCCAATAGTGACCCTAGTCCTTGGCTCCCACCCTTCTCAG
           :   : : :::::: ::::    : :  :::: : ::::: ::::::  : :  :::::
KIAA0165  GCCCATGCCCAGACAAGGAGAGTGACAAGGACCTTGGTCCTCGGCTCCAGCTCCCCTCAG
                4780       4790       4800       4810       4820       4830

                4760       4770       4780       4790       4800       4810
      K7  TTCCTGCACCCATTGATCTTTCTACCCTGGATTCCATTTCTGACTCATTGAGCATTGCTT
            :: : : ::: :: :::::::::::::: ::::: : ::::::  ::::  ::::::
KIAA0165  CCCCCGTAGCCACTGGTCTTTCTACCCTGGACTCCATCTGTGACTCCCTGAGTGTTGCTT
                4840       4850       4860       4870       4880       4890

                4820       4830       4840       4850       4860       4870
      K7  TCCGTGGTGTTAGTCACTGCCCTCCAAGTGGGCTCTATGCTCACCTCTGCCGCTTCTTGG
          :::: ::  :::::::::: ::::: :::::::::::::: ::::::::::::::: :::
KIAA0165  TCCGGGGCATTAGTCACTGTCCTCCTAGTGGGCTCTATGCCCACCTCTGCCGCTTCCTGG
                4900       4910       4920       4930       4940       4950

                4880       4890       4900       4910       4920       4930
      K7  CCCTGTGTCTGGGCCACCGAGATCCTTATGCCACTGCGTTCCTTGTTGCTGAGTCTATCT
          :: :::: :::::::::: :::::::::::::::::: ::::::::  : :::::: :::
KIAA0165  CCTTGTGCCTGGGCCACCGGGATCCTTATGCCACTGCTTTCCTTGTCACCGAGTCTGTCT
                4960       4970       4980       4990       5000       5010

                4940       4950       4960       4970       4980       4990
      K7  CCATCACCTGCCGTCACCAGCTGCTCACCCACTTGCACCGACAGCTCAGCAAGGCCCAGA
          :::::::::: :: :::::::::::::::::: : ::: ::::::::::::::::::::
KIAA0165  CCATCACCTGTCGCCACCAGCTGCTCACCCACCTCCACAGACAGCTCAGCAAGGCCCAGA
                5020       5030       5040       5050       5060       5070

                5000       5010       5020       5030       5040       5050
      K7  AGCAGCAAGAATCACCTGAACTGGCAGAGCATCTGCAGAGGCTGGACCTGAAGGAGAGGC
          :::: : :: ::::: :::: : ::::: :: :::::: ::::::  :::  :::::: ::
KIAA0165  AGCACCGAGGATCACTTGAAATAGCAGACCAGCTGCAGGGGCTGAGCCTTCAGGAGATGC
                5080       5090       5100       5110       5120       5130

                5060       5070       5080       5090       5100       5110
      K7  CCAGAGGTGTCCCACTGGCCCGTATCCAGCGCCTCTTTTCCTTTAAGGCTTTGGGATCTG
          :  ::: :::::: :::::::: :::::::::::::::::::: : :::::::: :::::
KIAA0165  CTGGAGATGTCCCCCTGGCCCGCATCCAGCGCCTCTTTTCCTTCAGGGCTTTGGAATCTG
                5140       5150       5160       5170       5180       5190

                5120       5130       5140       5150       5160       5170
      K7  GCTGCTTCCCCCAGGCAGAGAAGGAGAGTTTCCAGGAGCGCCTGGCTCTGATCCCTAGTG
          ::  :::::::::: : :: ::::::::::::::::::::::::::::::::::: ::::
KIAA0165  GCCACTTCCCCCAGCCTGAAAAGGAGAGTTTCCAGGAGCGCCTGGCTCTGATCCCCAGTG
                5200       5210       5220       5230       5240       5250

                5180       5190       5200       5210       5220       5230
      K7  GGGTGACTGTCTGTGTGTTGGCTCTGGCCACCCTCCAGCCTGGAACCTTGAGCAACACTC
          :::::::::: :::::::::: :::::::::::::::: :::::: :: :::::::  :
KIAA0165  GGGTGACTGTGTGTGTGTTGGCCCTGGCCACCCTCCAGCCCGGAACCGTGGGCAACACCC
                5260       5270       5280       5290       5300       5310

                5240       5250       5260       5270       5280       5290
      K7  TTTTACTGACCCGCCTGGAAAAGGACAACCCCCCAATTACTGTGAAGATCCCCACTGCCC
          :  : :::::::: ::::::::::::  :::::: : : :::: :::: :::::::: ::
KIAA0165  TCCTGCTGACCCGGCTGGAAAAGGACAGTCCCCCAGTCAGTGTGCAGATTCCCACTGGCC
                5320       5330       5340       5350       5360       5370

                5300       5310       5320       5330       5340       5350
      K7  AAAATAAGCTCCCTCTAAGCGCAGTGCTGAAAGAGTTTGATGCCATCCAGAAGGACCAAA
          : :: ::::: : :::  :  :::: ::::: ::::::::::::::::::::::  :: :
KIAA0165  AGAACAAGCTTCATCTGCGTTCAGTCCTGAATGAGTTTGATGCCATCCAGAAGGCACAGA
                5380       5390       5400       5410       5420       5430
```

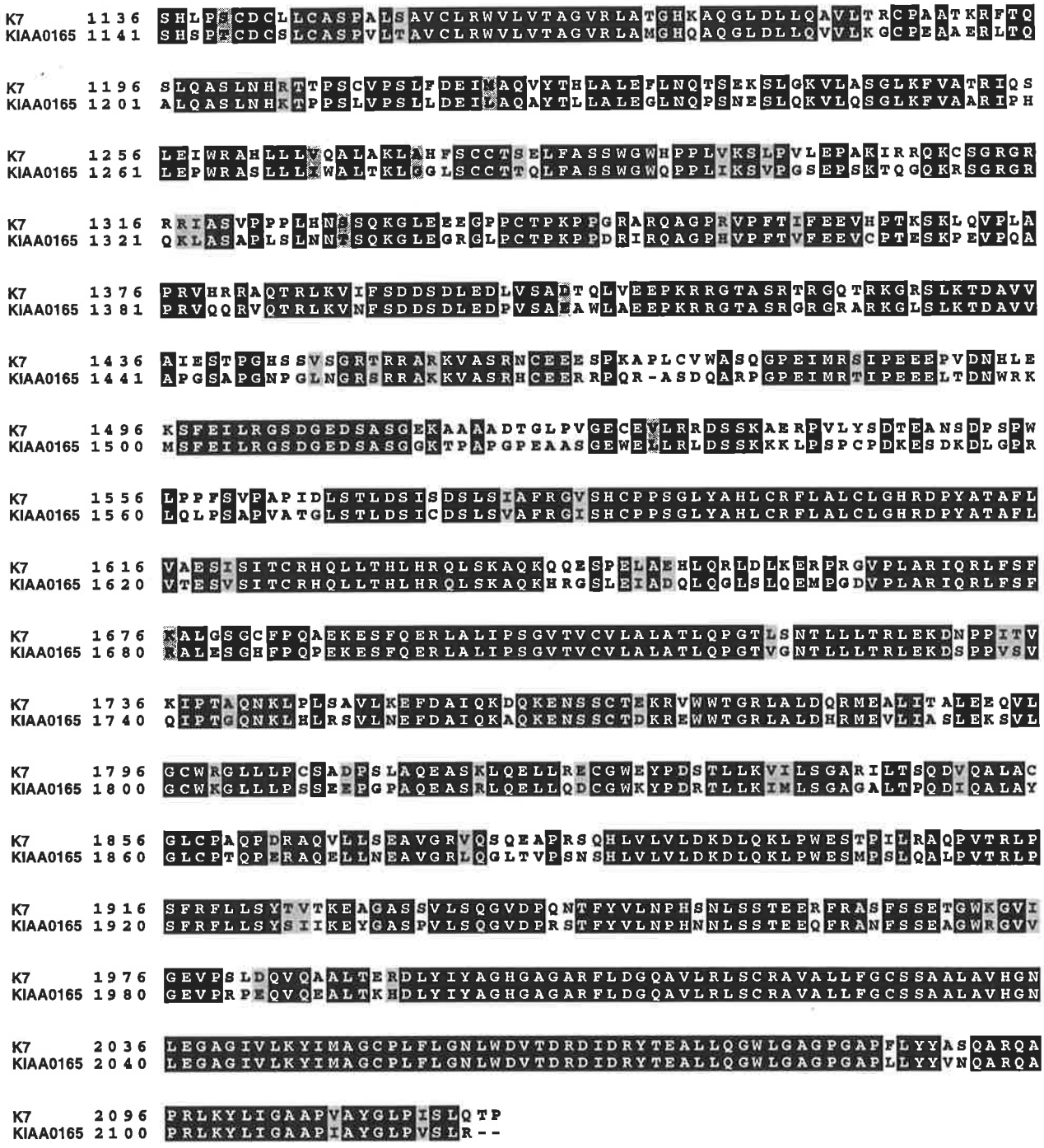
```
                5360       5370       5380       5390       5400       5410
      K7  AAGAAAACAGCAGCTGTACTGAGAAACGAGTATGGTGGACGGGCCGGCTGGCACTGGACC
          :::: :::::::::::::::: :: :::: ::::::::: :: ::::::::::::::::
KIAA0165  AAGAGAACAGCAGCTGTACTGACAAGCGAGAATGGTGGACAGGGCGGCTGGCACTGGACC
                5440       5450       5460       5470       5480       5490

                5420       5430       5440       5450       5460       5470
      K7  AAAGGATGGAGGCACTCATCACCGCCCTGGAGGAACAGGTGCTGGGCTGCTGGAGGGGGC
          : ::::::::::  :::::: :  :::: ::: :    ::::::::::::::::: :::::
KIAA0165  ACAGGATGGAGGTTCTCATCGCTTCCCTAGAGAAGTCTGTGCTGGGCTGCTGGAAGGGGC
                5500       5510       5520       5530       5540       5550

                5480       5490       5500       5510       5520       5530
      K7  TGCTGCTGCCATGCAGTGCTGACCCCAGCCTTGCTCAGGAGGCCTCCAAACTACAAGAGC
          :::::::::: : :::::  :: ::: ::: ::: ::::::::::::   ::::: ::::
KIAA0165  TGCTGCTGCCGTCCAGTGAGGAGCCCGGCCCTGCCCAGGAGGCCTCCCGCCTACAGGAGC
                5560       5570       5580       5590       5600       5610

                5540       5550       5560       5570       5580       5590
      K7  TGCTGCGGGAATGTGGCTGGGAATATCCTGACTCCACTCTCCTGAAAGTCATCCTTAGTG
          :::: : ::: ::::::::: ::::::::::  ::::::: ::::::: :::: :: ::::
KIAA0165  TGCTACAGGACTGTGGCTGGAAATATCCTGACCGCACTCTGCTGAAAATCATGCTCAGTG
                5620       5630       5640       5650       5660       5670

                5600       5610       5620       5630       5640       5650
      K7  GGGCCAGAATCCTCACCAGCCAGGATGTTCAGGCCTTGGCGTGTGGGCTGTGCCCAGCCC
          : ::: :   :::::::   :::::  :::::::: :::: :  :::::::::::: :::
KIAA0165  GTGCCGGTGCCCTCACCCCTCAGGACATTCAGGCCCTGGCCTACGGGCTGTGCCCAACCC
                5680       5690       5700       5710       5720       5730

                5660       5670       5680       5690       5700       5710
      K7  AGCCTGACCGAGCCCAAGTGCTCCTGAGCGAGGCAGTAGGACGGGTTCAGAGCCAGGAGG
          :::: :: :::::::: : ::::::::  :::::::::::::::  : ::: ::: :   :
KIAA0165  AGCCAGAGCGAGCCCAGGAGCTCCTGAATGAGGCAGTAGGACGTCTACAGGGCCTGACAG
                5740       5750       5760       5770       5780       5790

                5720       5730       5740       5750       5760       5770
      K7  CCCCTCGAAGTCAACATCTTGTATTGGTGTTAGACAAGGACCTGCAAAAGCTGCCTTGGG
            ::  : : :   :: ::::: :::::  :::::::::::: :::: :::::::: ::::
KIAA0165  TACCAAGCAATAGCCACCTTGTCTTGGTCCTAGACAAGGACTTGCAGAAGCTGCCGTGGG
                5800       5810       5820       5830       5840       5850

                5780       5790       5800       5810       5820       5830
      K7  AAAGCACGCCCATCCTCCGAGCACAGCCGGTCACCCGGCTGCCCTCCTTCCGCTTCCTGC
          :::::: ::::: ::::: ::::: ::: ::::::::::::::::::::::::::::: :
KIAA0165  AAAGCATGCCCAGCCTCCAAGCACTGCCTGTCACCCGGCTGCCCTCCTTCCGCTTCCTAC
                5860       5870       5880       5890       5900       5910

                5840       5850       5860       5870       5880       5890
      K7  TCAGCTACACAGTCACTAAAGAGGCTGGAGCCTCATCAGTGCTGAGCCAAGGTGTTGATC
          :::::::: :  :::  :::::::  ::: :::::  :::::::::: ::::: :: ::::
KIAA0165  TCAGCTACTCCATCATCAAAGAGTATGGGGCCTCGCCAGTGCTGAGTCAAGGGGTGGATC
                5920       5930       5940       5950       5960       5970

                5900       5910       5920       5930       5940       5950
      K7  CACAGAATACCTTCTATGTGCTAAACCCTCACAGTAACCTGTCGAGCACAGAGGAGAGAT
          :::  : :::::::::::: :: :::::::::: ::::::::: ::::::::::::  ::
KIAA0165  CACGAAGTACCTTCTATGTCCTGAACCCTCACAATAACCTGTCAAGCACAGAGGAGCAAT
                5980       5990       6000       6010       6020       6030

                5960       5970       5980       5990       6000       6010
      K7  TTCGAGCCAGTTTTAGCAGTGAAACTGGCTGGAAAGGAGTGATTGGGGAAGTGCCAAGCC
          ::::::::: ::: ::::::::: ::::::::: ::::::: ::::::: :::::::: :
KIAA0165  TTCGAGCCAATTTCAGCAGTGAAGCTGGCTGGAGAGGAGTGGTTGGGGAGGTGCCAAGAC
                6040       6050       6060       6070       6080       6090

                6020       6030       6040       6050       6060       6070
      K7  TCGATCAGGTGCAGGCTGCCCTGACAGAGCGGGATTTATATATCTACGCAGGGCACGGAG
            :: ::::::::::  :::::::: :::  ::::: :::::::: :::::::: :: :
KIAA0165  CTGAACAGGTGCAGGAAGCCCTGACAAAGCATGATTTGTATATCTATGCAGGGCATGGGG
                6100       6110       6120       6130       6140       6150

                6080       6090       6100       6110       6120       6130
      K7  CCGGTGCCCGCTTCCTTGATGGGCAGGCTGTCCTGCGGCTGAGCTGCCGGGCGGTGGCCC
          : ::::::::::::::::::::::::::::::::::::::::::: ::::: :::::::
KIAA0165  CTGGTGCCCGCTTCCTTGATGGGCAGGCTGTCCTGCGGCTGAGCTGTCGGGCAGTGGCCC
                6160       6170       6180       6190       6200       6210
```

```
          6140       6150       6160       6170       6180       6190
       K7 TGCTGTTCGGCTGCAGCAGTGCAGCCCTTGCCGTGCATGGAAACCTGGAGGGAGCTGGTA
          ::::::: ::::: :::::::: ::::: :: :::::::::::::::::::: ::::: :
KIAA0165  TGCTGTTTGGCTGTAGCAGTGCGGCCCTGGCTGTGCATGGAAACCTGGAGGGGGCTGGCA
               6220       6230       6240       6250       6260       6270

               6200       6210       6220       6230       6240       6250
       K7 TCGTGCTCAAGTACATCATGGCTGGCTGCCCCTTGTTTCTAGGTAATCTCTGGGATGTGA
          :::::::::::::::::::::::::: :::::::::::::: ::::::::::::::::::::
KIAA0165  TCGTGCTCAAGTACATCATGGCTGGTTGCCCCTTGTTTCTGGGTAATCTCTGGGATGTGA
               6280       6290       6300       6310       6320       6330

               6260       6270       6280       6290       6300       6310
       K7 CTGACCGTGACATTGACCGTTACACAGAGGCTTTGCTGCAGGGCTGGCTGGGAGCAGGCC
          ::::::: :::::::::::: ::::: :: ::: ::::::: :::::::: ::::::::::
KIAA0165  CTGACCGCGACATTGACCGCTACACGGAAGCTCTGCTGCAAGGCTGGCTTGGAGCAGGCC
               6340       6350       6360       6370       6380       6390

               6320       6330       6340       6350       6360       6370
       K7 CAGGGGCTCCCTTTCTCTACTATGCCAGCCAAGCCCGCCAGGCCCCTCGACTCAAGTACC
          ::::::: ::: ::::::::::::  : ::: :::::::: :: :: :::::::::::: :
KIAA0165  CAGGGGCCCCCCTTCTCTACTATGTAAACCAGGCCCGCCAAGCTCCCCGACTCAAGTATC
               6400       6410       6420       6430       6440       6450

               6380       6390       6400       6410       6420
       K7 TCATTGGTGCTGCACCTGTAGCCTACGGCCTACCTATCTCTCTGCAGACACCC--TGAAG
          : ::::: ::::::::: :::::::: ::: : ::: ::::::::: :  ::::  :: ::
KIAA0165  TTATTGGGGCTGCACCTATAGCCTATGGCTTGCCTGTCTCTCTGCGGTAACCCCATGGAG
               6460       6470       6480       6490       6500       6510

          6430       6440       6450       6460            6470       6480
       K7 CTGTCTTGTCTATGGTAGAAGCTGTGTGACT-----ACCTCCAAACTTGGTTTTGATTAT
          :::::::: :  ::: ::::::::    : :::     ::::::::  :: : ::: : :
KIAA0165  CTGTCTTATTGATGCTAGAAGCCTCATAACTGTTCTACCTCCAAGGTTAGATTT-AATCC
               6520       6530       6540       6550       6560       6570

            6490            6500       6510       6520       6530
       K7 TTAGAAT-----TTTTAAAGTGATTTTCCCTGGTGTTTTGTATGAAATATTTTCTTTA-A
          :::: ::     :::::::::::::::::  ::::::: :::::::: :::: :::: :
KIAA0165  TTAGGATAACTCTTTTAAAGTGATTTTCCCCAGTGTTTTATATGAAACATTTCCTTTTGA
               6580       6590       6600       6610       6620       6630

          6540            6550       6560
       K7 TGTGACCT-----TAATAAAGATAT-TCCAGG
          : : ::::     ::::::::::  ::
KIAA0165  TTTAACCTCAGTATAATAAAGATACATCATTT
               6640       6650       6660
```

Figure 4.6: Alignment of the predicted K7 and KIAA0165 protein sequences

Alignment of the deduced K7 amino acid sequence with the deduced KIAA0165 amino acid sequence, following removal of the 81 bp insertion, and conceptual translation of the resulting *KIAA0165* DNA sequence. The comparisons were carried out using the CLUSTALW and Boxshade programs (2.3.33). Amino acid positions are indicated to the left, identical residues are indicated by black shading and similar residues, as defined by default parameters of Thompson *et al.* (1994), are indicated by grey shading.


K7 1 MRNFKGVNFATLLCSKEETQQLLPDLKEFLSRSRTDFPSSRTDAERRQICDTILRACTQQ
KIAA0165 1 MRSFKRVNFGTLLSSQKEAEELLPDLKEFLSNPPAGFPSSRSDAERRQACDAILRACNQQ
K7 61 LTAKLDCPGHLRSILDLAELACDGYLLSTPQRPPLYLERILFILLRNGSTQGSPDTVLRL
KIAA0165 61 LTAKLACPRHLGSLLELAELACDGYLVSTPQRPPLYLERILFVLLRNAAAQGGSPEATLRL
K7 121 AQPLHACLVQNSGEAAPQDYEAVTRGSFSLFWKGAEALLERRAAFSTRLNALSFLVLLED
KIAA0165 121 AQPLHACLVQCSREAAPQDYEAVARGSFSLLWKGAEALLERRAAFAARLKALSFLVLLED
K7 181 GSVPCEVPHFASPTACRLVAAYQLYDATGQGLDEADADFLYEVLSRHLIRVLVGEGGSSP
KIAA0165 181 ESTPCEVPHFASPTACRAVAAHQLFDASGHGLNEADADFLDDLLSRHVIRALVGERGSSS
K7 241 GPLSPQRALCLLEITLEHCRRLCWNHHHRQAARAVERARNHLEKTSVAPSLQLCQMGVEL
KIAA0165 241 GLLSPQRALCLLELTLEHCRRFCWSAHHDKAISAVEKAHSYLRNTNLAPSLQLCQLGVKL
K7 301 LEAVEERPGAVAQLLRKAAAVLINSIKAPSPPLRALYDSCQFFLSGLERGIRRHCGLDAI
KIAA0165 301 LQVGEEGPQAVAKLLIKASAVLSKSMEAPSPPLRALYESCQFFLSGLERGTKRRYRLDAI
K7 361 LSLFAFLGGYSSLVRHLR--EVSEASSKQQQCLLQMHFQGFHLFTGIVYDFAQGCQAT--
KIAA0165 361 LSLFAFLGGYCSLLQQLRDDGVYGGSSKQQQSFLQMYFQGLHLYTVVVYDFAQGCQIVDL
K7 417 -ELAQLVDGCRSAAVWMLEALEGLSGGELADYLSMTASYTSNLAYSFFSQKLYEEACVIS
KIAA0165 421 ADLTQLVDSCKSTVVWMLEALEGLSGQELTDHMGMTASYTSNLAYSFYSHKLYAEACAIS
K7 476 EPVCQHLGSATLGACPEVPPEKLHRCFRLHVESLKKLGKQAQGCKMVTLWLAALKPYSLE
KIAA0165 481 EPLCQHLGLVKPGTYPEVPPEKLHRCFRLQVESLKKLGKQAQGCKMVILWLAALQPCSPE
K7 536 HMVEPVTFWVRVKMDASRAGDKELQLQTLRDSLSCWDPETQSLLLREELRAYKAVRADTG
KIAA0165 541 HMAEPVTFWVRVKMDAARAGDKELQLKTLRDSLSGWDPETLALLLREELQAYKAVRADTG
K7 596 QERFNIICDLLELSPEETAAGAWARATYLVELAQVLCYHNFTQQTNCSALDAVQEALQLL
KIAA0165 601 QERFNIICDLLELSPEETPAGAWARATHLVELAQVLCYHDFTQQTNCSALDAIREALQLL
K7 656 DSVSPEAQEQDRLLDDKAQALLWLYICTLEAKMQEGIERDRRAQAPSNLEEFEVNDLNYE
KIAA0165 661 DSVRPEAQARDQLLDDKAQALLWLYICTLEAKIQEGIERDRRAQAPGNLEEFEVNDLNYE
K7 716 DKLQEDRFLYSSIAFNLAADAAQSKCLDQALTLWKEVLTKGRAPAVRCLQQTAASLQILA
KIAA0165 721 DKLQEDRFLYSNIAFNLAADAAQSKCLDQALALWKELLTKGQAPAVRCLQQTAASLQILA
K7 776 AVYQLVAKPLQALETLLLLQIVSKRLQDHAKAASSSCQLTQLLLNLGCPSYAQLYLEEAE
KIAA0165 781 ALYQLVAKPMQALEVLLLLRIVSERLKDHSKAAGSSCHITQLLLTLGCPSYAQLHLEEAA
K7 836 SSLESLDQTSDACQLLSLTCALLGSQLCWACQKVTAGVSLLLSVLRDPALQKSSKAWYLL
KIAA0165 841 SSLKHLDQTTDTYLLLSLTCDLLRSQLYWTHQKVTKGVSLLLSVLRDPALQKSSKAWYLL
K7 896 RVQALQVLAFYLSLSSNLLSSALREQLWDQGWQTPETALIDAHKLLRSIIILLMGSDVLS
KIAA0165 901 RVQVLQLVAAYLSLPSNNLSHSLWEQLCAQGWQTPEIALIDSHKLLRSIILLLMGSDILS
K7 956 IQKAATESPFLDYGENLVQKWQVLTEVLTCSERLVGRLGRLGNVSEAKAFCLEALKLTTK
KIAA0165 961 TQKAAVETSFLDYGENLVQKWQVLSEVLSCSEKLVCHLGRLGSVSEAKAFCLEALKLTTK
K7 1016 LQIPRQCALFLVLKGELELARGDIDLCQSDLQQVLFLLESSTEFGVVTQHPDSVKKVHTQ
KIAA0165 1021 LQIPRQCALFLVLKGELELARNDIDLCQSDLQQVLFLLESCTEFGGVTQHLDSVKKVHLQ
K7 1076 KGKHKAQGPCFPPLSEEEPFLIGPALELVDTVLNEPGPIQSSVNSSPVLKTKPPPNPGFL
KIAA0165 1081 KGKQQAQVPCPPQLPEEELFLIGPALELVATVAKEPGPIAPSTNSSPVLKTKPQPIPNFL


K7 1136 SHLPSCDCLLCASPALSAVCLRWVLVTAGVRLATGHKAQGLDLLQAVLTRCPAATKRFTQ
KIAA0165 1141 SHSPTCDCSLCASPVLTAVCLRWVLVTAGVRLAMGHQAQGLDLLQVVLKGCPEAAERLTQ
K7 1196 SLQASLNHRTTPSCVPSLFDEIMAQVYTHLALEFLNQTSEKSLGKVLASGLKFVATRIQS
KIAA0165 1201 ALQASLNHKTPPSLVPSLLDEILAQAYTLLALEGLNQPSNESLQKVLQSGLKFVAARIPH
K7 1256 LEIWRAHLLLVQALAKLAHFSCCTSELFASSWGWHPPLVKSLPVLEPAKIRRQKCSGRGR
KIAA0165 1261 LEPWRASLLLVWALTKLGGLSCCTTQLFASSWGWQPPLIKSVPGSEPSKTQGQKRSGRGR
K7 1316 RRIASVPPPLHNSSQKGLEEEGPPCTPKPPGRARQAGPRVPFTIFEEVHPTKSKLQVPLA
KIAA0165 1321 QKLASAPLSLNNTSQKGLEGRGLPCTPKPPDRIRQAGPHVPFTVFEEVCPTESKPEVPQA
K7 1376 PRVHRRAQTRLKVIFSDDSDLEDLVSADTQLVEEPKRRGTASRTRGQTRKGRSLKTDAVV
KIAA0165 1381 PRVQQRVQTRLKVNFSDDSDLEDPVSAEAWLAEEPKRRGTASRGRGRARKGLSLKTDAVV
K7 1436 AIESTPGHSSVSGRTRRARKVASRNCEEESPKAPLCVWASQGPEIMRSIPEEEPVDNHLE
KIAA0165 1441 APGSAPGNPGLNGRSRRAKKVASRHCEERRPQR-ASDQARPGPEIMRTIPEEELTDNWRK
K7 1496 KSFEILRGSDGEDSASGEKAAAAADTGLPVGECEVLRRDSSKAERPVLYSDTEANSDPSPW
KIAA0165 1500 MSFEILRGSDGEDSASGGKTPAPGPEAASGEWELLRLDSSKKKLPSPCPDKESDKDLGPR
K7 1556 LPPFSVPAPIDLSTLDSISDSLSIAFRGVSHCPPSGLYAHLCRFLALCLGHRDPYATAFL
KIAA0165 1560 LQLPSAPVATGLSTLDSICDSLSVAFRGISHCPPSGLYAHLCRFLALCLGHRDPYATAFL
K7 1616 VAESISITCRHQLLTHLHRQLSKAQKQQESPELAEHLQRLDLKERPRGVPLARIQRLFSF
KIAA0165 1620 VTESVSITCRHQLLTHLHRQLSKAQKHRGSLEIADQLQGLSLQEMPGDVPLARIQRLFSF
K7 1676 KALGSGCFPQAEKESFQERLALIPSGVTVCVLALATLQPGTLSNTLLLTRLEKDNPPITV
KIAA0165 1680 RALESGHFPQPEKESFQERLALIPSGVTVCVLALATLQPGTVGNTLLLTRLEKDSPPVSV
K7 1736 KIPTAQNKLPLSAVLKEFDAIQKDQKENSSCTEKRVWWTGRLALDQRMEALITALEEQVL
KIAA0165 1740 QIPTGQNKLHLRSVLNEFDAIQKAQKENSSCTDKREWWTGRLALDHRMEVLIASLEKSVL
K7 1796 GCWRGLLLPCSADPSLAQEASKLQELLRECGWEYPDSTLLKVILSGARILTSQDVQALAC
KIAA0165 1800 GCWKGLLLPSSEEPGPAQEASRLQELLQDCGWKYPDRTLLKIMLSGAGALTPQDIQALAY
K7 1856 GLCPAQPDRAQVLLSEAVGRVQSQEAPRSQHLVLVLDKDLQKLPWESTPILRAQPVTRLP
KIAA0165 1860 GLCPTQPERAQELLNEAVGRLQGLTVPSNSHLVLVLDKDLQKLPWESMPSLQALPVTRLP
K7 1916 SFRFLLSYTVTKEAGASSVLSQGVDPQNTFYVLNPHSNLSSTEERFRASFSSETGWKGVI
KIAA0165 1920 SFRFLLSYSIIKEYGASPVLSQGVDPRSTFYVLNPHNNLSSTEEQFRANFSSEAGWRGVV
K7 1976 GEVPSLDQVQAALTERDLYIYAGHGAGARFLDGQAVLRLSCRAVALLFGCSSAALAVHGN
KIAA0165 1980 GEVPRPEQVQEALTKHDLYIYAGHGAGARFLDGQAVLRLSCRAVALLFGCSSAALAVHGN
K7 2036 LEGAGIVLKYIMAGCPLFLGNLWDVTDRDIDRYTEALLQGWLGAGPGAPFLYYASQARQA
KIAA0165 2040 LEGAGIVLKYIMAGCPLFLGNLWDVTDRDIDRYTEALLQGWLGAGPGAPLLYYVNQARQA
K7 2096 PRLKYLIGAAPVAYGLPISLQTP
KIAA0165 2100 PRLKYLIGAAPIAYGLPVSLR--

similar, as defined by the default parameters of Tatusova and Madden (1999), with homology extending across the entire protein sequence.

The K7 and KIAA0165 protein sequences also displayed substantial homology in a C-terminal region to gene products from three fungal organisms; *ESP1* from *Saccharomyces cerevisiae* (Baum *et al.*, 1988; McGrew *et al.*, 1992), *cut1* from *Schizosaccharomyces pombe* (Uzawa *et al.*, 1990), and *bimB* from *Aspergillus nidulans* (May *et al.*, 1992) (Figure 4.2; 4.7 and 4.8). Temperature sensitive mutation of these genes result in similar phenotypes. Mitosis and DNA replication become uncoupled such that multiple rounds of DNA replication occur in the absence of correct chromosome segregation, resulting in polyploid DNA contents and the production of multiple and aberrant spindle pole bodies, the yeast equivalent of the mammalian centrosomes. K7 displayed 87% identity and 94% similarity to KIAA0165, 42% identity and 56% similarity to *bimB*, 38% identity and 52% similarity to *cut1*, and 33.5% identity and 50% similarity to *ESP1*, in the C-terminal region (Figure 4.8). This C-terminal region also encompassed the only significant homology between the three fungal genes. *BimB* shares 50% identity and 60% similarity to *Cut1* and 40% identity and 58% similarity to *ESP1*, while Cut1 shares 43% identity and 56% similarity to *ESP1*. This implied that the C-terminal region is required for a conserved function of these genes during mitosis, and suggests a putative role for *K7* in this process in mammals.

4.3 CHROMOSOME LOCALISATION OF K7

The chromosome localisation of *K7* was determined using two *K7* cDNA probes, an 895 bp *Hind*III fragment corresponding to nucleotides 3663-4558 of the *K7* cDNA, and a 2012 bp *Hind*III/*Eco*RI fragment, corresponding to nucleotides 4558-6569 of the *K7* cDNA (Figure 4.1). Mapping of the physical location of the *K7* gene was carried out by Dr. Graham Webb (Department of Animal Science, University of Adelaide), using fluorescent *in situ* hybridisation (FISH) as described in Kamei *et al.* (2000). Slides of mitotic chromosomes with 5'-bromo-2'-deoxyuridine incorporated in early S-phase were prepared from mouse splenic lymphocytes for FISH analysis. The probes were biotinylated by nick translation, and hybridised to mouse chromosomes at a concentration of 10 ng/µl in the presence of 5 µg/µl of sonicated salmon sperm DNA. After FISH the chromosomes were banded by staining with propidium iodide by the

Figure 4.7: Schematic representation of protein sequence identity between K7, KIAA0165, bimB, Cut1 and Esp1

Diagrammatic depiction of regions of homology between K7 and KIAA0165, bimB, Cut1 and Esp1 proteins (grey shading). KIAA0165 displayed high sequence identity with K7 throughout the entire amino acid sequence. However, bimB, Cut1 and Esp1 displayed identity with K7 only within a carboxyl (C) terminal region of the protein.

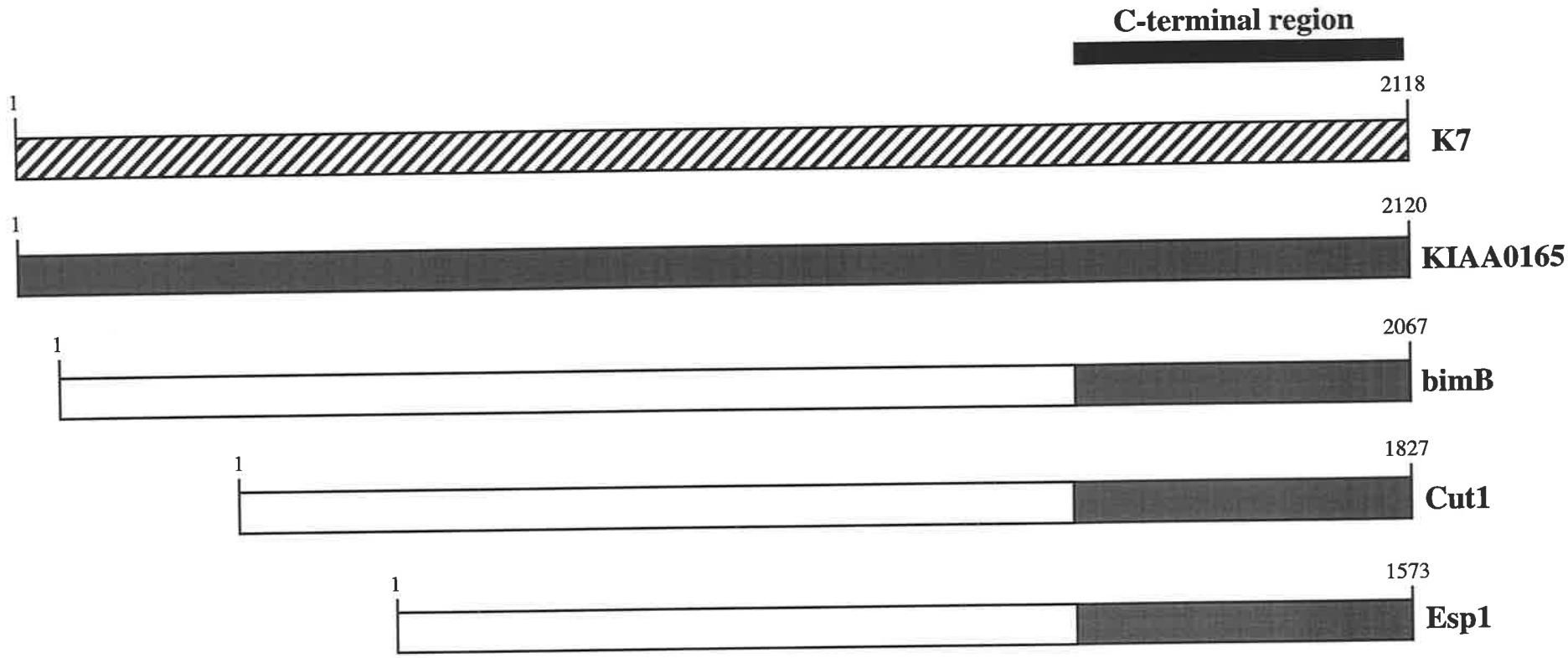
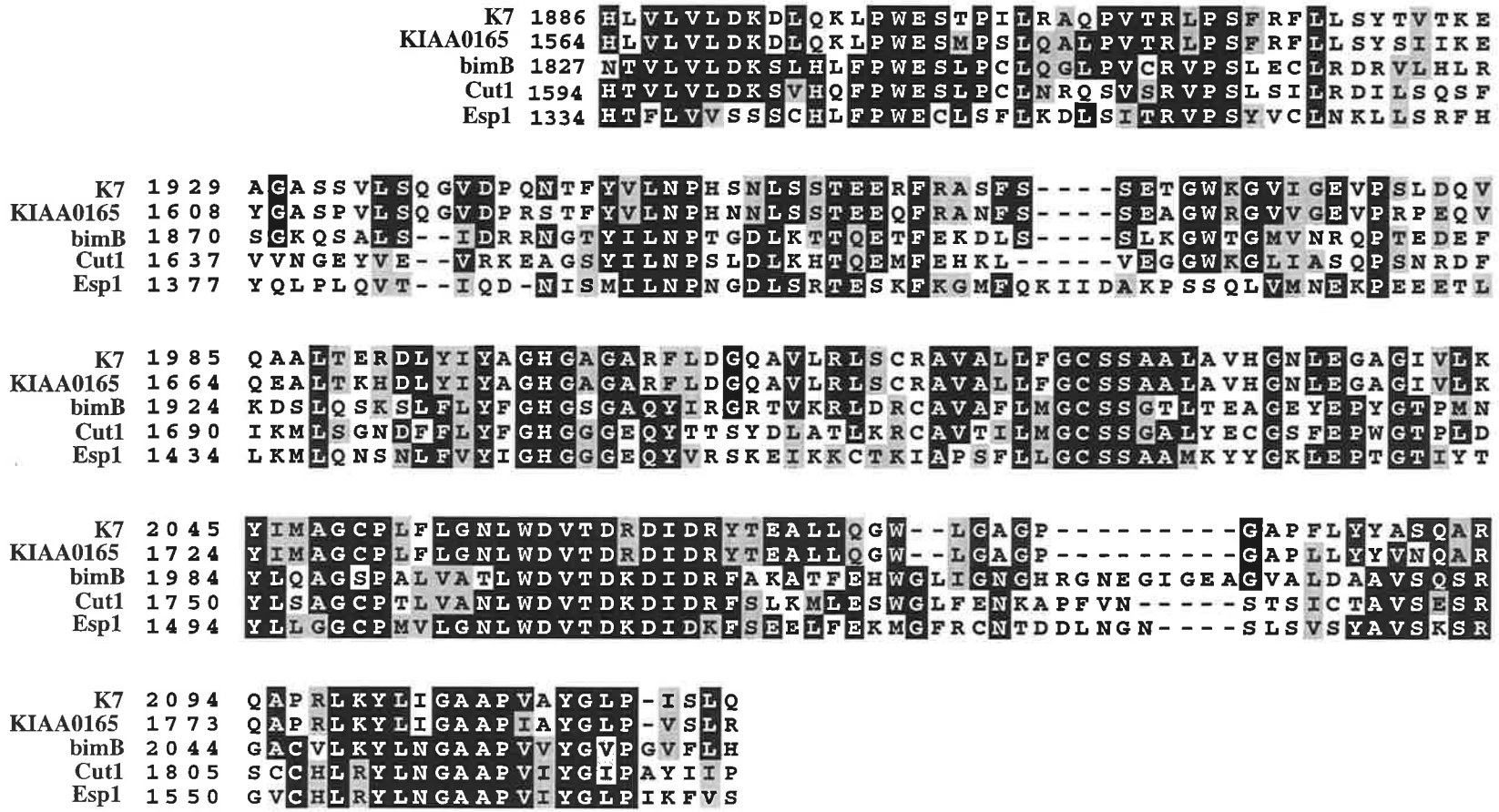

C-terminal region
1
2118
K7
1
2120
KIAA0165
1
2067
bimB
1
1827
Cut1
1
1573
Esp1

Figure 4.8: Alignment of the C-terminal regions of K7, KIAA0165, bimB, Cut1 and Esp1

Alignment of K7, KIAA0165, bimB, Cut1 and Esp1 amino acid sequences using the CLUSTALW and Boxshade programs (2.3.33). Amino acid positions are indicated to the left, identical residues are indicated by the black shading and similar residues, as defined by default parameters of Thompson *et al.* (1994) are indicated by the grey shading.


K7 1886 HLVLVLDKDLQKLPWESTPILRAQPVTRLPSFRFLLSYTVTKE
KIAA0165 1564 HLVLVLDKDLQKLPWESMPSLQALPVTRLPSFRFLLSYSIIKE
bimB 1827 NTVLVLDKSLHLFPWESLPCLQGLPVCRVPSLECLRDRVLHLR
Cut1 1594 HTVLVLDKSVHQFPWESLPCLNRQSVSRVPSLSILRDILSQSF
Esp1 1334 HTFLVVSSSCHLFPWECLSFLKDLSITRVPSYVCLNKLLSRFH
K7 1929 AGASSVLSQGVDPQNTFYVLNPHSNLSSTEERFRASFS----SETGWKGVIGEVPSLDQV
KIAA0165 1608 YGASPVLSQGVDPRSTFYVLNPHNNLSSTEEQFRANFS----SEAGWRGVVGEVPRPEQV
bimB 1870 SGKQSALS--IDRRNGTYILNPTGDLKTTQETFEKDLS----SLKGWTGMVNRQPTEDEF
Cut1 1637 VVNGEYVE--VRKEAGSYILNPSLDLKHTQEMFEHKL-----VEGGWKGLIASQPSNRDF
Esp1 1377 YQLPLQVT--IQD-NISMILNPNGDLSRTESKFKGMFQKIIDAKPSSQLVMNEKPEEETL
K7 1985 QAALTERDLYIYAGHGAGARFLDGQAVLRLSCRAVALLFGCSSAALAVHGNLEGAGIVLK
KIAA0165 1664 QEALTKHDLYIYAGHGAGARFLDGQAVLRLSCRAVALLFGCSSAALAVHGNLEGAGIVLK
bimB 1924 KDSLQSKSLFLYFGHGSGAQYIRGRTVKRLDRCAVAFLMGCSSGTLTEAGEYEPYGTPMN
Cut1 1690 IKMLSGNDFFLYFGHGGGEQYTTSYDLATLKRCAVTILMGCSSGALYECGSFEPWGTPLD
Esp1 1434 LKMLQNSNLFVYIGHGGGEQYVRSKEIKKCTKIAPSFLLGCSSAAMKYYGKLEPTGTIYT
K7 2045 YIMAGCPLFLGNLWDVTDRDIDRYTEALLQGW--LGAGP---------GAPFLYYASQAR
KIAA0165 1724 YIMAGCPLFLGNLWDVTDRDIDRYTEALLQGW--LGAGP---------GAPLLYYVNQAR
bimB 1984 YLQAGSPALVATLWDVTDKDIDRFAKATFEHWGLIGNGHRGNEGIGEAGVALDAAVSQSR
Cut1 1750 YLSAGCPTLVANLWDVTDKDIDRFSLKMLESWGLFENKAPFVN-----STSICTAVSESR
Esp1 1494 YLLGGCPMVLGNLWDVTDKDIDKFSEELFEKMGFRCNTDDLNGN----SLSVSYAVSKSR
K7 2094 QAPRLKYLIGAAPVAYGLP-ISLQ
KIAA0165 1773 QAPRLKYLIGAAPIAYGLP-VSLR
bimB 2044 GACVLKYLNGAAPVVYGVPGVFLH
Cut1 1805 SCCHLRYLNGAAPVIYGIPAYIIP
Esp1 1550 GVCHLRYLNGAAPVIYGLPIKFVS

method of Lemieux *et al.* (1992). Cells showing fluorescent signal under the FITC wavelength were photographed and scored (Figure 4.9).

FITC signal grains on the banded chromosomes were scored initially onto a standard mouse ideogram (Nesbitt and Francke, 1973) (Figure 4.10). The plots generated by each probe were very similar, and a total of 79 and 57 FITC grains for the 2012 bp and 895 bp *K7* probes respectively were scored over the whole karyotype. Of these, 63 grains in total (46% of all grains) were located over the very distal C3.3-C4 bands of Chromosome 16 (Figure 4.10). Of the 73 signals which did not localise to distal chromosome 16, 38.4% were over the centric heterochromatin of different chromosomes (Figure 4.10). This was interpreted as evidence that both probes contained a short length of sequence homologous to a satellite DNA sequence, since these sequences are found in regions of heterochromatin. There was no notable sub-peak of background grains over euchromatic regions of any other chromosome (Figure 4.10).

To confirm this mapping, FITC signal grains from both probes were replotted onto a detailed ideogram (Evans, 1989) of mouse Chromosome 16. Of a total of 74 FITC grains, 54 (73%) formed the two tallest peaks lying over the very distal bands 16C3.3-C4 (Figure 4.11 A), confirming these bands as the probable location of the mouse *K7* gene. Consistent with this, FITC signals were apparent on both chromatids of labelled Chromosome 16 in 13.6% of samples probed with the 895 bp probe (Figure 4.9 A) and 31.3% of samples probed with the longer 2012 bp *K7* probe (Figure 4.9 B). The tallest peak of fluorescent signal, 33 FITC grains (44.5%), resided over band 16C4, suggesting that this band is the possible point location of *K7* locus.

Alignment of the physical location of the *K7* locus with the corresponding physical and linkage maps for mouse the distal region of mouse Chromosome 16 (Figure 4.11 B; Cabin and Reeves, 1998) demonstrated that the only gene which had been localised by *in situ* hybridisation was the *sim2* gene, which also mapped to 16C3.3-C4 (Ema *et al.*, 1996).

The only mutation that mapped in the vicinity of 16C3.3-C4 was *weaver* (*wv*), a mutation which arose spontaneously in the C57BL/6J strain (Lane, 1964). The *wv* mutation results in degeneration of a subset of neuronal cell populations, including the granule cells of the cerebellum and dopaminergic neurons in the substantia nigra resulting in neurological and neuromuscular defects (reviewed by Hess, 1996). However, the mutation was mapped to a missense point mutation in the *Girk2* gene (*Kcnj6* locus; Figure 4.11 B), a gene that encodes a

Figure 4.9: Fluorescent *in situ* hybridisation of *K7* to mouse chromosome spreads

Fluorescent *in situ* hybridisation of mouse mitotic chromosome using two alternate *K7* probes; an 895 bp *Hind*III fragment (**A**) and a 2012 bp *Hind*III/*Eco*RI fragment (**B**). Hybridisation sites at the distal tip of chromosome 16 are indicated by arrows. Size bar indicates 10 μm.

A

B

Figure 4.10: Scoring plot of fluorescent signals observed on mitotic chromosomes probed with *K7* specific probes

Standard ideogram of the whole mouse karyotype (Nesbitt and Francke, 1973), scored for FITC signals following hybridisation with two *K7* probes. The 895 bp *Hind*III probe fluorescent signal is depicted using open circles **(A)** and encompassed 57 grains, while the 2012 bp *Hind*III/*Eco*RI probe florescent signal encompassed 79 grains and is depicted using closed circles **(B)**.


A
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
18
19
X
Y
16
B
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
18
19
X
Y
16

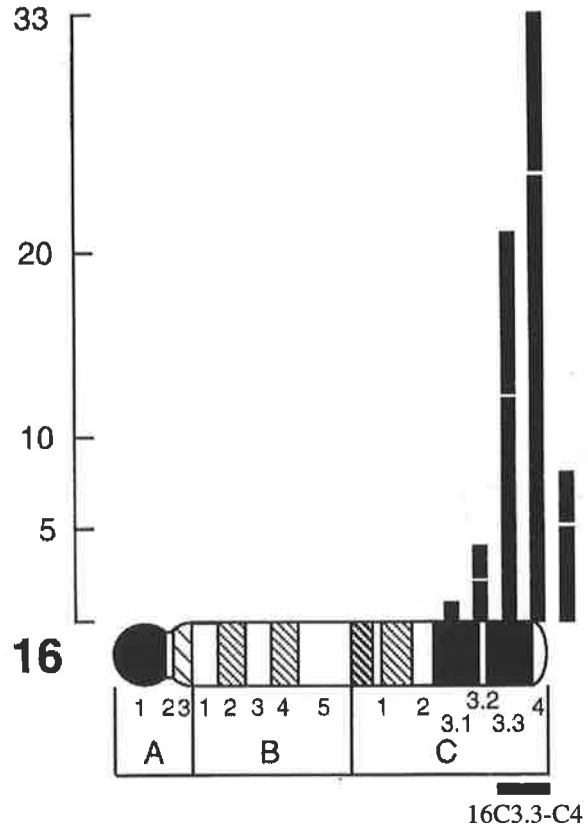
Figure 4.11: Localisation of the *K7* locus on mouse chromosome 16

(**A**) Schematic representation of *K7* chromosome localisation to mouse chromosome 16 as determined from the fluorescent signals from two *K7* specific probes. The lower and upper portion of each column refer to signals from the 2012 bp and 895 bp probes respectively. Localisation to bands 16C3.3-C4 is regarded as most probable.

(**B**) Linkage map of the distal region of mouse chromosome 16 (adapted from Cabin and Reeves, 1998). The linkage map positions are indicated in centimorgans. Genes physically mapped by *in situ* hybridisation are indicated with a star and underlined. Regions of synteny with human chromosome 21 are also indicated. Mapped DNA segments were omitted from this analysis.

Gene abbreviations: Sod1 = superoxide dismutase-1; Crbf4 = cytokine receptor family 2, member 4; Tiam1 = T-cell lymphoma invasion and metastasis 1; Cbfa2 = core binding factor 2; Gart = phosphoribosylglycinamide formyltransferase; Gas4 = growth arrest specific 4; Ifnar = interferon (alpha and beta;) receptor; Son = son cell proliferation protein; Kcne1 = potassium voltage-gated channel, Isk related subfamily; Infgr2 = interferon gamma receptor 2; Cbr = carbonyl reductase; Sim2 = single minded 2; Dyrk1a = dual-specificity tyrosine-(Y)-phosphorylation receptor kinase; Kcnj6 = potassium inwardly-rectifying channel; Erg = avian erythroblastosis virus E-26 (v-ets) oncogene; Ets2 = E26 avian leukaemia oncogene-2; Hmg14 = high mobility group protein 14; Pcp4 = purkinje cell protein-4; Mx1 = myxovirus (influenza virus) resistance gene 1; Mx2 = myxovirus (influenza virus) resistance gene 2; Tel16q = telomeric sequence, chromosome 16, q arm.

A



B

Mouse Chromosome 16	Band	Human Homolog and Band
61 Sod1		SOD1; 21q22.1
Crfb4		CRFB4; 21q22.1-q22.2
Tiam1		TIAM1; 21q22.1
62 Cbfa2		CBFA2; 21q22.1-q22.2
63 Gart		GART; 21q22.1
64 Gas4		
Ifnar		
Son		SON; 21q22.1-q22.2
Kcne1		KCNE1; 21q22.1-q22.2
65 Ifngr		IFNGT1; 21q22.1
66 Cbr		CBR1; 21q22.3
67 Sim2*	16C3.3-C4	SIM2; 21q22.2
Dyrk1a		DYRK; 21q22.2
68 Kcnj6		KCNJ6; 21q22.2
Erg		ERG; 21q22.2
Ets2		ETS2; 21q22.3
69 Hmg14		HMG14; 21q22.3
Pcp4		PCP4; 21q22.3
Mx1		MX1; 21q22.3
70 Mx2		MX2; 21q22.3
Tel16q		

G-protein-coupled inwardly rectifying K^+ channel (Patil *et al.*, 1995), which is important in neuronal cells to establish and maintain membrane excitability.

4.4 *K7* EXPRESSION IN FOETAL AND ADULT TISSUES

4.4.1 *K7* Expression During Gastrulation

K7 had previously been shown to be expressed during peri-implantation mouse development within pluripotent cells (3.2.3.3). To investigate if *K7* was expressed at the subsequent gastrulation stage of mouse development, whole mount *in situ* hybridisation (2.4.6) was carried out on embryos from 6.5-8.5 d.p.c. using sense and antisense DIG labelled *K7* riboprobes (2.4.5). *K7* expression was not detected at the onset of gastrulation (6.5 d.p.c.; Figure 4.12 A), during the late streak stage (7.5 d.p.c.; Figure 4.12 B) or during the early somite stage at approximately 8.5 d.p.c (Figure 4.12 C). Positive controls for the experiment (data not shown) included detection of pluripotent cell expression using an *Oct4* antisense riboprobe on 5.5 d.p.c. embryos and expression in progenitor cells of the developing heart using a *Nkx-2.5* antisense riboprobe on 8.5 d.p.c. embryos (Lints *et al.*, 1993). These results indicate that *K7* does not appear to be required for gastrulation or early organogenesis.

4.4.2 *K7* Expression During Later Mouse Embryogenesis

Analysis of *K7* expression during later stages of mouse embryogenesis was carried out using RNAse protection assays (2.3.21.2). RNAse protection analysis was performed using two riboprobes, an antisense riboprobe derived from a 370 bp *Hind*III/*Sty*I fragment of the *K7* cDNA (2.3.21.1; Figure 4.1), and a mouse *glyceraldehyde phosphate dehydrogenase* (*mGap,* 240 bp; 2.3.21.1) riboprobe as a loading control. *mGap* is expressed at relatively constant levels among different cell types, and probes were synthesised at lower specific activities to compensate for the greater expression of this gene compared to *K7* (2.3.21.2) (Rathjen *et al.*, 1990b). tRNA samples were used in the RNAse protection procedure to confirm the specificity of protected bands.

Post-gastrulation embryonic RNA samples from 10.5 d.p.c. to 17.5 d.p.c were screened for *K7* expression by RNAse protection (Figure 4.13). *K7* was barely detectable at 10.5 and 11.5 d.p.c., which was consistent with the absence of *K7* expression during early

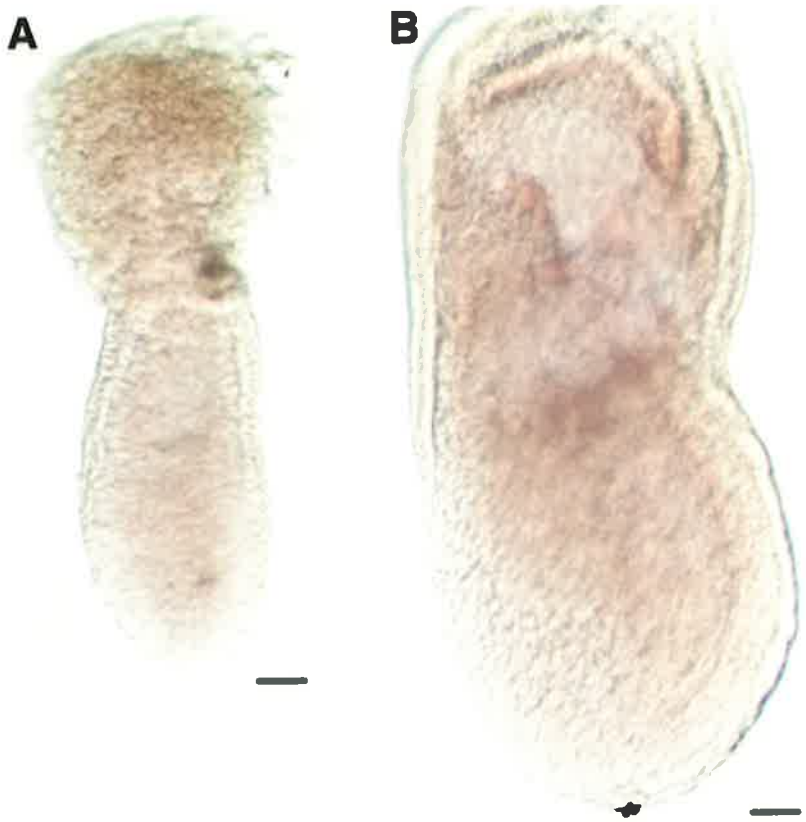
Figure 4.12: Expression analysis of *K7* in post-implantation embryos from 6.5 to 8.5 d.p.c. by whole mount *in situ* hybridisation

Analysis of *K7* expression in post-implantation embryos, from day 6.5 to day 8.5 of embryogenesis, was carried out by wholemount *in situ* hybridisation using a 232 bp DIG labelled antisense riboprobe (2.4.5/6). Embryo stains were developed for greater than 6 hours. Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the Hoffman modulation contrast system. The size bars indicate 100 µm.

A) 6.5 d.p.c. embryo

B) 7.5 d.p.c. embryo

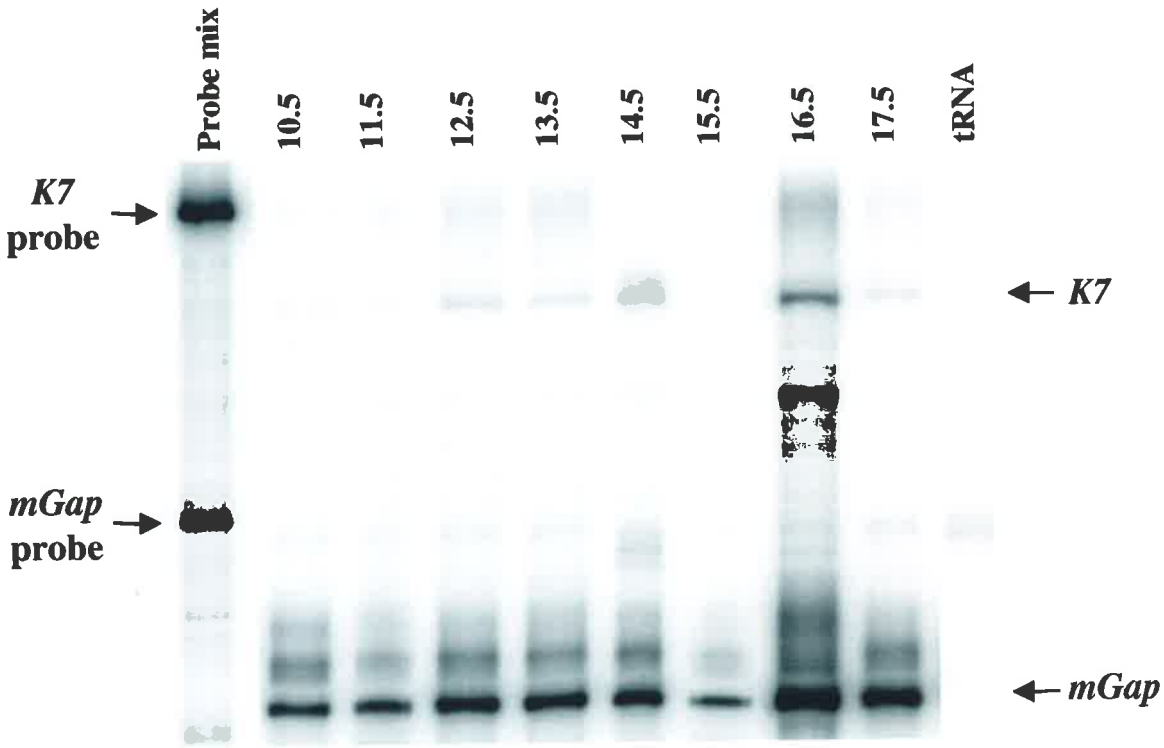
C) 8.5 d.p.c. embryo


A
B


C

Figure 4.13: Analysis of *K7* expression from 10.5 d.p.c. to 17.5 d.p.c. by RNAse Protection

K7 expression was detected using a 376 bp antisense *K7* riboprobe (2.3.21.1). RNAse protection analysis (2.3.21.2) was performed on 10 µg of total RNA isolated from mouse embryos at 10.5 d.p.c. to 17.5 d.p.c., at 1 day intervals, and 10 µg of tRNA. Mouse *Glyceraldehyde phosphate dehydrogenase* (*mGap*) was used as a loading control. Arrows on the left hand side indicate the size of *K7* and *mGap* unprotected probes, while arrows to the right of the figure indicate the protected bands for *K7* and *mGap* respectively.


Probe mix
10.5
11.5
12.5
13.5
14.5
15.5
16.5
17.5
tRNA
K7 probe
mGap probe
K7
mGap

organogenesis. Low level expression was observed at day 12.5 and day 13.5 of embryogenesis. *K7* expression was upregulated at 14.5 d.p.c., with highest level expression detected at 16.5 d.p.c., before downregulation at day 17.5 of embryogenesis. Therefore, *K7* was expressed in later stages of mouse embryogenesis in a developmentally regulated manner.

RNAse protection analysis of RNA isolated from various 16.5 d.p.c. embryonic tissues was carried out to determine the embryonic sites of *K7* expression in foetal tissues. *K7* was expressed in all tissues analysed by RNAse protection (Figure 4.14 A). Quantitation of *K7* expression levels compared to *mGap* expression levels by volume integration (2.3.25) indicated an approximate 4 fold variation in expression levels, with expression highest in the foetal liver, stomach, lung and kidney, intermediate in the intestine, skin and limb, and low in the foetal heart and brain (Figure 4.14 B). Therefore, *K7* was differentially expressed in a tissue specific manner at 16.5 d.p.c. of mouse embryogenesis.

4.4.3 *K7* Expression in Adult Tissues

Differential regulation of *K7* expression was also observed in adult tissues (Figure 4.15). In contrast to the expression pattern observed in foetal tissues, the majority of adult tissues analysed did not exhibit detectable *K7* expression, indicating that the *K7* gene is more tightly regulated within the adult. For example, *K7* was detectable at high levels in the kidney and liver at 16.5 d.p.c., but expression was not detectable in these tissues in the adult. Highest level expression was detected in adult testis, with expression also detected in haematopoietic tissues of the bone marrow, spleen, and thymus, but not mesenteric lymph nodes, and also in the stomach. *K7* expression in the stomach and at the site of haematopoiesis, which occurs in the foetal liver at 16.5 d.p.c. and in the bone marrow of adults, was therefore conserved between embryo and adult.

Since *K7* expression was detected in tissues of the haematopoietic system, mouse haematopoietic cell lines, representing different haematopoietic lineages, were screened by RNAse protection to determine if *K7* was expressed by all lineages, or if differential regulation was present within haematopoietic cell lineages. RNA derived from cell lines representing both lymphoid and myeloid precursors (Figure 4.16), including a lymphoblast-like cell line (P388D1), a B-cell line (Baf-3), a T-cell line (CTLL), promyelocyte-like cell lines (WEHI-3B and FDC-P1),

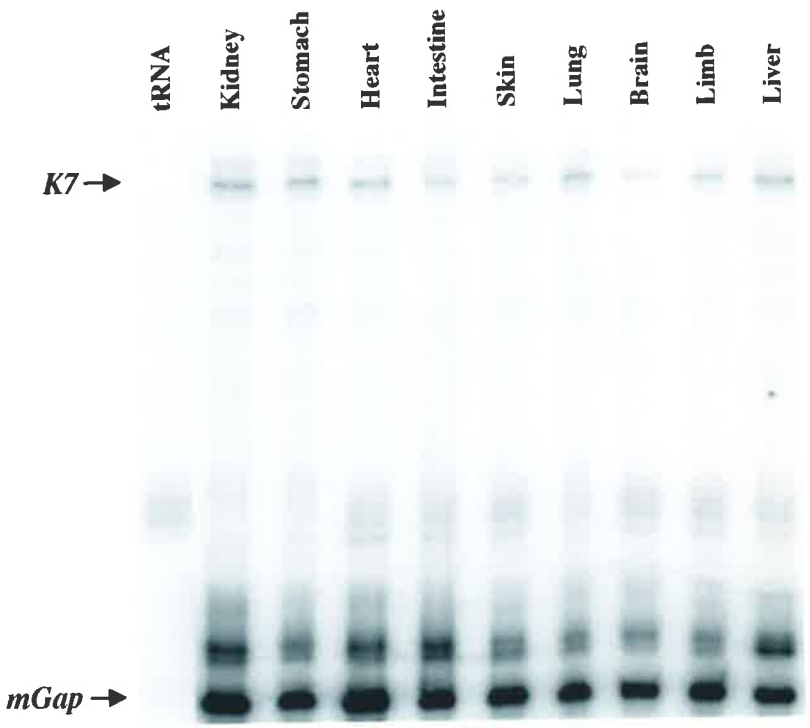
Figure 4.14: Expression analysis of *K7* in 16.5 d.p.c. embryonic tissues by RNAse Protection

A) *K7* expression was detected using a 376 bp antisense *K7* riboprobe (2.3.21.1). RNAse protection (2.3.21.2) was performed on 10 µg of total RNA isolated from embryonic tissues of 16.5 d.p.c. mouse embryos, and 10 µg of tRNA. Mouse *Glyceraldehyde phosphate dehydrogenase* (*mGap*) was used as a loading control.

B) Graphical representation of *K7* expression levels, expressed as a percentage of *mGap* expression, following quantitation by volume integration (2.3.25).

Ki = kidney, St = stomach, He = heart, Int = intestine, Sk = skin, Lu = lung, Br = brain, Liv = liver

A


tRNA
Kidney
Stomach
Heart
Intestine
Skin
Lung
Brain
Limb
Liver
K7
mGap

B

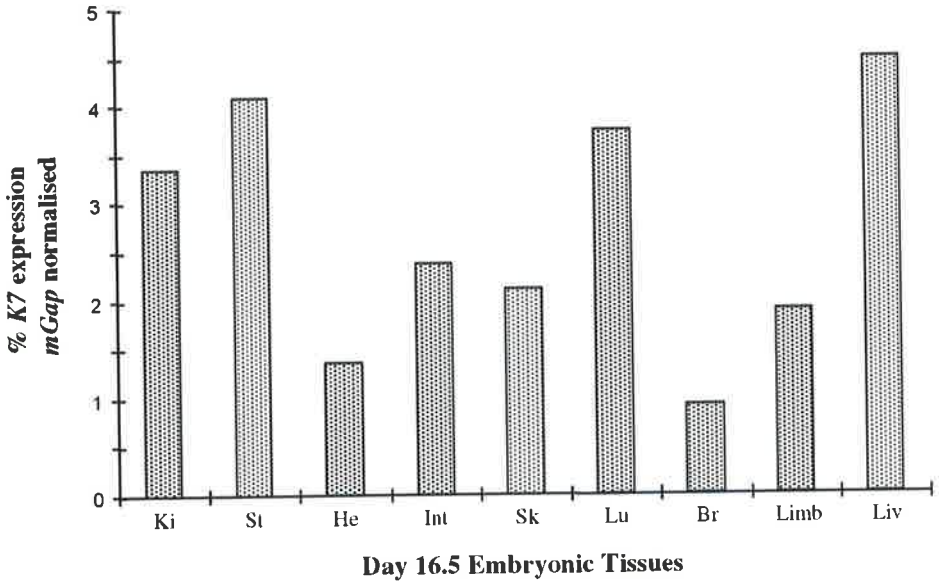

% K7 expression
mGap normalised
0
1
2
3
4
5
Ki
St
He
Int
Sk
Lu
Br
Limb
Liv
Day 16.5 Embryonic Tissues

Figure 4.15: Expression analysis of *K7* in adult tissues by RNAse Protection

K7 expression was detected using a 376 bp antisense *K7* riboprobe (2.3.21.1). RNAse protection (2.3.21.2) was performed on 10 µg of total RNA isolated from adult mouse tissues, and 10 µg of tRNA. Mouse *Glyceraldehyde phosphate dehydrogenase* (*mGap*) was used as a loading control. The separate panel represents a second RNAse protection assay of 10 µg of RNA isolated from adult thymus and 10 µg of tRNA.

Te = testis, Ov = ovary, Pl = placenta, Br = brain, He = heart, Li = liver, Mu = muscle, St = stomach, Kid = kidney, Sp = spleen, BM = bone marrow, MLN = mesenteric lymph nodes, Th = thymus, t = tRNA

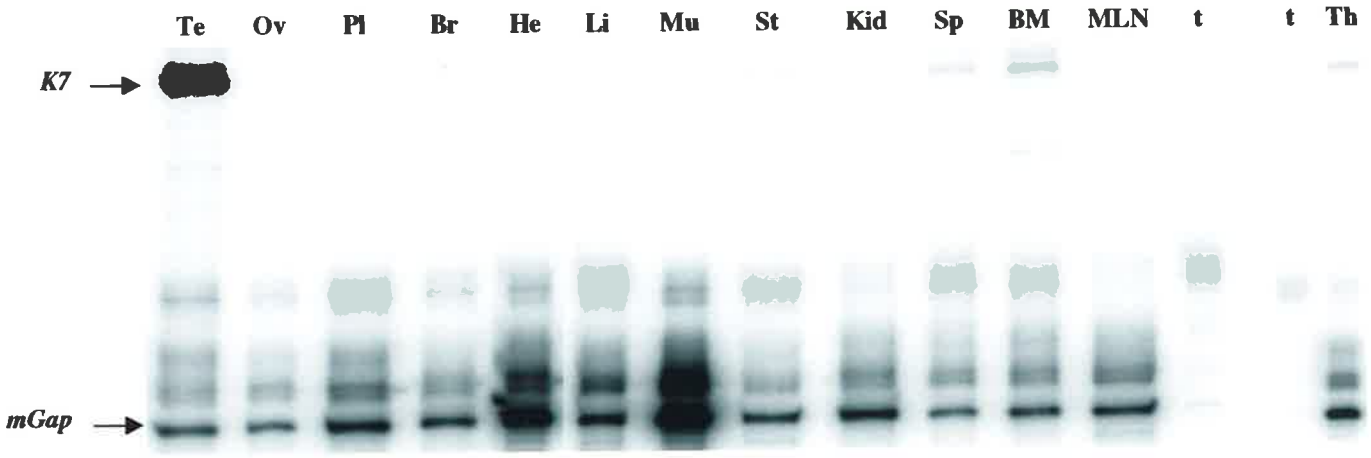
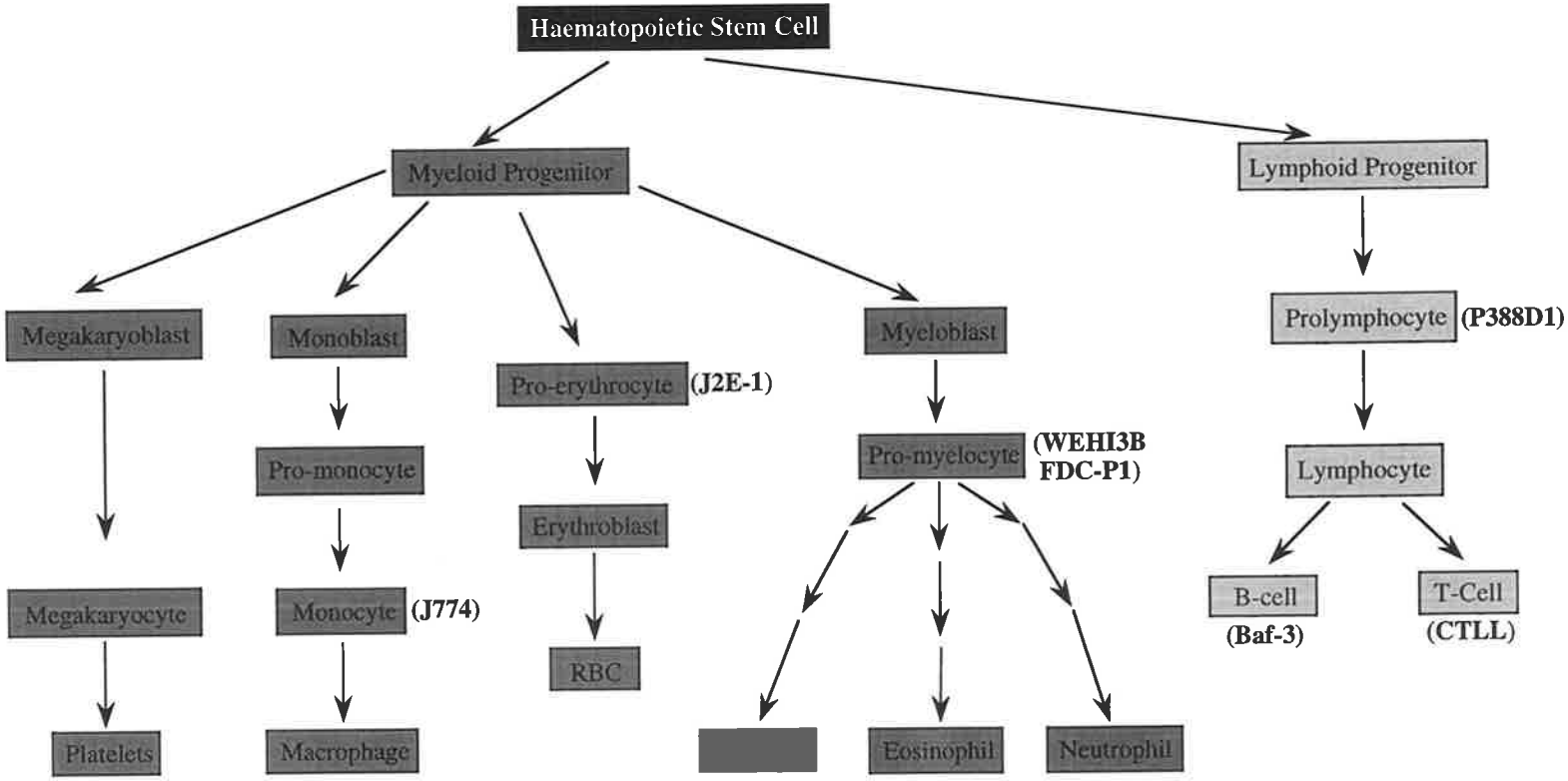

Te
Ov
Pl
Br
He
Li
Mu
St
Kid
Sp
BM
MLN
t
t
Th
K7
mGap

Figure 4.16: Mouse haematopoiesis

Schematic representation of the cell lineages derived from haematopoietic stem cells. Haematopoietic stem cells differentiate, in response to different cytokines, into myeloid or lymphoid progenitors. Lymphoid progenitors then differentiate into either B-cells in the bone marrow, or T-cells in the thymus. Myeloid progenitors give rise to a number of lineages in the bone marrow including megakaryocytes which differentiate into platelets, monocytes which differentiate into macrophages, erythrocytes which differentiate into red blood cells (RBC) and myelocytes which differentiate into granulocytes (basophils, eosinophils and neutrophils). The cell lines used in analysis of *K7* expression in the haematopoietic lineages *in vitro* are indicated besides the various lineages that they represent.

Adapted from Abbas *et al.* (1991)


Haematopoietic Stem Cell
Myeloid Progenitor
Lymphoid Progenitor
Megakaryoblast
Monoblast
Pro-erythrocyte
(J2E-1)
Myeloblast
Prolymphocyte
(P388D1)
Megakaryocyte
Pro-monocyte
Erythroblast
Pro-myelocyte
(WEHI3B
FDC-P1)
Lymphocyte
Monocyte
(J774)
RBC
B-cell
(Baf-3)
T-Cell
(CTLL)
Platelets
Macrophage
Eosinophil
Neutrophil

a monocyte cell line (J774) and a pro-erythrocyte cell line (J2E-1), was analysed by RNAse protection analysis (2.3.21.1/2). While *K7* expression was detected in all the haematopoietic cell lines analysed, substantially higher expression was observed in the pro-erythrocyte cell line (J2E-1) (Figure 4.17).

4.5 DISCUSSION

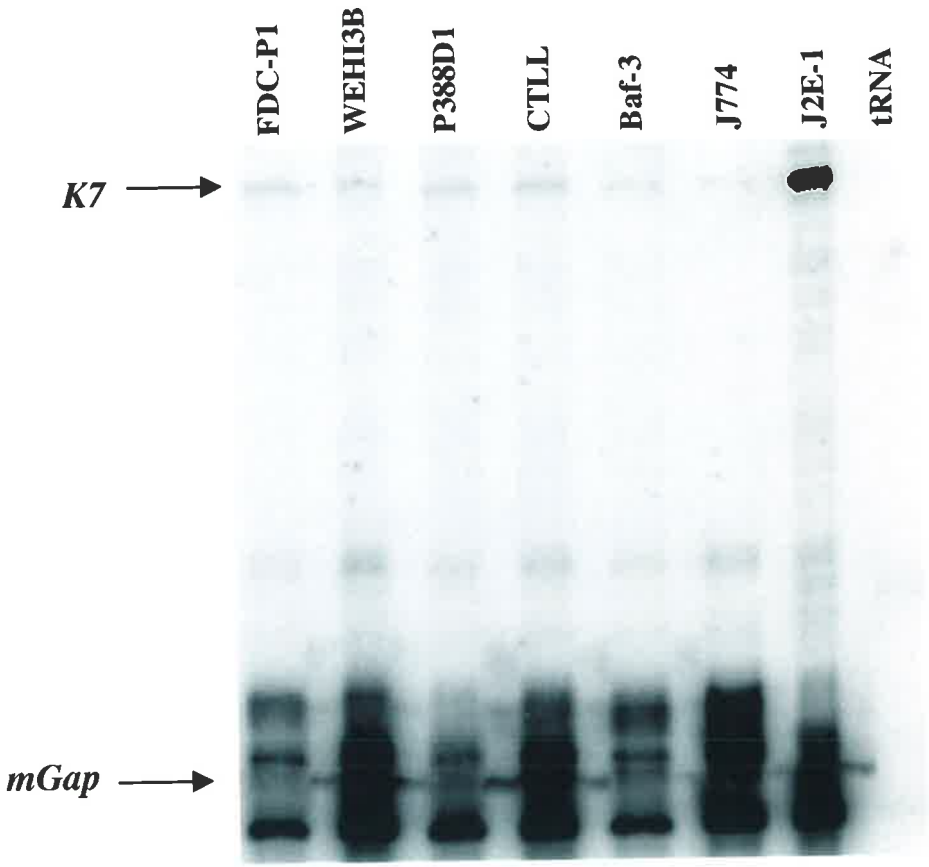
4.5.1 The *K7* cDNA

The complete *K7* cDNA sequence was compiled from the original dd-PCR clone and three different partial cDNA clones isolated from a λ ZAP II D3 ES cDNA library (Clontech). The cDNA sequence of 6583 bp in length was similar to the size of approximately 7 kb transcript estimated from northern analysis. The nature of the second, approximately 4.5 kb transcript, was not addressed and further investigation using RNAse protection and genomic analysis will be required to determine if the two *K7* transcripts identified by northern blot result from alternate transcription or alternate splicing events.

The DNA sequence was found to be 79.3% identical to a human gene termed *KIAA0165*, isolated as one of many cDNAs from the immature myeloid cell line, KG-1 (Nagase *et al.*, 1996). Analysis of the 5' region of the two DNA sequences indicated that the *KIAA0165* gene contained 81 bp of additional sequence, which shared no homology to the *K7* sequence, but sequence conservation between the two genes was maintained both 5' and 3' of the additional sequence. The additional sequence also contained a conserved 5' splice site and a near consensus 3' splice site, indicating that the sequence was likely to represent an unspliced intron in the *KIAA0165* cDNA, and that the first in frame ATG codon, conserved between *K7* and *KIAA0165*, was likely to be functional in the human. However, the additional sequence may simply represent a difference between the two genes within this region, or be a consequence of alternatively spliced exons which results in the addition of this sequence into the *KIAA0165* cDNA. RNAse protection analysis of *K7* and *KIAA0165*, using probes designed across this region and genomic analysis, or identification of translational start sites and analysis of endogenous proteins produced by these two genes will be required to determine the nature of this difference between the two genes.

Figure 4.17: Expression analysis of *K7* in haematopoietic cell lines by RNAse Protection

K7 expression was detected using a 376 bp antisense *K7* riboprobe (2.3.21.1). RNAse protection analysis (2.3.21.2) was performed on 10 μg of total RNA isolated from various mouse haematopoietic cell lines, and 10 μg of tRNA. Mouse *Glyceraldehyde phosphate dehydrogenase* (*mGap*) was used as a loading control.


FDC-P1
WEHI3B
P388D1
CTLL
Baf-3
J774
J2E-1
tRNA
K7
mGap

The *K7* cDNA encoded a 2118 amino acid protein of an estimated molecular weight of 233,185 Daltons. The amino acid sequence also shared high level identity with the *KIAA0165* gene product. However, possible functions of *KIAA0165* have not been investigated and consequently, little information regarding a possible function for *K7* could be determined from this considerable homology. Northern analysis of *KIAA0165* expression in adult human tissues (Nagase *et al.*, 1996) indicated similar expression to that observed for *K7*. The human transcript was expressed in a tissue specific manner, with highest level expression detected in the adult testis, and expression also detected in the thymus, peripheral blood leukocytes, small intestine and colon. Given the similar expression patterns of both genes in cells of the testis, haematopoietic system and gastrointestinal tract, and the high level of sequence identity at the DNA and protein level, it appears likely that *K7* is the mouse homologue of the human *KIAA0165* gene.

Fluorescent *in situ* hybridisation was used to identify the probable physical location of the *K7* locus at the distal end of chromosome 16, within bands C3.3-C4. Genes in the very distal bands of mouse Chromosome 16 (refer Figure 4.11 B) are homeologous with the distal region of human chromosome 21, bands q22.2-q22.3 (Cabin and Reeves, 1998). However, chromosome localisation of *KIAA0165* has been reported on human chromosome 8 (Nagase *et al.*, 1996), and recently revised to human chromosome 12 (Dr. T Nagase, personal communication). This identification was achieved using PCR, which did not reveal the sub-chromosomal location of the gene. Therefore, although *K7* and *KIAA0165* have been interpreted as homologues based on a high level of sequence identity at both the DNA and protein level and similar expression patterns in the adult, the localisation of the genes to non-syntenic chromosomes is inconsistent with this conclusion. However, the extreme distal localisation of *K7* on mouse chromosome 16 may provide an explanation for why the *K7* and *KIAA0165* genes reside on non-syntenic chromosomes. Bernard *et al.* (1996), described a similar phenomenon following physical localisation of the mouse *Kiz/Limk1* gene to the distal region of mouse chromosome 5, and the human *Kiz/Limk1* gene to the distal end of human chromosome 17 and not chromosome 7, as predicted from synteny data. Taken together, localisation of both genes to the distal ends of non-syntenic chromosomes provides evidence for extreme distal genes being subject to chromosomal exchange. If this is the case for *K7*, then *KIAA0165* would be expected to localised at the distal

tip of human chromosome 12. To resolve this question, fluorescent *in situ* hybridisation using a *KIAA0165* probe on human chromosome spreads could be carried out to determine the physical localisation of this gene.

Presently, the most distally mapped gene on mouse Chromosome 16 (*Mx2,* Cabin and Reeves, 1998; Figure 4.11 B) shares synteny with the distal end of chromosome 21 (21q22.3). Therefore, *K7* would be predicted to lie further distal on mouse Chromosome 16 than *Mx2*, as any genes distal to *K7* or *KIAA0165,* depending on which gene underwent chromosome exchange, would also be expected to have undergone translocation.

4.5.2 Potential Function of *K7* in Development and Adult tissues

K7 expression was detected in pluripotent cells both *in vitro* and *in vivo,* with expression within the embryo correlating with specific events during peri-implantation development. *K7* expression was upregulated *in vivo* at the stage when proliferation was proposed to increase prior to gastrulation (see 3.4.2). *K7* may therefore be required at the time when this increased cellular proliferation commences, as the developmental strategy changes from being cleavage based to a proliferative based regime. *In situ* analysis immediately prior to gastrulation and during early and later gastrulation stages did not detect *K7* transcripts, indicating that when proliferation within the primitive ectoderm is at its most rapid (Snow, 1977), *K7* is most likely not playing a role. However, the possibility that K7 protein persists following downregulation of the transcript cannot be excluded.

RNAse protection analysis during later development indicated that *K7* transcripts were developmentally regulated and differentially expressed from day 10.5 to day 17.5. *K7* transcripts were expressed at variable levels in all tissues analysed at day 16.5. Within adult tissues *K7* expression was more tightly regulated than at 16.5 days of embryogenesis. High level adult expression was detected in the testis, and lower level expression detected in the stomach and in haematopoietic tissues of the bone marrow, thymus and spleen. Interestingly, the testis, bone marrow and the gastrointestinal tract are tissues which contain stem cells required for tissue regeneration. *K7* expression within these stem cell containing tissues of the adult was consistent with early embryonic expression since all are sites of proliferation, required for elaboration of embryonic tissues during development and cell replenishment in the adult. This prediction was

further evidenced by two other observations. Firstly, *K7* expression was detected in adult thymus, a tissue which directs the development of immature thymic progenitors set aside during embryogenesis, into mature T-lymphocytes (Ikuta *et al.*, 1990). However, *K7* was not in adult mesenteric lymph nodes, which contain only mature lymphocytes. Secondly, high level expression of *K7* in the pro-erythrocyte cell line *in vitro* was consistent with rapid proliferation of these cells, whose *in vivo* equivalents are required to replenish the red blood cell population. *K7* may therefore function in stem cell maintenance/renewal or in accelerated progenitor cell proliferation. This was also consistent with a proposed role for *K7* in mitosis, following database comparisons at the protein level (4.2.3.2).

High level *K7* in the testis suggested a mitotic role for *K7* during spermatogenesis. This may reflect self-renewal of stem cells, termed type-A spermatogonia or rapid cell division of an intermediate cell type, type-B spermatogonia. However, differentiation of type-B spermatogonia into primary spermatocytes is followed by meiosis. During the second meiotic division sister chromatid separation occurs, producing gametes with a haploid number of chromosomes. Therefore, a potential involvement of *K7* in meiotic division can also be envisaged and investigated by mapping *K7* expression at a cellular level (see 6.3.1.5).

4.5.3 A Potential Cellular Function for *K7*

The C-terminal region of the deduced K7 protein sequence displayed substantial homology to the C-termini of fungal proteins identified during temperature sensitive based screens for mitotic defects. Three distinct fungal genes, *bimB* (*A. nidulans*), *cut1* (*S. pombe*) and *ESP1* (*S. cerevisiae*) have similar phenotypes when mutated (Baum *et al.*, 1988; Uzawa *et al.*, 1990; McGrew *et al.*, 1992; May *et al.*, 1992). Analysis of temperature sensitive *bimB, Cut1* and *ESP1* mutants at the restrictive temperature demonstrated that mutant cells contained abnormally large nuclei in which DNA replication continued to occur in the absence of chromosome segregation. Other cell cycle events were unaffected by the mutation and occurred with normal timing such that following cytokinesis, one daughter cell often received most if not all the genetic information and both spindle pole bodies. Mutant cells therefore contained polyploid DNA contents and multiple spindle pole bodies, or were anucleate. The protein sequence similarity between the three fungal genes was only observed within the C-terminus of the respective

proteins, suggesting that this region conferred a common function. This indicated potential involvement of K7 and KIAA0165 proteins in chromosome segregation in mouse and humans respectively.

4.5.3.1 Chromosome segregation in yeast

In eukaryotes, chromosomes are duplicated during S (synthesis) phase and segregated during mitosis (M phase) of the cell cycle. A temporal delay exists between S phase and M phase and is known as the G2 phase (gap phase 2). To ensure that chromosomes are segregated equally, sister chromatids are physically linked to each other by a multi-subunit complex known as the cohesin complex, from the time of replication, through G2 phase, until segregation during M phase. In *S. cerevisiae*, the cohesin complex is made up of at least four subunits, Scc1, Scc3, Smc1 and Smc3 (Michaelis *et al.*, 1997; Tóth *et al.*, 1999). Physical association of sister chromatids is essential for attachment at the kinetochore region to microtubules, which emanate from the microtubule organising centres, termed spindle pole bodies, at opposite poles of the cell. At the metaphase to anaphase transition of mitosis the attachment between sister chromatids must be dissolved in order for sister chromatids to separate. Two of the subunits, Scc1 and Scc3 lose chromatin attachment at this stage, leading to chromosome segregation as sister chromatids are pulled to opposite poles of the cell by the microtubules (Tóth *et al.*, 1999).

The onset of anaphase is also the stage at which the cell begins exit from M phase and move into the subsequent gap phase (G1), prior to entry into a new round of DNA replication. This transition is mediated by specific proteolysis of regulatory proteins, such as cyclin B, following ubiquitination of destruction box sequences within these proteins by the anaphase promoting complex (APC) (Holloway *et al.*, 1993).

Loss of sister chromatid cohesion also involves APC dependent, ubiquitin mediated proteolysis, since mutations in any one of at least 10 APC subunits prevents sister chromatid segregation. This is thought to ensure that mitotic exit and chromosome segregation occur coordinately. However, mutational studies indicate that proteolysis of known APC targets, such as cyclin B, are not required for chromosome segregation (Holloway *et al.*, 1993), indicating that additional protein(s) must be destroyed at the onset of anaphase to induce chromosome segregation. These protein(s) could themselves be part of the cohesin complex or a protein which

negatively regulates an activator of chromosome segregation (Holloway *et al.*, 1993). Analysis of *cut1* in *S. pombe* and *ESP1* in *S. cerevisiae* suggested that the later hypothesis was correct (Funabiki *et al.*, 1996a and b; Ciosk *et al.*, 1998). No further investigation of the function of *bimB* in sister chromatid separation has been reported.

4.5.3.2 The role of cut1 in sister chromatid separation in S. pombe

A mechanistic understanding of *cut1* activity began to emerge following identification of a second temperature sensitive *S. pombe* mutant, *cut2*. Mutation of *cut2* resulted in a similar phenotype to the *cut1* mutation, as multiple rounds of entry into the cell cycle were observed in the absence of nuclear division (Uzawa *et al.*, 1990; Funabiki *et al.*, 1996a). *Cut2* shared no sequence similarity to known proteins (Uzawa *et al.*, 1990). However, expression of Cut1 protein could rescue *cut2* mutants, and *cut1-cut2* double mutants were inviable, indicating that these two proteins were likely to be part of a pathway involved in the regulation of sister chromatid separation (Funabiki *et al.*, 1996a). A direct interaction between Cut1 and Cut2 was established by immunoprecipitation and co-sedimentation following centrifugation on sucrose gradients (Funabiki *et al.*, 1996a).

Cut1 and Cut2 displayed similar sub-cellular localisations during metaphase of mitosis, being distributed along the short mitotic spindle during metaphase, and Cut1 also localised to the spindle pole body (Funabiki *et al.*, 1996a; Kumada *et al.*, 1998). Cut2 disappeared from the elongating spindle at the onset of anaphase, while Cut1 remained localised to the shorter mitotic spindle during early anaphase, before disappearing from longer spindles during mid-late anaphase (Funabiki *et al.*, 1996a and b; Kumada *et al.*, 1998). The interaction of Cut1 with the early mitotic spindle appeared dependent on the Cut1-Cut2 interaction as localisation of Cut1 to mitotic spindles and spindle pole body occurred only in the presence of functional Cut2. Furthermore, deletion studies indicated that fragments of Cut1 which were unable to interact with Cut2 could not localise to the mitotic spindle, suggesting that Cut2 may be required to load Cut1 onto the mitotic spindle (Kumada *et al.*, 1998).

Analysis of protein stability during the cell cycle indicated that the Cut1 protein was present at relatively constant levels throughout the cell cycle. In contrast, Cut2 was degraded at anaphase and unstable in G1. The N-terminal sequence of Cut2 contained two sequences with

similarity to the mitotic cyclin B destruction boxes, and these sequence were found to mediate ubiquitination of Cut2 at the onset of anaphase by the APC, which targeted the protein for proteolysis by the proteosome (Funabiki *et al.*, 1997). This proteolysis was essential for sister chromatid separation as mutation of the *cut2* destruction box sequences led to a dominant negative phenotype in which sister chromatid separation was completely inhibited (Funabiki *et al.*, 1996b; Kumada *et al.*, 1998).

4.5.3.3 ESP1 function in S. cerevisiae

Accumulating evidence suggested that the mechanism of chromosome segregation was conserved between *S. pombe* and *S cerevisiae*. The *ESP1* gene product of *S. cerevisiae* gene was found to interact with Pds1 (Ciosk *et al.*, 1998). While *Pds1* shared no significant similarity with *cut2* at the sequence level, it did contain two destruction box sequences which targeted the protein for degradation at the metaphase to anaphase transition, following ubiquitination by the APC (Yamamoto *et al.*, 1996; Ciosk *et al.*, 1998). Furthermore, localisation of Esp1 to the mitotic spindle and spindle pole body during metaphase was dependent on the presence of Pds1 (Ciosk *et al.*, 1998). Association of Pds1 with Esp1 negatively regulated Esp1, as following Pds1 proteolysis at the onset of anaphase, Esp1 triggered the dissociation of a cohesin subunit, Scc1, from chromatin (Ciosk *et al.*, 1998). This effect was blocked in non-degradable *pds1* mutant cells (Ciosk *et al.*, 1998). Further experiments indicated that Esp1 not only dissociated Scc1 from chromatin, but the cohesin subunit underwent proteolytic cleavage in response to Esp1 (Uhlmann *et al.*, 1999). Therefore, Esp1 was proposed to function in one of two ways; as a protease which mediates Scc1 cleavage, or as an activator of a novel protease. Recent analysis of the C-terminal region conserved between Esp1, Cut1, bimB, and KIAA0165 identified an invariant histidine and cysteine residue which were flanked by amino acids which shared some similarity to cysteine endopeptidases of the CD subclass (Nasmyth *et al.*, 2000). This subclass includes caspases, legumains and two bacterial cysteine proteases (Chen *et al.*, 1998). Therefore, Esp1 may be the protease which cleaves Scc1, however, this proposition has yet to be tested experimentally. The invariant histidine and cysteine residues are also conserved in the K7 C-terminal protein sequence, indicating that K7 could also contain protease activity.

Therefore, the Esp1/Pds1 complex of *S. cerevisiae* appeared functionally analogous with the Cut1/Cut2 complex of *S. pombe* (Figure 4.18). Interaction of Pds1 or Cut2 with Esp1 or Cut1 respectively, inhibited the function of Esp1 and Cut1, until degradation of Pds1 and Cut2 at the metaphase to anaphase transition by APC mediated ubiquitination. Loss of Esp1 negative regulation allows this protein to induce chromosome segregation, most likely by proteolytic cleavage of the cohesin subunit, Scc1 (Uhlmann *et al.*, 1999).

4.5.3.4 Sister chromatid separation in vertebrates

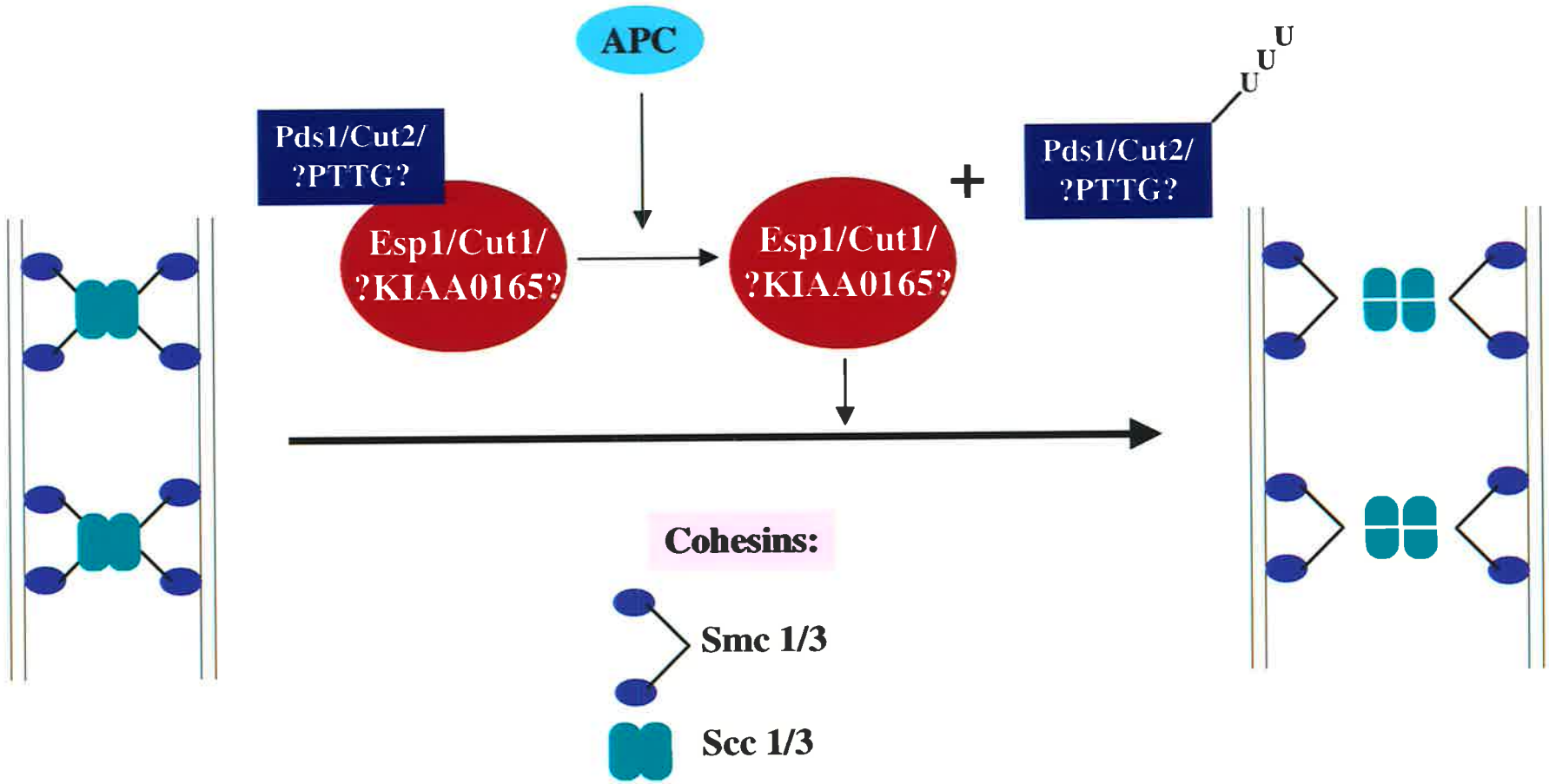
Immunoprecipitation analysis using antibodies raised to the putative KIAA0165 protein and purification of proteins which interacted with KIAA0165 in lysates of metaphase arrested cells, but not extracts prepared four hours following release from the nocodozole block, lead to purification of a known protein encoded by the *pituitary tumour transforming gene* (PTTG) (Zou *et al.*, 1999). PTTG was originally identified in the rat and was cloned based on its high level expression in pituitary tumours (Pei and Melmed, 1997). Human PTTG (hPTTG) was also cloned and found to be expressed at high levels in the testis, and at lower levels in the thymus, spleen, colon, and small intestine (Zhang *et al.*, 1999). hPTTG was thus expressed in similar tissues described for KIAA0165, indicating that hPTTG has the potential to interact with KIAA0165 *in vivo*. Stable overexpression of rPTTG in NIH 3T3 cells inhibited cell proliferation, consistent with a role in mitosis, and induced transformation *in vitro* and tumourogenesis upon injection of these cells *in vivo* (Pei and Melmed, 1997). Many transformed cells demonstrated chromosome instability, resulting in aneuploidy. The tumourogenic nature of cells overexpressing rPTTG suggests that defects in chromosome segregation could potentially underlie some cancers (Lengauer *et al.*, 1998).

Investigation of *Xenopus* PTTG (xPTTG) and hPTTG protein levels during the cell cycle demonstrated that the PTTG protein oscillated in a cell cycle dependent manner. The protein accumulated from the onset of S phase, peaked during G2/M phase, before degrading coincident with cyclin B degradation late in mitosis (Zou *et al.*, 1999). The presence of destruction box sequences in the PTTG protein resulted in degradation of the protein, as mutation of this sequence prevented xPTTG proteolysis. Furthermore, exogenous addition of a dominant negative xPTTG, in which the destruction box had been mutated, to metaphase arrested *Xenopus* egg

Figure 4.18: Schematic representation of chromosomal segregation at the metaphase to anaphase transition during mitosis.

From the end of S phase until the onset of anaphase, sister chromatids are held together by the cohesin complex, which includes the structural maintenance of chromosome proteins (SMC) 1 and 3, which mediates binding to chromosomes, and the Scc1 and 3 proteins, which mediate association between SMC proteins bound to alternate sister chromatids. During this time, Cut1 or Esp1 activity is inhibited by association with destruction box containing proteins, Cut2 or Pds1 respectively. Anaphase promoting complex (APC) mediated ubiquitination of Cut2/Pds1 at the onset of anaphase targets these proteins for degradation, allowing Cut1/Esp1 to induce sister chromatid separation by triggering the proteolysis of Scc1. Recently, the gene product of *KIAA0165*, which displays C-terminal similarity to Cut1 and Esp1 protein sequences, was demonstrated to interact with PTTG, a destruction box containing protein which is degraded late in mitosis in human cells. Therefore, interaction of KIAA0165 with PTTG in humans may be analogous to the Cut1/Cut2 and Esp1/Pds1 interactions in *S. pombe* and *S. cerevisiae*.

Adapted from Uhlmann *et al.* (1999).


Metaphase
Anaphase
APC
Pds1/Cut2/
?PTTG?
Esp1/Cut1/
?KIAA0165?
Esp1/Cut1/
?KIAA0165?
+
Pds1/Cut2/
?PTTG?
U
U
U
Cohesins:
Smc 1/3
Scc 1/3

extracts inhibited chromosome segregation at anaphase (Zou *et al.*, 1999), suggesting a negative regulatory role for this protein in the control of chromosome segregation in mammals.

Therefore, interaction of PTTG with KIAA0165 in humans was reminiscent of the Cut1/Cut2 and Esp1/Pds1 interaction in *S. pombe* and *S. cerevisiae* (Figure 4.18). Furthermore, homology of K7 and KIAA0165 with the C-terminal region of Cut1 and Esp1 suggested a role for these mammalian proteins in the control of chromosome segregation, possibly by proteolysis of cohesin subunits. This proposal was also consistent with expression studies which postulated a mitotic role for *K7* in stem cell renewal or rapid proliferation associated with progenitor cell populations, and perhaps a meiotic role for *K7* during spermatogenesis in the testis, presumably in chromosome segregation, which occurs twice during meiosis.

CHAPTER 5:

FUNCTIONAL ANALYSIS OF THE *K7* GENE

5.1 INTRODUCTION

K7 shared 79.3% identity at the DNA level with a human gene of unknown function, termed *KIAA0165*, and demonstrated an almost identical expression pattern to *KIAA0165* in adult tissues, suggesting that these genes were homologues. The deduced amino acid sequence of K7 and KIAA0165 also displayed similarity in a C-terminal region with proteins from three different fungi, Cut1 (*S. pombe*), Esp1 (*S. cerevisiae*) and bimB (*A. nidulans*), which have been implicated in sister chromatin separation during mitosis (Baum *et al.*, 1988; Uzawa *et al.*, 1990; McGrew *et al.*, 1992; May *et al.*, 1992). Cut1 and Esp1 have also been demonstrated to localise to the mitotic spindle and spindle pole body, the yeast equivalent of centrosomes, consistent with a role for these proteins in sister chromatid separation at the metaphase to anaphase transition (Funabiki *et al.*, 1996a; Kumada *et al.*, 1998; Ciosk *et al.*, 1998). In vertebrates, the protein encoded by the *pituitary tumour transforming gene* (*PTTG*) has been shown to interact with KIAA0165 by immunoprecipitation (Zou *et al.*, 1999). Furthermore, exogenous addition of a dominant negative PTTG protein, which contained a mutated destruction box, to *Xenopus* oocytes synchronised in metaphase prevented chromosome segregation, indicating an inhibitory role of PTTG in chromosome segregation (Zou *et al.*, 1999). This was reminiscent of negative regulation of Cut1 and Esp1 via interaction with Cut2 and Pds1 respectively (Funabiki *et al.*, 1996b; Ciosk *et al.*, 1998).

This chapter therefore investigated a possible function for *K7* during mitosis in three ways. Firstly, immunoprecipitation analysis was used to identify an *in vivo* interaction between K7 and mouse PTTG. Secondly, K7 and mPTTG protein localisation was analysed in transfected Cos-1 cells, and finally, analysis of cell cycle changes in Cos-1 cells transfected with K7 and mPTTG expressing constructs was used to provide functional information. In order to investigate functional sequence conservation across species, the C-terminal sequence of *ESP1* was replaced with the *K7* C-terminal sequence to determine if the K7 C-terminus was capable of rescuing an *ESP1* temperature sensitive mutant strain of *S. cerevisiae*.

5.2 *K7* FUNCTION IN MAMMALIAN CELLS

5.2.1 K7 Interacts with Mouse PTTG protein

Investigation of protein-protein interactions by immunoprecipitation requires antibodies specific for the proteins of interest. Since antibodies specific for K7 and mPTTG were not available, K7 and mPTTG containing mammalian expression constructs were generated as fusion proteins with EGFP (Clontech) and FLAG (Sigma) moieties respectively. Production of tagged fusion constructs for K7 and mPTTG also allowed sub-cellular protein localisation to be determined using fluorescent microscopy, in the case of EGFP-K7, and by immunohistochemistry using anti-FLAG antibodies, in the case of mPTTG-FLAG.

5.2.1.1 Cloning of mPTTG and construction of a mPTTG-FLAG containing mammalian expression vector

Cloning of mPTTG was carried out by RT-PCR (2.3.16/17) of RNA isolated from the mouse testis, using 5' and 3' oligonucleotides designed against the unpublished mPTTG sequence obtained from GenBank (Accession number AF069051; 2.2.7.3). The PCR product which contained the mPTTG open reading frame was ligated into the pGemT-Easy vector according to the manufacturer's instructions (Promega). The resulting clone, termed pmPTTG-GemT-Easy was then sequenced (2.3.13). The PCR product contained two sequence discrepancies when compared to the GenBank sequence; a nine bp insertion at nucleotide position 363, and a C to G substitution at nucleotide position 432 of the RT-PCR sequence, which resulted in substitution of isoleucine for methionine (Figure 5.1). However, the published rat PTTG sequence (Pei and Melmed, 1997; Accession number U73030) contained an equivalent nine base pair insertion and encoded a methionine at nucleotide position 432 bp. Therefore, the mouse sequence amplified by RT-PCR was used in all subsequent analyses.

mPTTG was re-PCR amplified using 5' *Eco*RI mPTTG and 3' FLAG mPTTG primers (2.2.7.3) from the pmPTTG-GemT-Easy construct, to incorporate a FLAG tag at the 3' end of the protein (Figure 5.2). The PCR product was digested with *Eco*RI and ligated into a mammalian expression vector, pXMT2 (Rathjen *et al.*, 1990a), to produce the clone pmPTTG-FLAG-X.

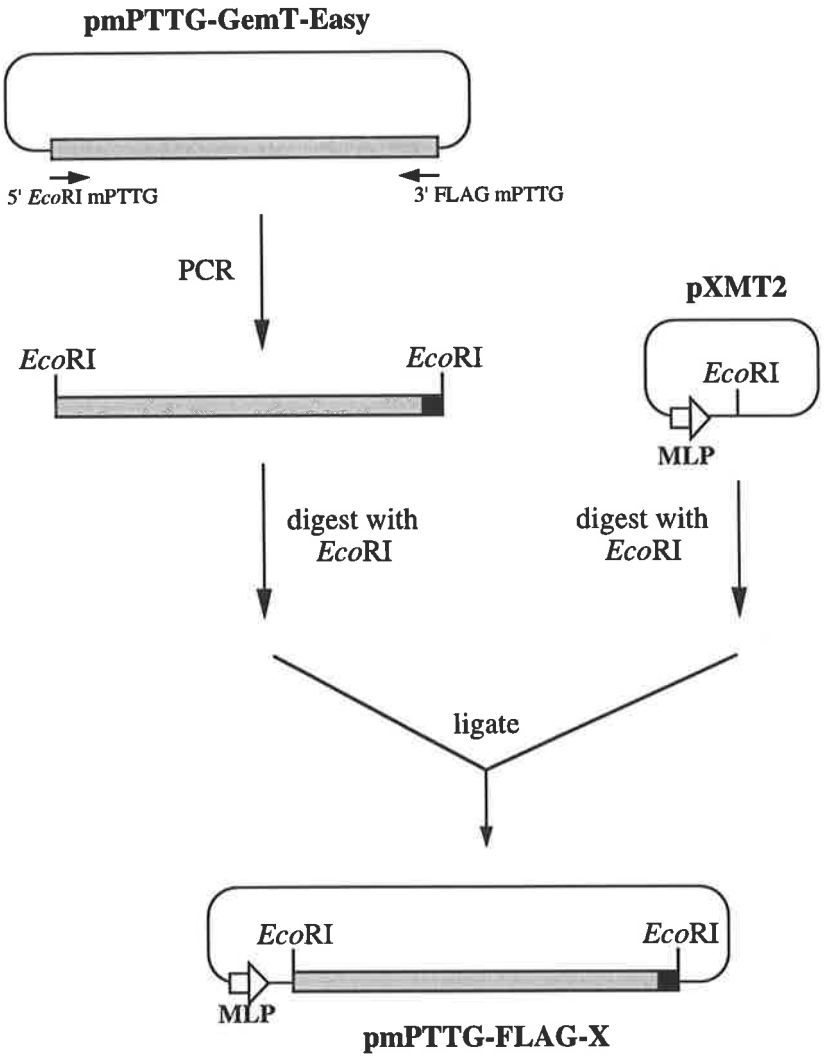
Figure 5.1: Alignment of PTTG sequence amplified by RT-PCR with the PTTG sequence of mouse and rat from GenBank

Alignment of the PTTG sequence obtained by RT-PCR of mouse testis RNA with the unpublished mouse PTTG (mPTTG) cDNA sequence in GenBank (Accession number AF069051) and the published rat PTTG (rPTTG) sequence (Pei and Melmed, 1997; Accession number U73030). A 9 bp insertion at position 363 bp and base pair change at nucleotide position 432 are indicated by dark shading. Two further sequence discrepancies between the RT-PCR sequence and mPTTG and/or rPTTG, at nucleotide positions 297 and 363 (star), were observed but did not alter the amino acids at these positions. The sequence differences between rPTTG and mPTTG are indicated by light shading. Comparisons were carried out using the CLUSTALW and Boxshade programs (2.3.33).


RT-PCR 1 ATGGCTACTCTTATCTTTGTTGATAAGGATAATGAAGAACCCGGCCGCCGTTTGGCATCT
mPTTG 1 ATGGCTACTCTTATCTTTGTTGATAAGGATAATGAAGAACCCGGCCGCCGTTTGGCATCT
rPTTG 1 ATGGCTACTCTGATCTTTGTTGATAAGGATAACGAAGAGCCAGGCAGCCGTTTGGCATCT
RT-PCR 61 AAGGATGGGTTGAAGCTGGGCACTGGTGTCAAGGCCTTAGATGGGAAATTGCAGGTTTCA
mPTTG 61 AAGGATGGGTTGAAGCTGGGCACTGGTGTCAAGGCCTTAGATGGGAAATTGCAGGTTTCA
rPTTG 61 AAGGATGGATTGAAGCTGGGCTCTGGTGTCAAAGCCTTAGATGGGAAATTGCAGGTTTCA
RT-PCR 121 ACGCCTCGAGTCGGCAAAGTGTTCAATGCTCCAGCCGTGCCTAAAGCCAGCAGAAAGGCT
mPTTG 121 ACGCCTCGAGTCGGCAAAGTGTTCAATGCTCCAGCCGTGCCTAAAGCCAGCAGAAAGGCT
rPTTG 121 ACGCCACGAGTCGGCAAAGTGTTCGGTGCCCCAGGCTTGCCTAAAGCCAGCAGGAAGGCT
RT-PCR 181 TTGGGGACAGTCAACAGAGTTGCCGAAAAGCCTATGAAGACTGGCAAACCCCTCCAACCA
mPTTG 181 TTGGGGACAGTCAACAGAGTTGCCGAAAAGCCTATGAAGACTGGCAAACCCCTCCAACCA
rPTTG 181 CTGGGAACTGTCAACAGAGTTACTGAAAAGCCAGTGAAGAGTAGTAAACCCCTGCAATCG
RT-PCR 241 AAACAGCCGACCTTGACTGGGAAAAAGATCACCGAGAAGTCTACTAAGACACAAAGTTCT
mPTTG 241 AAACAGCCGACCTTGACTGGGAAAAAGATCACCGAGAAGTCTACTAAGACACAAAGCTCT
rPTTG 241 AAACAGCCGACTCTGAGTGTGAAAAAGATCACCGAGAAGTCTACTAAGACACAAGGCTCT
RT-PCR 301 GTTCCTGCTCCTGATGATGCCTACCCAGAAATAGAAAAGTTCTTCCCTTTCAATCCTCTA
mPTTG 301 GTTCCTGCTCCTGATGATGCCTACCCAGAAATAGAAAAGTTCTTCCCTTTCAATCCTCTA
rPTTG 301 GCTCCTGCTCCTGATGATGCCTACCCAGAAATAGAAAAGTTCTTCCCCTTCGATCCTCTA
RT-PCR 361 GACTTTTGAGAGTTTTGACCTGCCTGAGGAGCACCAGATCTCACTTCTCCCCTTGAATGGC
mPTTG 361 GA---------TTTTGACCTGCCTGAGGAGCACCAGATCTCACTTCTCCCCTTGAATGGC
rPTTG 361 GATTTTGAGAGTTTTGACCTGCCTGAAGAGCACCAGATCTCACTTCTCCCCTTGAATGGA
RT-PCR 421 GTGCCTCTCATGACCCTGAATGAAGAGAGAGGGCTGGAGAAGCTGCTGCATCTGGGCCCC
mPTTG 412 GTGCCTCTCATCACCCTGAATGAAGAGAGAGGGCTGGAGAAGCTGCTGCATCTGGGCCCC
rPTTG 421 GTGCCTCTCATGATCCTGAATGAAGAGAGGGGGCTTGAGAAGCTGCTGCACCTGGACCCC
RT-PCR 481 CCTAGCCCTCTGAAGACACCCTTTCTATCATGGGAATCTGATCCGCTGTACTCTCCTCCC
mPTTG 472 CCTAGCCCTCTGAAGACACCCTTTCTATCATGGGAATCTGATCCGCTGTACTCTCCTCCC
rPTTG 481 CCTTCCCCCTCTGCAGAAGCCCTTCCTACCGTGGGAATCTGATCCGTTGCCGTCTCCTCCC
RT-PCR 541 AGTGCCCTCTCCACTCTGGATGTTGAATTGCCGCCTGTTTGTTACGATGCAGATATTTAA
mPTTG 532 AGTGCCCTCTCCACTCTGGATGTTGAATTGCCGCCTGTTTGTTACGATGCAGATATTTAA
rPTTG 541 AGCGCCCTCTCCGCTCTGGATGTTGAATTGCCGCCTGTTTGTTACGATGCAGATATTTAA

Figure 5.2: Construction of a mPTTG-FLAG containing expression plasmid

To incorporate a FLAG tag epitope at the C-terminus of the mPTTG protein, the mPTTG sequence obtained by RT-PCR of mouse testis RNA and cloned into pGemT-Easy (pmPTTG-GemT-Easy), was re-PCR amplified using 5' *Eco*RI mPTTG and 3' FLAG mPTTG primers (2.2.7.3). The PCR product obtained was digested with *Eco*RI and cloned into *Eco*RI digested pXMT2 (Rathjen *et al.*, 1990a), an expression vector which drives expression in mammalian cells using the adenovirus late promoter (MLP).


pmPTTG-GemT-Easy
5' EcoRI mPTTG
3' FLAG mPTTG
PCR
EcoRI
EcoRI
pXMT2
EcoRI
MLP
digest with EcoRI
digest with EcoRI
ligate
EcoRI
EcoRI
MLP
pmPTTG-FLAG-X

5.2.1.2 Cloning of a full length K7 cDNA construct and pEGFP-K7 mammalian expression plasmid

A full length *K7* cDNA construct was created from the three independent *K7* cDNA clones, S1, 9.1.1 and 7A (4.2.1; refer Figure 4.1). The cloning strategy is illustrated in Figure 5.3. Briefly, the S1 cDNA clone in pBluescript SK (S1pBs-SK) was digested with *Bam*HI, and the excised insert ligated into the 9.1.1 cDNA clone (9.1.1pBs-SK), which had also been digested with *Bam*HI. This clone, termed 9.1.1/S1pBs-SK, which contained *K7* sequence from 625 bp to 6561 bp, was then digested with *Eco*RI and the insert cloned in reverse orientation into *Eco*RI cut pBluescript KS to produce the clone 9.1.1/S1pBs-KS. *K7* cDNA clone 7A (7ApBs-SK) was used as a template for PCR amplification of the 5' end using 5' *Sac*I and 3' *Xba*I primers (2.2.7.3) (Figure 5.3 B). This step removed the 5' UTR region which contained an in frame stop codon and would prevent read through from the N-terminal EGFP moity. The PCR product was digested with *Sac*I/*Xba*I and was ligated into *Sac*I/*Xba*I digested 9.1.1/S1pBs-KS, to produce the clone K7pBs-KS. The full length *K7* sequence was then excised by *Sac*I/*Eco*RI digestion and ligated into *Sac*I/*Eco*RI digested pEGFP-C2, producing the clone pEGFP-K7 which contained the entire K7 open reading frame fused in frame and C-terminal to the EGFP sequence.

5.2.1.3 Immunoprecipitation analysis of K7 and mPTTG

The pEGFP-K7 and pmPTTG-FLAG-X constructs were transfected into Cos-1 cells (2.6.7ii) and analysed for protein expression in cell lysates (2.3.27) by western blot (2.3.29/30), using anti-GFP and anti-FLAG antibodies (2.2.8) respectively. Proteins of the expected molecular weights were obtained for both pEGFP-K7 (approximately 250 kDa) and pmPTTG-FLAG-X (approximately 26 kDa) (Figure 5.4 A and B, arrows). A low abundance, larger protein resulting from transfection of pEGFP-K7 was identified, and may represent a post-translational modification to the EGFP-K7 protein. The presence of a 31 kDa FLAG-containing protein in pmPTTG-FLAG-X transfected Cos-1 cells was attributed to either translation initiation from an in frame CTG codon within pXMT2 vector, 115 bp upstream of the mPTTG translation initiation site (Voyle, 1999), or from a post-translational modification such as ubiquitination of the mPTTG protein. Smaller bands detected in the pmPTTG-FLAG-X lysate may represent degradation

Figure 5.3: Construction of a full length *K7* cDNA clone as a C-terminal fusion protein with enhanced green fluorescent protein (EGFP)

A) The first step of a three step cloning strategy to create a full length *K7* cDNA plasmid. The *K7* cDNA clone S1pBs-SK was digested with *Bam*HI and ligated into *Bam*HI digested 9.1.1pBs-SK to produce the clone 9.1.1/S1pBs-SK. In order to clone in the 5' end of the *K7* cDNA, the orientation of the construct was reversed by *Eco*RI digestion of 9.1.1/S1pBs-SK, and cloning into *Eco*RI cut pBluescript KS (9.1.1/S1pBs-KS).

B) The 5' end of the *K7* cDNA, excluding the 5' untranslated region, was amplified by PCR using the *K7* cDNA clone 7ApBs-SK as a template and 5' *Sac*I and 3' *Xba*I primers (2.2.7.3). Following digestion with *Sac*I and *Xba*I the fragment was cloned into *SacI/Xba*I digested 9.1.1/S1pBs-KS. The full length *K7* cDNA (K7pBs-KS) was then excised using *Sac*I and *Eco*RI and cloned into a *SacI/Eco*RI digested enhanced green fluorescent protein (EGFP) containing mammalian expression vector (Clontech), termed pEGFP-C2, as a C-terminal fusion protein to EGFP.

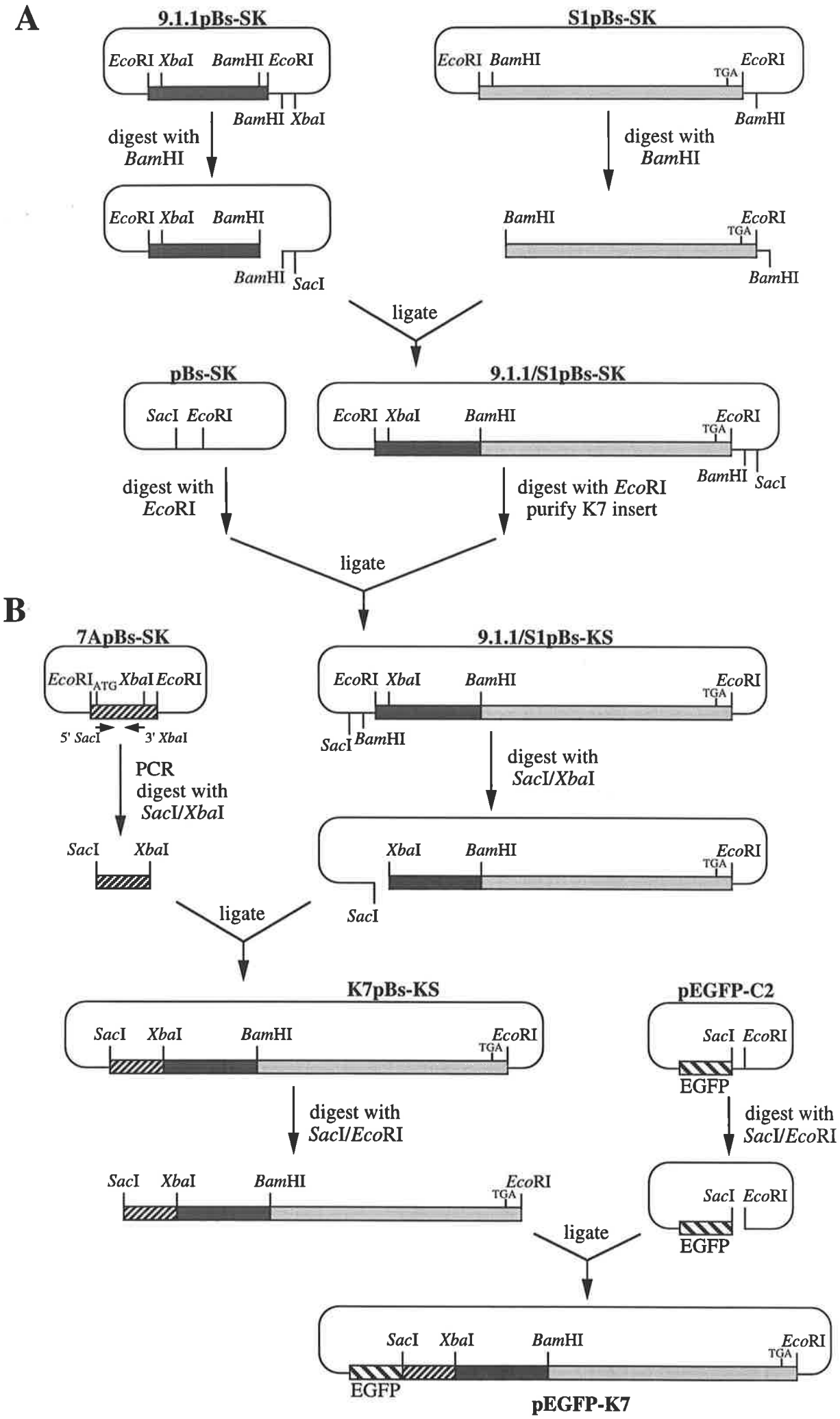
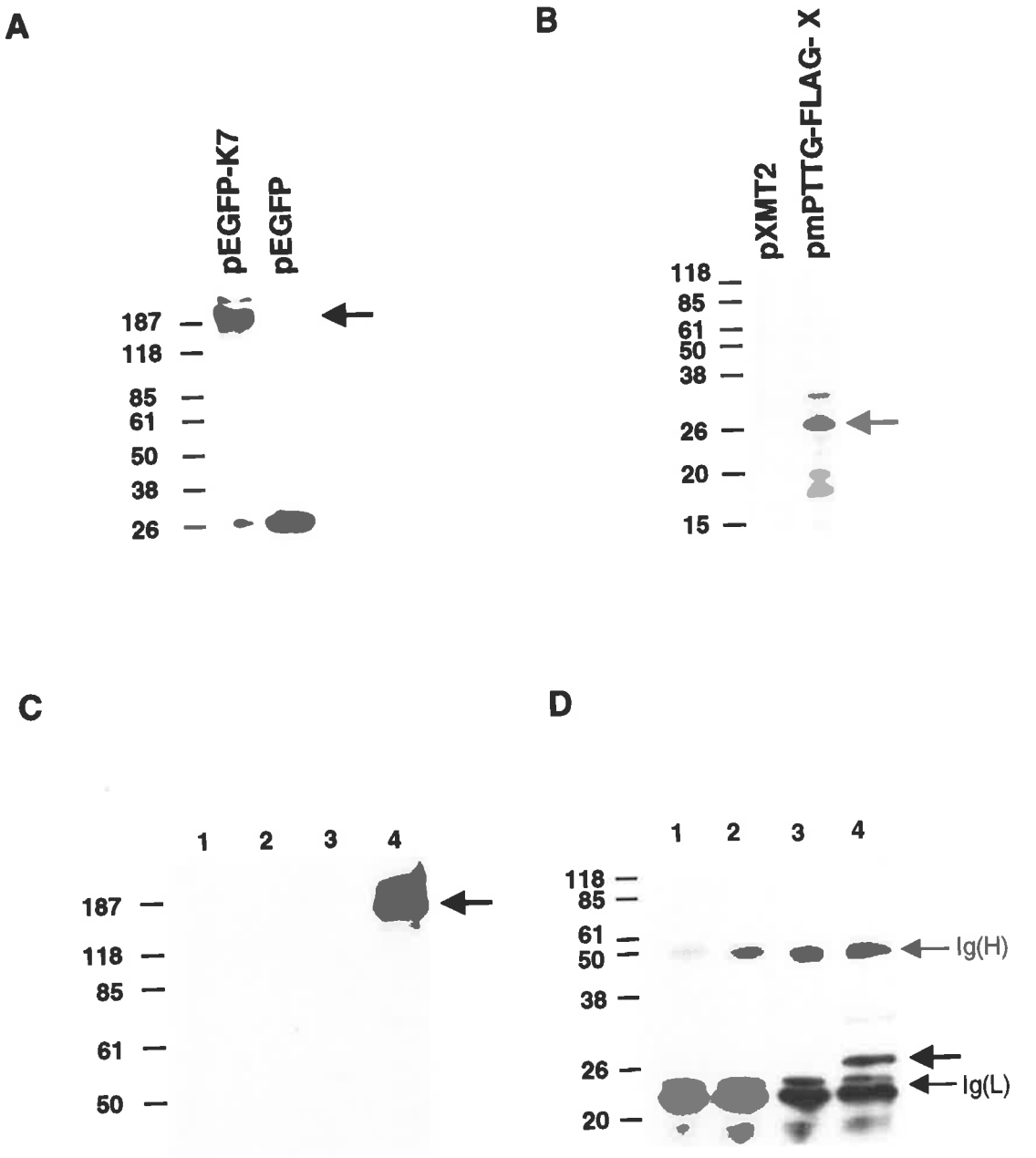

A
9.1.1pBs-SK
EcoRI XbaI BamHI EcoRI
BamHI XbaI
digest with BamHI
EcoRI XbaI BamHI
BamHI SacI
S1pBs-SK
EcoRI BamHI
EcoRI
TGA
BamHI
digest with BamHI
BamHI
EcoRI
TGA
BamHI
ligate
pBs-SK
SacI EcoRI
9.1.1/S1pBs-SK
EcoRI XbaI BamHI
EcoRI
TGA
BamHI SacI
digest with EcoRI
digest with EcoRI
purify K7 insert
ligate
B
7ApBs-SK
EcoRI ATG XbaI EcoRI
5' SacI
3' XbaI
PCR
digest with
SacI/XbaI
SacI XbaI
9.1.1/S1pBs-KS
EcoRI XbaI BamHI
EcoRI
TGA
SacI BamHI
digest with
SacI/XbaI
XbaI BamHI
EcoRI
TGA
SacI
ligate
K7pBs-KS
SacI XbaI BamHI
EcoRI
TGA
digest with
SacI/EcoRI
SacI XbaI BamHI
EcoRI
TGA
pEGFP-C2
SacI EcoRI
EGFP
digest with
SacI/EcoRI
SacI EcoRI
EGFP
ligate
SacI XbaI BamHI
EcoRI
TGA
EGFP
pEGFP-K7

Figure 5.4: Immunoprecipitation analysis of K7 and mPTTG in Cos-1 cells

A & B) Cos-1 cells, grown in 6 cm dishes, were transfected with 2 μg of plasmid DNA and lysed 48 hours post-transfection (2.3.27). 10 μl of the protein extract was loaded onto an 8% SDS-polyacrylamide gel **(A)** or a 12.5% SDS-polyacrylamide gel **(B)**, along with 10 μl of Benchmark prestained protein ladder (2.2.12), and electrophoresed (2.3.29), prior to western blotting (2.3.30). Filter A was incubated with anti-GFP antibody (diluted 1/5000) and filter B with anti-FLAG antibody (diluted 1/500). Following incubation with the appropriate horse radish peroxidase (HRP)-conjugated secondary antibody (2.2.8), the filters were developed using chemiluminescence (2.3.30).

C & D) Cos-1 cells, grown in 6 cm dishes, were co-transfected with 1 μg of the appropriate plasmid DNA and analysed for an interaction between K7 and mPTTG by immunoprecipitation (2.3.28) and western blotting (2.3.30). mPTTG was immunoprecipitated using 13.2 μg of anti-FLAG antibody (C) and K7 protein was immunoprecipitated using 1 μg of anti-GFP antibody (D). Immunoprecipitations and 10 μl of Benchmark prestained protein ladder (2.2.12) were electrophoresed (3.2.29) on 8% **(C)** and 15% **(D)** SDS-polyacrylamide gels, and transferred to nitrocellulose. Filter C was incubated with anti-GFP antibody (diluted 1/5000) to detect any GFP (K7) protein in anti-FLAG immunoprecipitates. Filter D was incubated with anti-FLAG antibody (diluted 1/500) in order to detect FLAG-containing protein (mPTTG) in anti-GFP immunoprecipitates. Following incubation with the appropriate HRP-conjugated secondary antibody (2.2.8), the filters were developed using chemiluminescence (2.3.30). Protein bands present in all lanes of filter D represent immunoglobulin light (Ig(L)) and heavy chains (Ig(H)) recognised non-specifically by the goat anti-mouse HRP-conjugated secondary antibody (2.2.8). Lane 1 represents immunoprecipitations from cells co-transfected with pXMT2 and pEGFP, lane 2 represents immunoprecipitations from cells co-transfected with pmPTTG-FLAG-X and pEGFP, lane 3 represents immunoprecipitations from cells co-transfected with pmPTTG-FLAG-X and pEGFP-K7, and lane 4 represents immunoprecipitations from cells co-transfected with pXMT2 and pEGFP-K7.

EGFP-K7 protein of filters A and C, and mPTTG-FLAG protein on filters B & D are indicated by bold arrows.


A
pEGFP-K7
pEGFP
187
118
85
61
50
38
26
B
pXMT2
pmPTTG-FLAG- X
118
85
61
50
38
26
20
15
C
1
2
3
4
187
118
85
61
50
D
1
2
3
4
118
85
61
50
38
26
20
Ig(H)
Ig(L)

products of the mPTTG protein, as this protein has been demonstrated to be unstable and proteolysed at the onset of anaphase during mitosis (Zou *et al.*, 1999).

Expression constructs were co-transfected into Cos-1 cells and analysed by immunoprecipitation and western analysis (2.3.28-30) for an interaction between K7 and mPTTG. Immunoprecipitation using anti-FLAG antibodies and western blotting using anti-GFP antibodies is shown in Figure 5.4 C. No products were detected in anti-FLAG immunoprecipitates from Cos-1 cells co-transfected with both expression vectors alone (lane 1; Figure 5.4 C), or pXMT2 and pEGFP in concert with either pEGFP-K7 or pmPTTG-FLAG-X respectively (lanes 2 and 3; Figure 5.4 C). However, immunoprecipitation of cell lysates co-transfected with pEGFP-K7 and pmPTTG-FLAG-X using anti-FLAG antibodies, followed by western analysis using anti-GFP antibodies resulted in detection of protein bands of the molecular weight expected for EGFP-K7 protein (lane 4; Figure 5.4 C), indicating that mPTTG and K7 interact in mammalian cells.

The reciprocal experiment confirmed this interaction as immunoprecipitation of K7 interacting proteins using anti-GFP antibodies from cell lysates co-transfected with pEGFP-K7 and pmPTTG-FLAG-X and western analysis using anti-FLAG antibodies, detected a band of the size expected for mPTTG-FLAG protein (lane 4; Figure 5.4 D). This protein was not detected in extracts from cell lysates co-transfected with both expression vectors alone (lane 1; Figure 5.4 D), or pXMT2 and pEGFP in concert with pEGFP-K7 or pmPTTG-FLAG-X respectively (lanes 2 and 3; Figure 5.4 D).

5.2.2 K7 Protein Localises to Centrosomes

5.2.2.1 Sub-cellular localisation of K7

The sub-cellular localisation of proteins often provides insight into possible protein function by determining the cellular compartment occupied by the protein. Localisation of Cut1 protein in a synchronous cell population in yeast demonstrated that this protein was completely excluded from the nucleus during interphase, but became localised to the mitotic spindle and spindle pole body during early mitosis, before being lost from the elongating spindle during mid-late anaphase (Funabiki *et al.*, 1996a; Kumada *et al.*, 1998). This was consistent with the suggested role for Cut1 in the control of chromosome segregation during mitosis.

Two attempts were made to generate anti-K7 polyclonal antibodies from two distinct regions of the K7 protein, as C-terminal fusions with glutathione S-transferase (GST) protein from the helminth *Schistosoma japonicum* (data not shown; see section 2.5). However, analysis of polyclonal antisera obtained from rabbits immunised with either GST-K7 peptide, using western blotting and immunohistochemistry of Cos-1 cells overexpressing K7, demonstrated that no anti-K7 antibodies were produced following immunisation with either K7 peptide (data not shown). Therefore, K7 protein localisation was investigated using the pEGFP-K7 construct.

The sub-cellular localisation of EGFP or EGFP-K7 protein was visualised by fluorescent microscopy at an excitation wavelength of 465-495 nm, 24 and 48 hours following transfection of the pEGFP or pEGFP-K7 construct into Cos-1 cells. Nuclear morphology was also examined by fluorescent microscopy under the UV wavelength (330-380 nm), following methanol fixation for 2 minutes and Hoechst-33258 DNA staining. pEGFP transfected cells displayed uniform fluorescence which extended throughout all parts of the nucleus and cytoplasm (Figure 5.5), consistent with the observations of others (Russell, 1999). By contrast, K7 displayed two distinct sub-cellular localisations. The first EGFP-K7 localisation pattern, observed in approximately 70% of transfected cells, is illustrated by the fluorescing cell on the left of Figure 5.5 E. Fluorescence was relatively uniform throughout the cytoplasm, extending to the cellular extremities and was excluded from the nucleus. The second EGFP-K7 localisation pattern, observed in approximately 30% of cells and illustrated by the fluorescent cell on the right hand side of Figure 5.5 E, was associated with a punctate distribution, concentrated in the peri-nuclear region. This localisation pattern was reminiscent of centrosome localised proteins such as CG-NAP, a PKN-serine/threonine protein kinase associated protein (Takahashi *et al.*, 1999). Cells demonstrating the peri-nuclear punctate staining pattern of EGFP-K7 (Figure 5.5 F) also exhibited an irregular nuclear morphology.

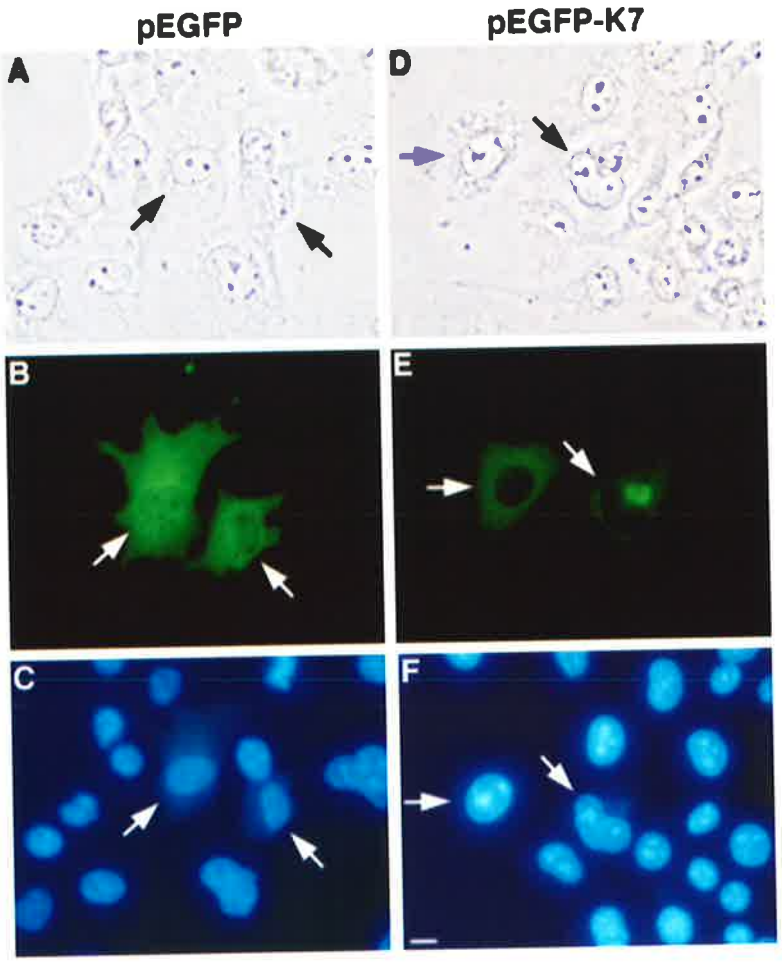
5.2.2.2 Co-localisation of EGFP-K7 and γ-tubulin, a centrosomal marker

Microtubule nucleation is required for formation of the mitotic spindle and occurs by assembly of α- and β-tubulin at microtubule organising centres, such as the centrosome in metazoans, and spindle pole body in yeast (reviewed in Schiebel, 2000). Assembly of microtubules at the centrosome is initiated by complexes containing γ-tubulin, and as a result γ-

Figure 5.5: Sub-cellular localisation of the K7 protein in transfected Cos-1 cells

Localisation of EGFP and EGFP-K7 in Cos-1 cells transfected with 1 µg of pEGFP or pEGFP-K7. pEGFP and pEGFP-K7 transfected cells were viewed 48 hours post-transfection by bright field microscopy (**A**) and (**D**), fluorescent microscopy under the FITC wavelength (excitation 465-495 nm) (**B**) and (**E**), and fluorescent microscopy under the UV wavelength (excitation 330-380 nm) (**C**) and (**F**). The different staining patterns for EGFP-K7 are represented on the left (pattern 1) and right (pattern 2) of the field of view of panel E.

Photography was carried out on a Nikon Eclipse TE300 inverted microscope using the TE-FM Epi-fluorescence attachment. Size bar indicates 10 µm and arrows indicate transfected cells.


pEGFP
pEGFP-K7
A
D
B
E
C
F

tubulin is used as a marker for centrosomes (Schiebel, 2000). The mechanism by which γ-tubulin complexes nucleate microtubules is not clear, however models suggest that γ-tubulin and two associated proteins directly interact with tubulin to mediate nucleation, whilst other ancillary proteins play a structural role or bind γ-tubulin to the centrosome (Zheng *et al.*, 1995; Oegema *et al.* 1999; Schiebel, 2000). In somatic cells, 80% of γ-tubulin is present in the cytoplasm, associated with intermediate filaments, and 20% is associated with the centrosome (Salas, 1999).

Centrosomal localisation of EGFP-K7 in transfected Cos-1 cells was confirmed by co-localisation with γ-tubulin using an anti-γ-tubulin monoclonal antibody and an anti-mouse TRITC-conjugated secondary antibody (2.2.8). Cells were visualised by fluorescent confocal laser scanning microscopy through the equatorial plane (2.3.31). Cells were stained 24 hours post-transfection using two alternate fixation protocols (2.3.31), as detection of protein localisation by immunohistochemistry can be influenced by fixation protocols (Dr. S. Dalton, personal communication). EGFP-K7 protein, uniformly distributed throughout the cytoplasm (pattern 1), did not co-localise with γ-tubulin in methanol fixed cells (Figure 5.6 a-c) or PFA fixed cells (Figure 5.6 j-l). In Cos-1 cells displaying the peri-nuclear, punctate distribution of EGFP-K7, punctate EGFP-K7 protein appeared to co-localise with γ-tubulin at the centrosome, the region from which microtubules emanate, following both methanol and PFA fixation (arrow; Figure 5.6 d-f & m-o). Additional sites of K7 and γ-tubulin co-localisation could be observed in peri-nuclear regions not associated with the centrosome (Figure 5.6 d-f & m-o), however, the significance of these additional sites is not known. EGFP-K7 and γ-tubulin did not co-localise in cytoplasmic regions (Figure 5.6 f). EGFP protein did not co-localise with γ-tubulin in either fixative regime (Figure 5.6 g-i & p-r), and showed no evidence of localisation at the centrosome (arrow; Figure 5.6 q), confirming the specificity of EGFP-K7 localisation at the centrosome.

5.2.3 mPTTG Protein Localisation

The sub-cellular localisation of mPTTG was determined by immunofluorescence (2.3.31), using anti-FLAG monoclonal antibody and an anti-mouse FITC-conjugated secondary antibody (2.2.8), in Cos-1 cell cultures transfected for 24 hours with pmPTTG-FLAG-X and fixed using methanol. mPTTG was found to be localised in both the nucleus and cytoplasm of the cell (Figure 5.7 A), indicating that K7 and mPTTG were present in the same cellular compartment

Figure 5.6: Co-localisation of K7 and γ-tubulin to the centrosome

Cos-1 cells, transfected with 1 μg of either pEGFP-K7 or pEGFP, were co-stained 24 hours post-transfection using an anti-γ-tubulin antibody (Clone GTU-88, Sigma) and an anti-mouse-TRITC conjugated secondary antibody (2.2.8), following fixation with either methanol (**A**) or 3.7% PFA (**B**) (see section 2.3.31). Confocal laser scanning microscopy was carried out through the equatorial plane to visualise protein localisations. For EGFP visualisation, excitation and emission wavelengths were 488/10 nm and 522/35 nm respectively. For detection of TRITC fluorescence, excitation and emission wavelengths were 568/10 and 605/32 respectively. Overlaid images were prepared using Confocal Assistant 4.0. Size bars represent 10 μm.

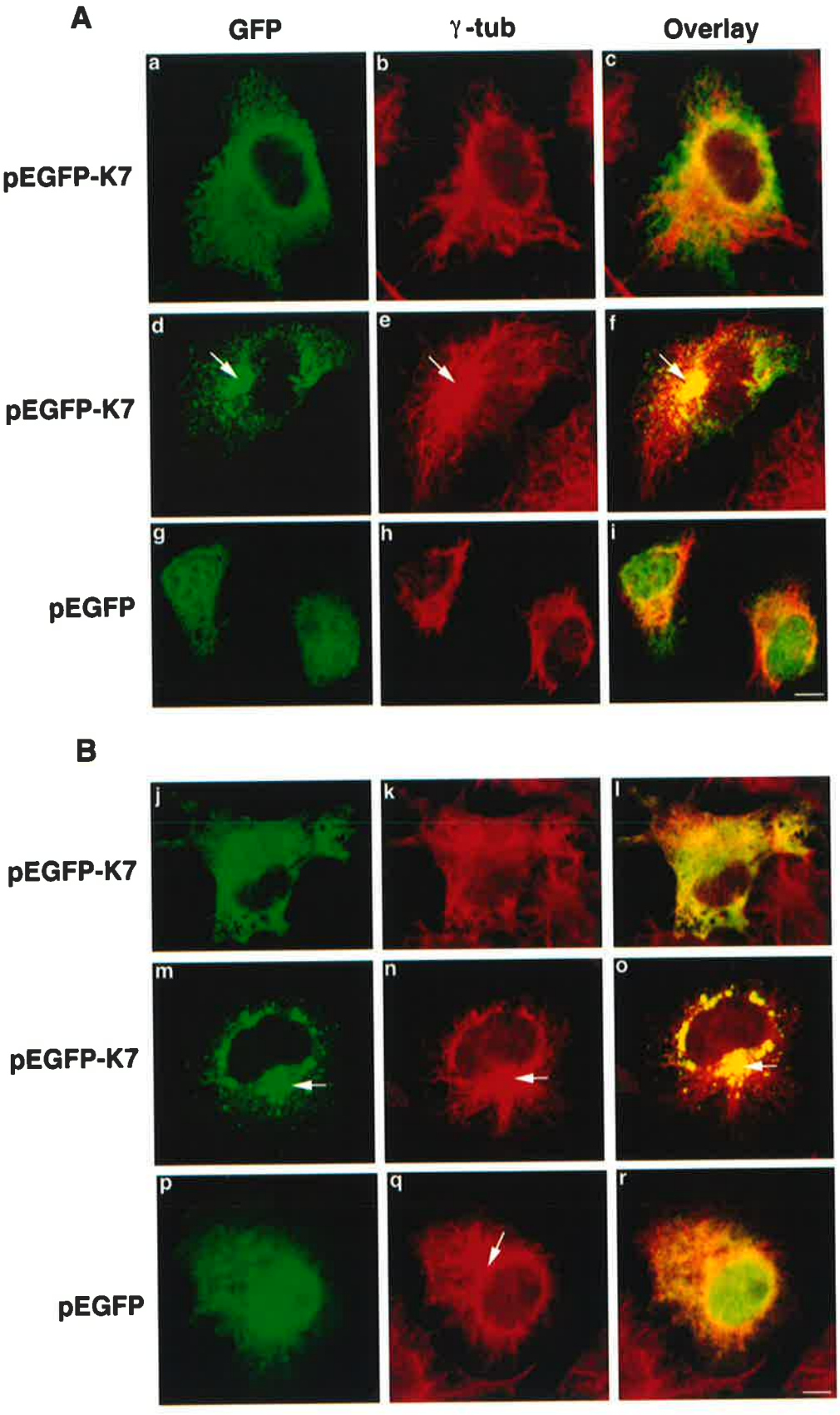
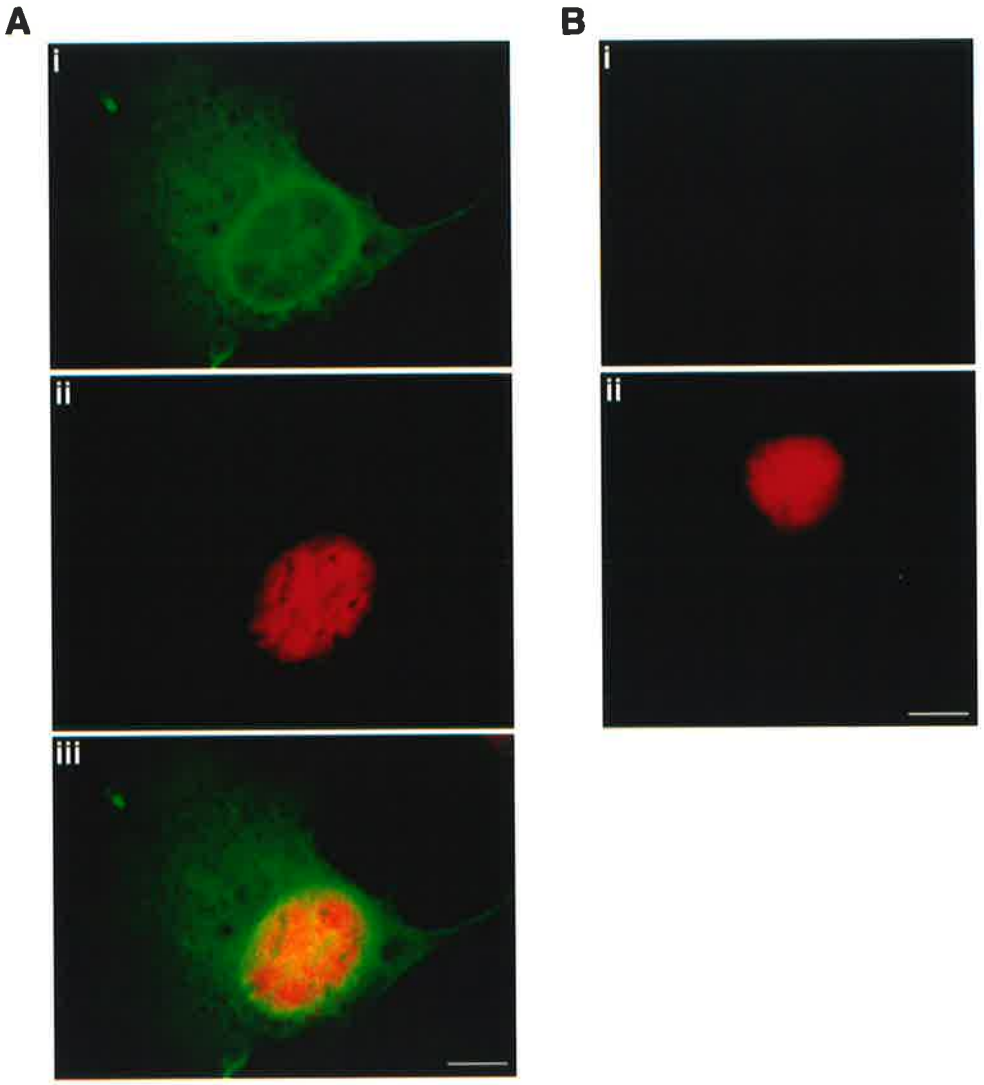

A
GFP
γ-tub
Overlay
pEGFP-K7
pEGFP-K7
pEGFP
B
pEGFP-K7
pEGFP-K7
pEGFP

Figure 5.7: Localisation of mPTTG protein in transfected Cos-1 cells

Localisation of mPTTG in Cos-1 cells transiently transfected with 1 μg of pmPTTG-FLAG-X (**A**) and pXMT2 constructs (**B**). Cells were stained by immunohistochemistry (2.3.31), 24 hours post-transfection, using an anti-FLAG monoclonal antibody (Sigma) and an anti-mouse FITC-conjugated secondary antibody (2.2.8) (i). DNA was stained with propidium iodide to visualise the nucleus (ii). Confocal laser scanning microscopy was carried out through the equatorial plane to visualise protein localisations. For visualisation of FITC fluorescence (i), excitation and emission wavelengths were 488/10 nm and 522/35 nm respectively. For detection of propidium iodide fluorescence (ii), excitation and emission wavelengths were 568/10 and 605/32 respectively. Overlaid images (iii) were prepared using Confocal Assistant 4.0. Size bars represent 10 μm.


A
i
ii
iii
B
i
ii

and capable of interaction. mPTTG was detected at lower levels in the nucleus compared with the cytoplasm and appeared to be concentrated in a ring around the nucleus. Centrosomal staining was not observed. Immuno-staining of Cos-1 cells transfected with pXMT2 alone did not demonstrate fluorescence (Figure 5.7 B), indicating that the staining pattern shown in Figure 5.7 A was a result of pmPTTG-FLAG expression.

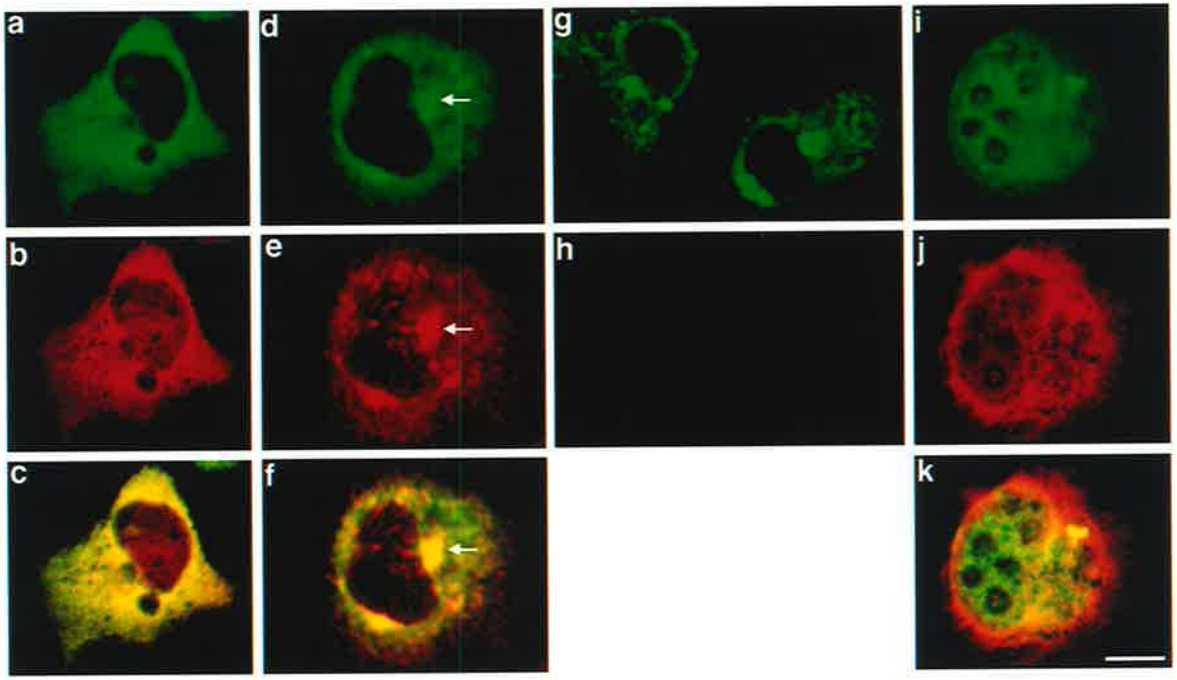
Co-localisation of mPTTG-FLAG and EGFP-K7 proteins was assessed following co-transfection of pmPTTG-FLAG-X and pEGFP-K7 for 24 hours (2.6.7ii). Transfected cells were methanol fixed and analysed by immunohistochemical staining using an anti-FLAG monoclonal antibody and an anti-mouse TRITC-conjugated secondary antibody (2.2.8), followed by confocal laser scanning microscopy (2.3.31). Co-localisation of mPTTG-FLAG and EGFP-K7 proteins was observed when EGFP-K7 was distributed throughout the cytoplasm (pattern 1; Figure 5.8 a-c) and when EGFP-K7 was localised in a punctate peri-nuclear manner (pattern 2; Figure 5.8 d-f). Re-distribution of the uniform mPTTG localisation to a peri-nuclear punctate staining pattern by expression with EGFP-K7 (Figure 5.8 d-f) suggested that association of mPTTG with K7 may be required for mPTTG to be localised to the centrosome.

5.2.4 Overexpression of mPTTG Blocks Cells in G2/M phase

K7 was shown to interact with mPTTG, a destruction box-containing protein which in *Xenopus* was specifically degraded during late mitosis coincident with cyclin B degradation. Furthermore, addition of a non-degradable mutant PTTG protein to metaphase arrested *Xenopus* egg extracts inhibited chromosome segregation (Zou *et al.*, 1999). K7 also shared C-terminal sequence homology with the C-termini of Cut1 and Esp1 (4.2.3.2), yeast proteins which are regulated in a negative manner by the destruction box-containing proteins, Cut2 and Pds1 respectively (Funabiki *et al.*, 1996b; Ciosk *et al.*, 1998). Overexpression of destruction box-containing fragments of Cut2 in *S. pombe* arrests cells in metaphase by preventing both chromosome segregation and inactivation of cyclin B activity, via sequestration of components of the APC ubiquitination pathway (Funabiki *et al.*, 1997). If mPTTG binding to K7 was playing a negative regulatory role, similar to Cut2, overexpression of mPTTG in mammalian cells may inhibit mitotic exit.

Figure 5.8: Co-localisation of mPTTG and K7 protein in transfected Cos-1 cells

Co-localisation of mPTTG and EGFP-K7 in Cos-1 cells, transfected with 1 µg of pmPTTG-FLAG-X and pEGFP-K7 (**a-f**), pXMT2 and pEGFP-K7 (**g,h**), or pmPTTG-FLAG-X and pEGFP (**i-k**). mPTTG protein was detected by immunohistochemistry (2.3.31), 24 hours post-transfection, using an anti-FLAG monoclonal antibody, diluted 1/200, and an anti-mouse TRITC-conjugated secondary antibody (2.2.8). Confocal laser scanning microscopy was carried out through the equatorial plane to visualise protein localisations. For visualisation of EGFP fluorescence (**a,d,g,i**), excitation and emission wavelengths were 488/10 nm and 522/35 nm respectively. For detection of TRITC fluorescence (**b,e,h,j**), excitation and emission wavelengths were 568/10 and 605/32 respectively. Overlaid images (**c,f,k**) were prepared using Confocal Assistant 4.0. Size bar represents 10 µm.


a
d
g
i
b
e
h
j
c
f
k

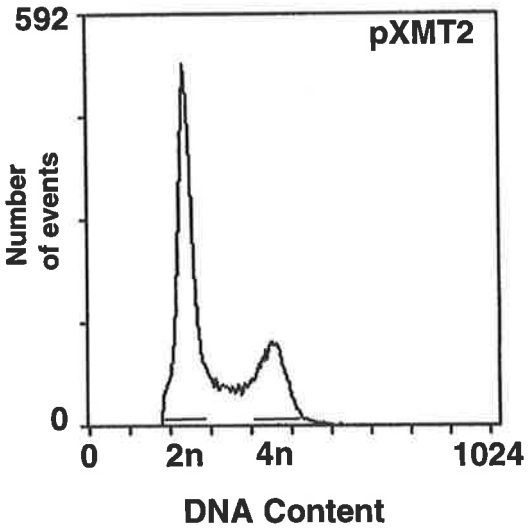
Cos-1 cells were transfected with pXMT2 or pmPTTG-FLAG-X for 30 hours (2.6.7ii), fixed and the DNA was stained using propidium iodide (2.3.32.1). The DNA content of the entire population of Cos-1 cells was analysed by flow cytometry by measurement of propidium iodide fluorescence using the FL3 filter. A flow cytometry profile of Cos-1 cells transfected with pXMT2 alone is depicted in Figure 5.9 A. A DNA content of 2n represents cells in the G1 phase of the cell cycle (56%), a DNA content between 2n and 4n represents S phase cells (22.4%) and a 4n DNA content represents cells in the G2/M phases (20.2%). This profile was typical of the cell cycle characteristics for a population of differentiated cells grown in culture (E. Stead, personal communication). Transfection of Cos-1 cells with pmPTTG-FLAG-X resulted in an increase in the total number of cells in the G2/M phase to 28%, an increase in the number of cells in S phase (27.8%) and a decrease in the total number of cells in G1 phase (40.3%) (Figure 5.9 B).

Analysis of the cell cycle characteristics of the mPTTG-FLAG transfected cells within the population was carried out to determine whether mPTTG-FLAG expression enriched the proportion of cells in G2/M (2.3.32.2). Cos-1 cells were transfected with pmPTTG-FLAG-X or pXMT2 for 30 hours, harvested, fixed in ethanol, and incubated overnight with anti-FLAG monoclonal antibody (2.2.8). Following incubation for 1 hour with an anti-mouse FITC-conjugated secondary, the DNA was stained using propidium iodide. Flow cytometric analysis was then carried out to count the DNA content of FITC stained cells, detected using the FL1 filter. Transfection with pXMT2 and staining with anti-FLAG and the FITC-conjugated secondary antibody determined that non-specific background staining was observed in Cos-1 cells using the anti-FLAG antibody, at 10 units on the FL1 log scale (Figure 5.10 A). In addition to the fluorescent peak at 10 units of the FL1 log scale attributed to background, cells transfected with pmPTTG-FLAG-X produced a higher intensity peak of fluorescence at 100 units on the FL1 log scale (Figure 5.10 B). This FITC peak was not observed in Cos-1 cells overexpressing pmPTTG-FLAG-X which had been incubated with the FITC-conjugated secondary antibody (Figure 5.10 C), indicating that it represented Cos-1 cells expressing the mPTTG-FLAG fusion protein. Analysis of the DNA content of this transfected population (Figure 5.10 J) demonstrated that 69.1% of Cos-1 cells expressing mPTTG-FLAG were in the G2/M cell cycle phases (Figure 5.10 B). This indicated that the increase in the proportion of cells in the G2/M cell cycle phases

Figure 5.9: Overexpression of mPTTG increases the number of cells in the G2/M phase of the cell cycle

Flow cytometric analysis of a population of Cos-1 cells transfected with pXMT2 (**A**), and pmPTTG-FLAG-X (**B**) for 30 hours. Cells were prepared for flow cytometric analysis following fixation and DNA staining using propidium iodide (2.3.32.1). Each profile represents the proportion of cells in each stage of the cell cycle according to the relative DNA content, determined by measurement of incorporated propidium iodide. Percentages of cells in different cell cycle stages were calculated by volume integration using algorithms incorporated into the flow cytometer XL program. The position of cells containing 2n and 4n DNA contents are indicated. Cells in G1 phase have a 2n DNA content, cells in G2/M phase have a 4n DNA content and cells in S phase have an intermediate DNA content. Cells with a sub-2n DNA content represent apoptotic cells.

A


592
pXMT2
Number of events
0
0
2n
4n
1024
DNA Content

B

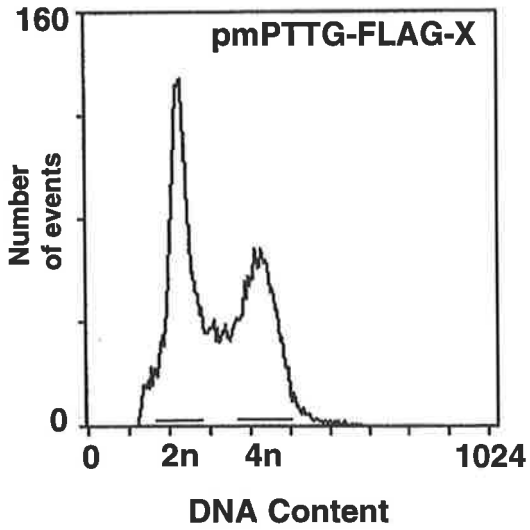

160
pmPTTG-FLAG-X
Number of events
0
0
2n
4n
1024
DNA Content

Figure 5.10: Cells transfected with mPTTG arrest in G2/M phases of the cell cycle

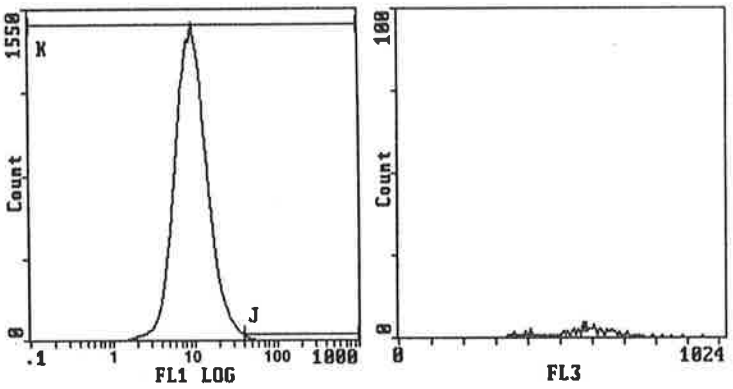
Flow cytometric analysis of Cos-1 cells transfected with pXMT2 (**A**) or pmPTTG-FLAG-X (**B,C**), gated for transfected cells stained with an anti-FLAG monoclonal antibody (2.3.32.2). Two 10 cm plates of Cos-1 cells were transfected (2.6.7ii) with either construct for 30 hours, fixed, permealised and incubated with anti-FLAG antibody, diluted 1/400, and an anti-mouse FITC-conjugated secondary antibody (2.2.8). The DNA was then stained using propidium iodide (PI). FITC and PI fluorescent signals were detected simultaneously using the FL1 and FL3 filters respectively. Detection of FITC fluorescence in cells transfected with pXMT2, following incubation with both primary and secondary antibodies (**A**), determined the background staining of the anti-FLAG antibody in Cos-1 cells. Detection of FITC fluorescence in Cos-1 cells transfected with pmPTTG-FLAG-X and stained with the secondary antibody alone (**C**), determined the background staining of the secondary antibody. Detection of FITC fluorescence in Cos-1 cells transfected with pmPTTG-FLAG-X and incubated with primary and secondary antibodies (**B**), allowed visualisation of the mPTTG-FLAG expressing cells within the culture. Gating for this population (**J**) and plotting its DNA content allowed the cell cycle characteristics of the mPTTG-FLAG transfected cells within the culture to be determined. Proportions of cells in each cell cycle phase were calculated using algorithms incorporated into the flow cytometer XL program.

FITC Fluorescence of entire population

DNA content of Transfected cells

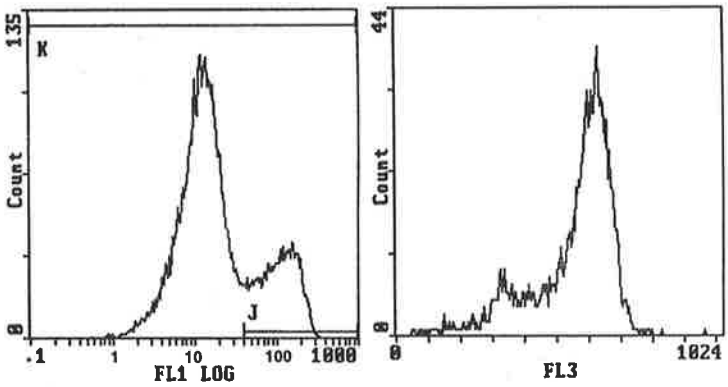
A

pXMT2
+ primary
+secondary



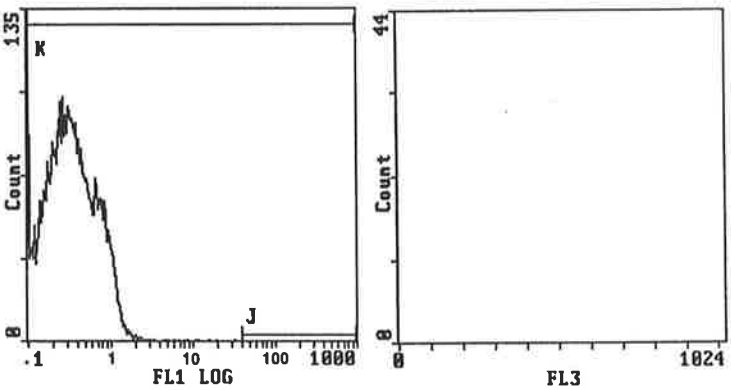
B

pmPTTG-FLAG-X
+ primary
+ secondary



C

pmPTTG-FLAG-X
No primary
+ secondary



within the entire population cells transfected with pmPTTG-FLAG-X (Figure 5.9 B) was a result of mPTTG-FLAG expressing cells being arrested in G2/M.

5.2.5 Analysis of K7 Overexpression on Cell Cycle Progression

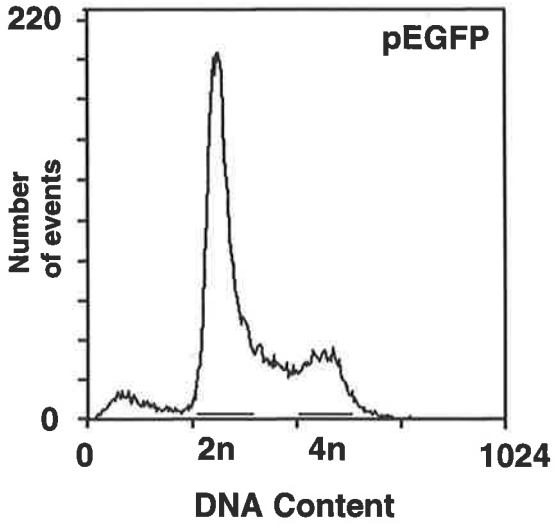
Overexpression of mPTTG inhibited cell cycle progression through the G2/M phases. Therefore, since mPTTG interacted with K7, a potential effect of K7 overexpression in mammalian cells on the cell cycle was also investigated. Cos-1 cells were transfected with pEGFP or pEGFP-K7 for 30 hours (2.6.7ii), fixed and the DNA was stained using propidium iodide. The DNA content of the Cos-1 cells was analysed by flow cytometry using the FL3 filter to detect propidium iodide fluorescence (2.3.32.1), allowing determination of the cell cycle characteristics of the entire population. The flow cytometry profile obtained from transfection of Cos-1 cells with pEGFP is depicted in Figure 5.11 A. 48.3% of cells were in G1 phase, 27.3% of cells were in S phase and 15.7% of cells were in the G2/M phase. Transfection of Cos-1 cells with pEGFP-K7 resulted in an increase in the total number of cells in the G2/M phase (24.8%), a similar proportion of cells in S phase (25.9%) and a decrease in the total number of cells in G1 phase (41.4%) (Figure 5.11 B).

To determine whether Cos-1 cells transfected with pEGFP-K7 had an increased proportion of cells in G2/M, the DNA content of the transfected cells within the population was determined. Following transfection of pEGFP or pEGFP-K7 for 30 hours (2.6.7ii), cells were harvested and fixed in 80% ethanol overnight. The DNA was stained with propidium iodide and the cell cycle characteristics of Cos-1 cells overexpressing EGFP or EGFP-K7 within the population were determined by flow cytometric analysis using the FL1 filter (2.3.32.2). For EGFP, fluorescence was detected just above 10 units on the FL1 log scale and a sub-peak was detected at 100 units on the FL1 log scale (Figure 5.12 A). Two peaks of fluorescence were observed following overexpression of EGFP-K7, the first just under 1 unit on the FL1 log scale and the second at approximately 100 units on the FL1 log scale (Figure 5.12 B). Minimal fluorescence was observed in the untransfected Cos-1 cell control, indicating that these cells have a low level of auto-fluorescence and that the fluorescent peaks observed were a result of EGFP expression within the cells (Figure 5.12 C). Difference in the fluorescent levels observed between EGFP and EGFP-K7, and within the K7 transfected cells, may reflect variability of fluorescence

Figure 5.11: Overexpression of EGFP-K7 increases the number of cells in the G2/M phase of the cell cycle

Flow cytometric analysis of a population of Cos-1 cells transfected with pEGFP (**A**) or pEGFP-K7 (**B**) for 30 hours. Cells were prepared for flow cytometric analysis following fixation and DNA was stained using propidium iodide (2.3.32). Each profile represents the proportions of cells in each stage of the cell cycle according to the relative DNA content determined by measurement of incorporated propidium iodide. Percentages of cells in different cell cycle stages were calculated by volume integration using algorithms incorporated into the flow cytometer XL program. The position of cells containing 2n and 4n DNA contents are indicated. Cells in G1 phase have a 2n DNA content, cells in G2/M phase have a 4n DNA content and cells in S phase have an intermediate DNA content. Cells with a sub-2n DNA content represent apoptotic cells.

A


220
pEGFP
Number of events
0
0
2n
4n
1024
DNA Content

B

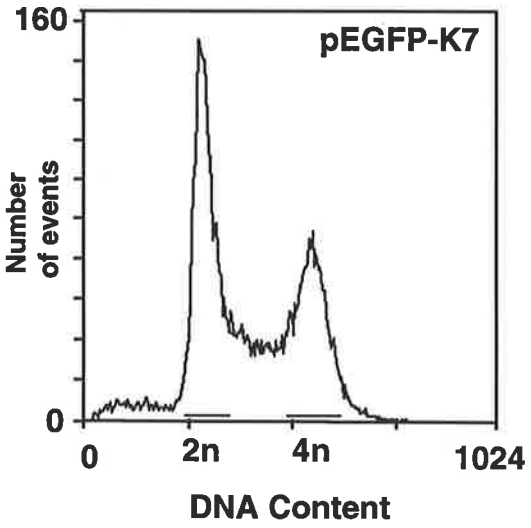
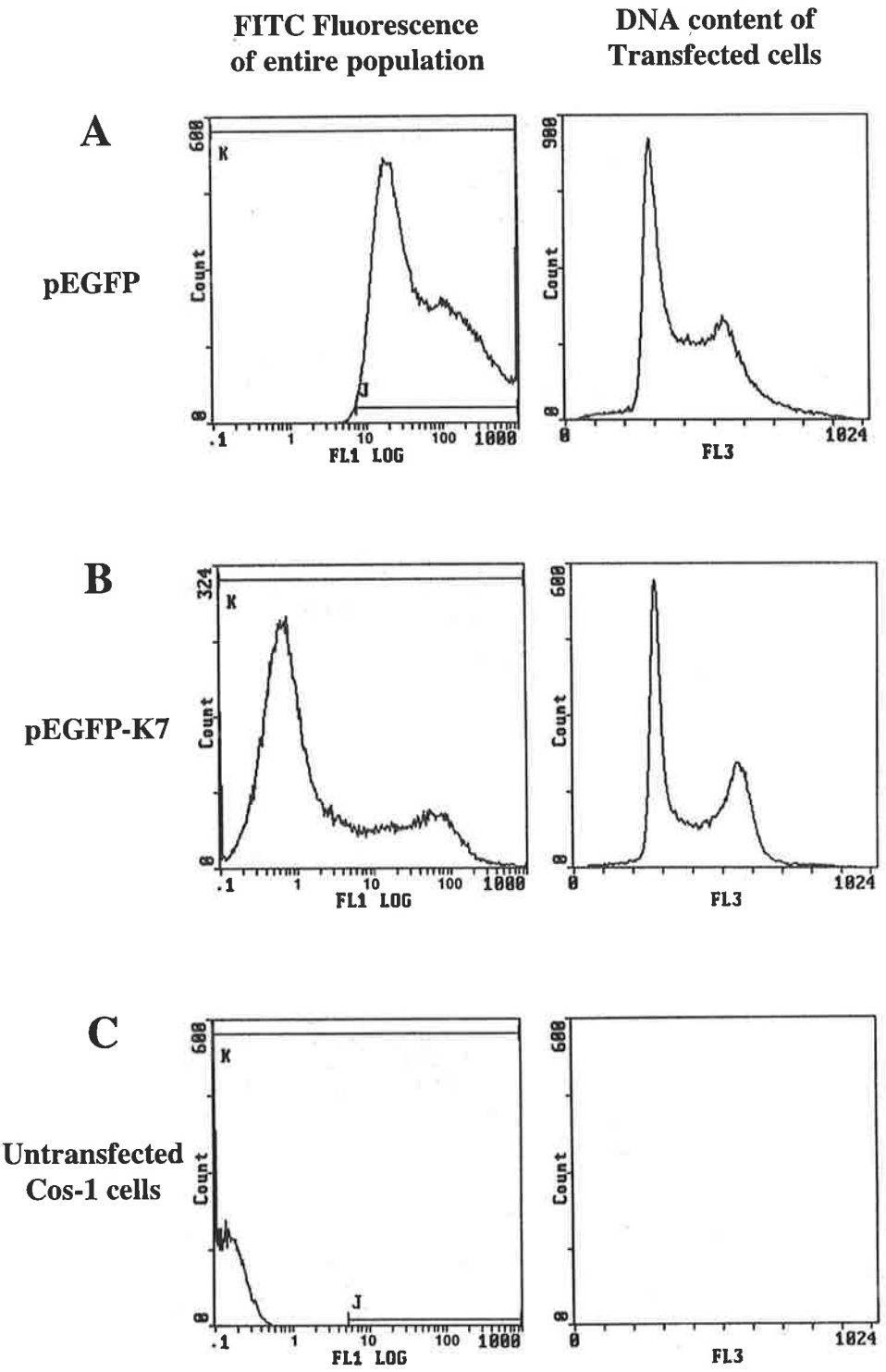

160
pEGFP-K7
Number of events
0
0
2n
4n
1024
DNA Content

Figure 5.12: Cell cycle analysis of Cos-1 cells expressing pEGFP-K7

Flow cytometric analysis of EGFP expressing Cos-1 cells, following transfection with pEGFP (**A**) or pEGFP-K7 (**B**). Two 10 cm plates of Cos-1 cells were transfected (2.6.7ii) with either construct for 30 hours and fixed in 80% ethanol overnight. The DNA was then stained using propidium iodide (PI) (2.3.32.2). Two 10 cm plates of similar density Cos-1 cells were harvested, fixed in 80% ethanol overnight and stained with PI to provide a Cos-1 cell auto-fluorescence control (**C**). EGFP and PI fluorescent signals were detected simultaneously using the FL1 and FL3 filters respectively. The DNA content of the EGFP expressing population (**J**) was graphed and allowed the cell cycle characteristics of the pEGFP transfected cells within the culture to be determined. The DNA content of the EGFP-K7 expressing population (**K**) was graphed and allowed the cell cycle characteristics of the pEGFP-K7 transfected cells within the culture to be determined. Proportions of cells in each phase of the cell cycle was calculated with the use of algorithms incorporated into the flow cytometer XL program.


FITC Fluorescence
of entire population
DNA content of
Transfected cells
A
pEGFP
B
pEGFP-K7
C
Untransfected
Cos-1 cells
Count
FL1 LOG
FL3
K
J

intensity following overexpression of these constructs in culture. EGFP expressing cells and pEGFP-K7 transfected Cos-1 cells demonstrating the punctate, peri-nuclear staining pattern fluoresced more brightly when observed by fluorescence microscopy than EGFP-K7 expressing cells in which demonstrated diffuse cytoplasmic staining (data not shown).

The DNA content of all cells expressing pEGFP (Figure 5.12 J) or pEGFP-K7 (Figure 5.12 K) was analysed. The cell cycle characteristics of the EGFP expressing populations were 41% in G1 phase, 25.7% in S phase and 21.8% in G2/M phases. The cell cycle characteristics of Cos-1 cells overexpressing EGFP-K7 were 39% in G1 phase, 24.9% in S phase and 28.8% in G2/M phase. Therefore, a 7% increase in the proportion of G2/M cells expressing EGFP-K7 compared with EGFP was observed, suggesting that overexpression of EGFP-K7 in Cos-1 cells did not have a significant effect on progression through the G2/M phase of the cell cycle.

5.3 *ESP1* AND K7 C-TERMINAL SEQUENCES ARE NOT FUNCTIONALLY INTER-CHANGEABLE

5.3.1 Rationale

Mutation of *bimB*, *cut1* and *ESP1* resulted in an uncoupling of mitosis and DNA replication due to a failure to segregate sister chromatids at the onset of anaphase (Baum *et al.*, 1988; Uzawa *et al.*, 1990; McGrew *et al.*, 1992; May *et al.*, 1992). The homology between these genes was confined to the C-terminal regions of the encoded proteins, suggesting that this region conferred functional importance, perhaps in comprising a proteolytic active site of a cysteine endopeptidase which mediates the proteolytic cleavage of cohesin subunits (Uhlmann *et al.*, 1999; Nasmyth *et al.*, 2000). The N-terminus and central region of the proteins have been implicated in mediating interactions with diverse regions of the mitotic spindle and spindle pole body, and in mediating association with negative regulatory proteins (Funabiki *et al.*, 1996a). Even though *cut1* and *ESP1* demonstrate almost identical phenotypes when mutated in their respective yeasts, *cut1* was unable to rescue the *ESP1* phenotype when expressed in an *ESP1* temperature sensitive mutant strain (Uzawa *et al.*, 1990). In hindsight this was not surprising given that interaction of these proteins with non-conserved negative regulators is functionally important for correct sub-cellular localisation of Cut1 and Esp1 to the mitotic spindle and spindle pole body during early metaphase (Funabiki *et al.*, 1996a; Kumada *et al.*, 1998; Ciosk *et al.*,

1998). Experiments using only the C-terminal sequence of *cut1*, exchanged for the homologous region of *ESP1*, have not been reported.

K7 shared similar homology in the C-terminal region to *bimB*, *cut1* and *ESP1* gene products, suggesting that the K7 protein may also act to mediate cleavage of cohesin subunits. Since the N-terminal and central regions of the proteins had been implicated in organism specific interactions with regulatory proteins important for the correct cellular distribution of these proteins, a chimeric protein was constructed in which the C-terminal region of K7 was fused to the non-conserved N-terminus and central region of Esp1. This construct was then expressed in an *ESP1* temperature sensitive (*ESP1*^{ts}) *S. cerevisiae* mutant strain in order to determine if the C-terminal region of K7 could functionally replace the equivalent region in *ESP1*.

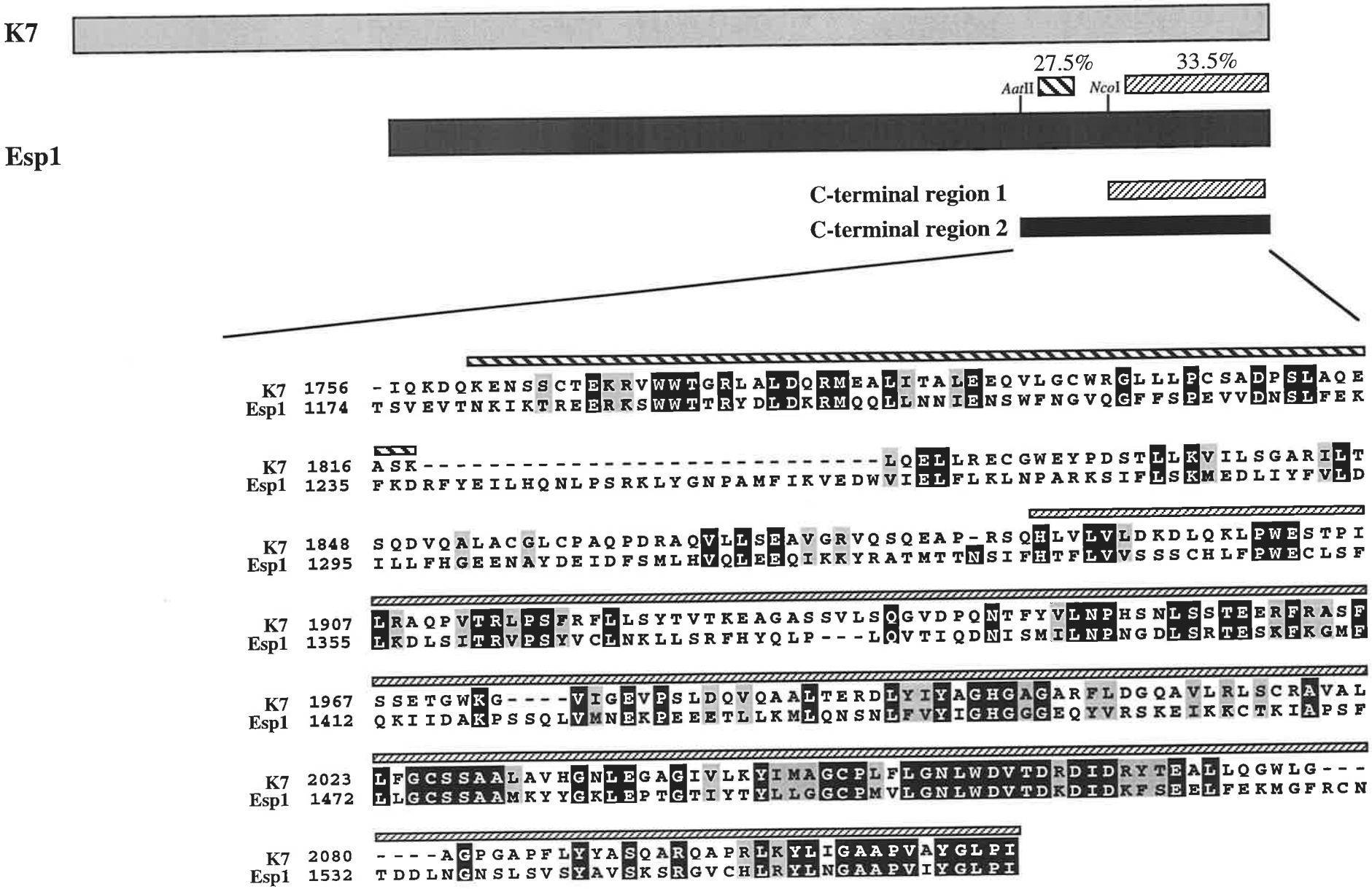
5.3.2 Substitution of Two C-terminal *K7* Sequences for Corresponding C-terminal Sequences of *ESP1* and Expression in S252 *ESP1*^{ts} Strain does not Rescue the Mutant Phenotype

Homology searches using the BlastP program indicated that Esp1 shared similarity with K7 in two C-terminal regions. C-terminal region 1, from amino acid 1886 to 2114 of the K7 protein was 33.5% similar to the Esp1 protein sequence (Figure 5.13). A short region of K7, from amino acid position 1764 to 1818, also displayed 27.5% identity with Esp1 (Figure 5.13). Therefore, a second, slightly longer K7 C-terminal region (C-terminal region 2), extending from the 27.5% homologous region to the end of the K7 protein was also analysed. Unique *Nco*I and *Aat*II sites within the *ESP1* nucleotide sequence were found to lie 109 bp and 19 bp upstream of the corresponding *ESP1* C-terminal region 1 and C-terminal region 2 sequences respectively, and were exploited for PCR amplification and exchange of the equivalent regions of K7 into the *ESP1* sequence.

An *ESP1* containing plasmid, obtained from Dr B. Byers (McGrew *et al.*, 1992), was digested with *Bsr*FI to excise the 6.5 kb *ESP1* open reading frame, endfilled and blunt cloned into the *Sma*I site of pBluescript KS to produce the clone termed ESP1pBs-KS. A positive control, termed ESP1pYES2, was generated by ligation of the *Bsr*FI *ESP1* fragment, containing the entire *ESP1* open reading frame, into the endfilled *Hind*III site of the yeast expression vector,

Figure 5.13: Schematic representation of K7 and Esp1 homology in the C-terminus.

K7 shared 33.5% identity and 50% similarity with a 298 amino acid C-terminal region of Esp1 which was designated C-terminal region 1 (light cross-hatch). A second, smaller region of homology (57 amino acids, dark cross-hatch) was identified 68 amino acids upstream of C-terminal region 1, which shared 27.5% identity and 36% similarity with Esp1. C-terminal region 2 therefore began from the start of this second region of homology and extended to the end of the protein. The K7 C-terminal sequence, from amino acid 1756, has been aligned with the corresponding C-terminal region of Esp1. Dark shading represents identical residues, while light shading represents similar amino acids. The dark cross-hatch above the alignment indicates the position of the 27.5% homologous region and the lighter cross-hatch represents C-terminal region 1.


K7
Esp1
27.5%
33.5%
AatII
NcoI
C-terminal region 1
C-terminal region 2
K7 1756 -IQKDQKENSSCTEKRVWWTGRLALDQRMEALITALEEQVLGCWRGLLLPCSADPSLAQE
Esp1 1174 TSVEVTNKIKTREERKSWWTTRYDLLDKRMQQLLNNIENSWFNGVQGFFSPEVVDNSLFEK
K7 1816 ASK----------LQELLRECGWEYPDSTLLKVILSGARILT
Esp1 1235 FKDRFYEILHQNLPSRKLYGNPAMFIKVEDWVIELFLKLNPARKSIFLSKMEDLIYFVLD
K7 1848 SQDVQALACGLCPAQPDRAQVLLSEAVGRVQSQEAP-RSQHLVLVLDKDLQKLPWESTPI
Esp1 1295 ILLFHGEENAYDEIDFSMLHVQLEEQIKKYRATMTTNSIFHTFLVVSSSCHLFPWECLSF
K7 1907 LRAQPVTRLPSFRFLLSYTVTKEAGASSVLSQGVDPQNTFYVLNPHSNLSSTEERFRASF
Esp1 1355 LKDLSITRVPSYVCLNKLLSRFHYQLP---LQVTIQDNISMILNPNGDLSRTESKFKGMF
K7 1967 SSETGWKG----VIGEVPSLDQVQAALTERDLYIYAGHGAGARFLDGQAVLRLSCRAVAL
Esp1 1412 QKIIDAKPSSQLVMNEKPEEETLLKMLQNSNLFVYIGHGGGEQYVRSKEIKKCTKIAPSF
K7 2023 LFGCSSAALAVHGNLEGAGIVLKYIMAGCPLFLGNLWDVTDRDIDRYTEALLQGWLG---
Esp1 1472 LLGCSSAAMKYYGKLEPTGTIYTYLLGGCPMVLGNLWDVTDKDIDKFSEELFEKMGFRCN
K7 2080 ----AGPGAPFLYYASQARQAPRLKYLIGAAPVAYGLPI
Esp1 1532 TDDLNGNSLSVSYAVSKSRGVCHLRYLNGAAPVIYGLPI

pYES2 (Invitrogen), which directs galactose inducible expression of the protein of interest from the Gal1 promoter, and contains the *Ura3* gene for selection of the plasmid in yeast.

C-terminal region 1 of *K7* was amplified from the partial *K7* cDNA clone, S1pBs-SK (4.2.1), using 5' and 3' PCR primers containing *Nco*I and *Sac*I restriction sites at the 5' and 3' ends respectively (2.2.7.3). The PCR product was digested with *Nco*I/*Sac*I and cloned into *Nco*I/*Sac*I cut ESP1pBs-KS (Figure 5.14 A). The ESP1-K7 fusion fragment was then directionally cloned following digestion with *Kpn*I/*Sac*I, into *Kpn*I/*Sac*I cut pYES2, for galactose inducible expression in yeast. The clone was termed ESP1/*Nco*I-K7pYES2 (Figure 5.14 A).

K7 C-terminal region 2 sequence was PCR amplified from the partial *K7* cDNA clone, S1pBs-SK (4.2.1), using PCR primers containing 5' *Aat*II and 3' *Sac*I restriction sites (2.2.7.3). The PCR product was cloned into ESP1pYES2 following *Aat*II/*Sac*I digestion of both the PCR product and ESP1pYES2 vector (Figure 5.14 B). The clone was termed *ESP1*/*Aat*II-K7pYES2. Both clones were sequenced using 5' ESP1 6175 and 3' K7 5702 sequencing primers (2.2.7.2) to ensure that no errors had been incorporated into the sequence during the PCR reaction and to check that the *ESP1* and *K7* sequences were in frame (data not shown).

A full length *K7* cDNA construct in pYES2 (K7pYES2) was also generated by ligating a *Sac*I/*Eco*RI *K7* cDNA fragment (see Figure 5.3 B) into *Sac*I/*Eco*RI digested pYES2 vector (not shown).

Esp1pYES2, *ESP1*/*Nco*I-K7pYES2, *ESP1*/*Aat*II-K7pYES2 and K7pYES2 were transformed (2.3.26) into the wild type *S. cerevisiae* strain, W303-1a, and a temperature sensitive *ESP1* (*ESP1*^{ts}) mutant strain, S252 (2.2.10). Tranformants were grown using 2% glucose as the carbon source, at the permissive temperature of 25°C for 3 days. Transformants were streaked to single cells on plates containing 2% galactose to induce fusion protein expression from pYES2, and grown at the restrictive temperature of 37°C for 3 days, in order to determine if C-terminal sequences of K7 could rescue the mutant phenotype by allowing growth at 37°C. As expected wild type strains transformed with all plasmids were capable of growth at 37°C (data not shown). Transformation of S252 *ESP1*^{ts} strain with ESP1pYES2 allowed growth at 37°C, indicating that expression of *ESP1* from the pYES2 high copy number plasmid occurred in the presence of galactose, allowing rescue of the *ESP1* temperature sensitive mutant phenotype (Figure 5.15 #3). However, mutant strains transformed with K7pYES2, ESP1/*Aat*II-K7pYES2

Figure 5.14: Construction of Esp1-K7 fusion proteins

A) Construction of a yeast expression plasmid containing C-terminal region 1 of *K7* fused to *ESP1*, from the *Nco*I site in the *ESP1* sequence. The K7 C-terminus was PCR amplified from the S1pBs-SK (refer Figure 4.1), using 5' *Nco*I and 3' *Sac*I primers (2.2.7.3), to engineer an *Nco*I site and *Sac*I site onto the 5' and 3' ends of the *K7* PCR product respectively. The PCR product was then digested *Nco*I and *Sac*I and cloned into *Nco*I/*Sac*I digested ESP1pBs-KS. The resulting ESP1/*Nco*I K7 clone in pBluescript was sub-cloned into the yeast expression vector, pYES2, for protein expression under the control of the Gal1 promoter.

B) Construction of a yeast expression plasmid containing the C-terminal region 2 sequence of *K7* fused to *ESP1* from the *Aat*II site in the *ESP1* sequence. The K7 C-terminal region 2 was PCR amplified from the *K7* cDNA plasmid S1pBs-SK, using 5' *Aat*II and 3' *Sac*I primers (2.2.7.3), to engineer an *Aat*II site and *Sac*I site onto the 5' and 3' ends of the PCR product respectively. Following digestion of the PCR product with *Aat*II and *Sac*I, the K7 C-terminal region 2 fragment was directionally cloned into *Aat*II/*Sac*I digested ESP1pYES2 for protein expression under the control of the Gal1 promoter.

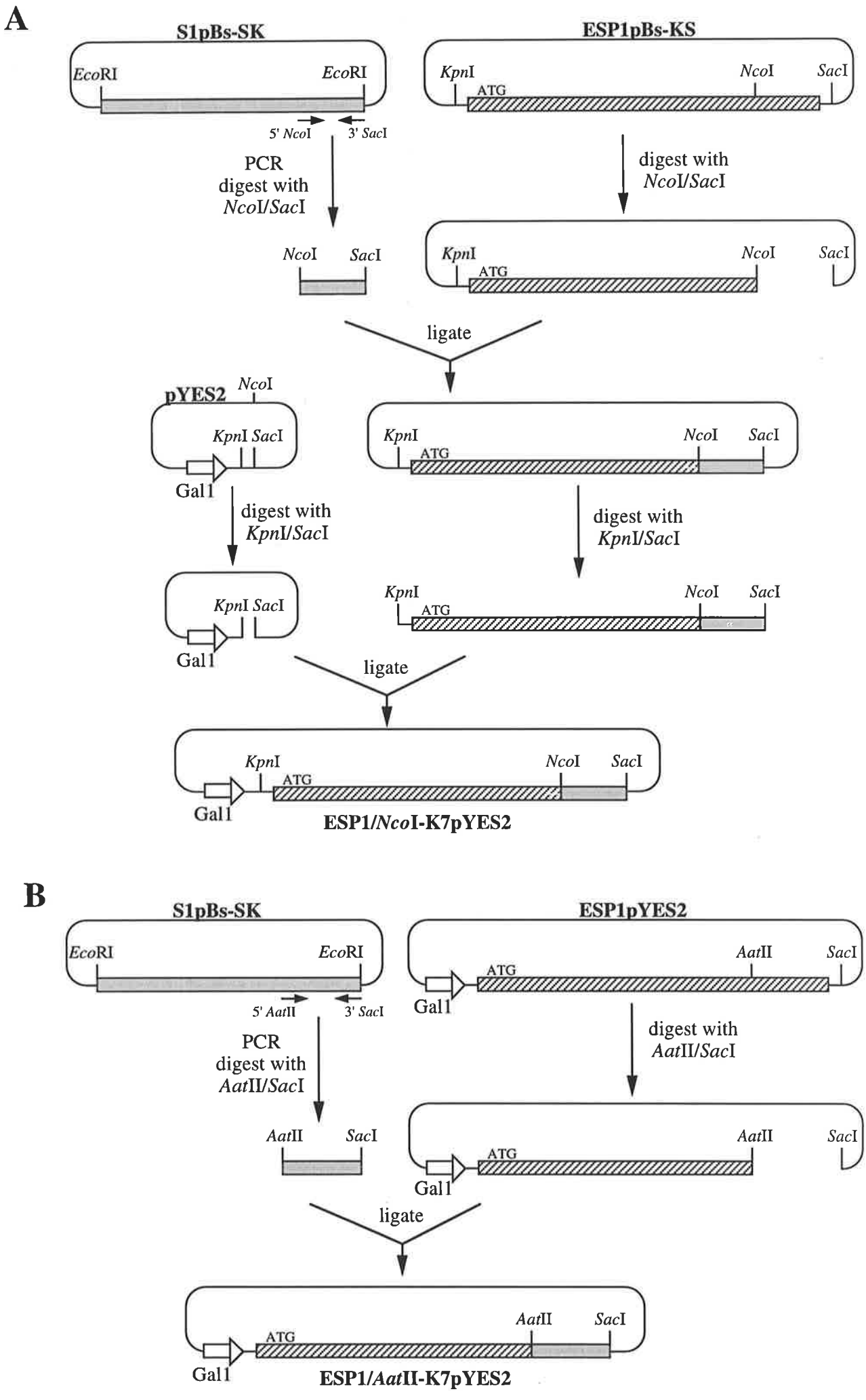
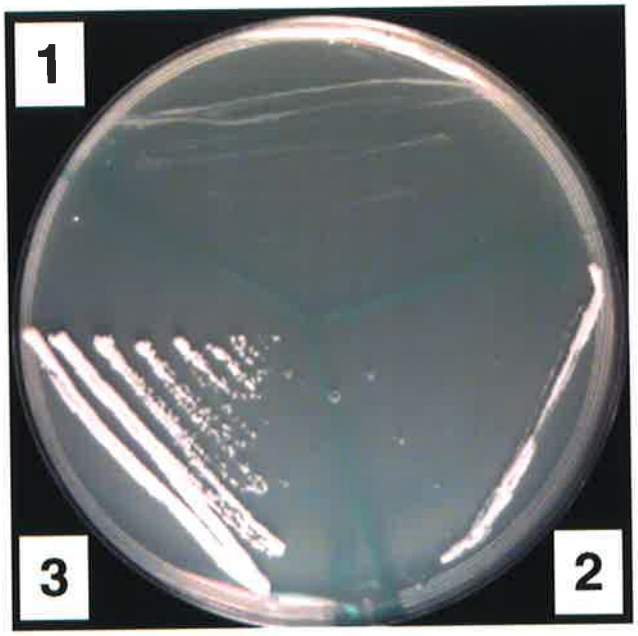
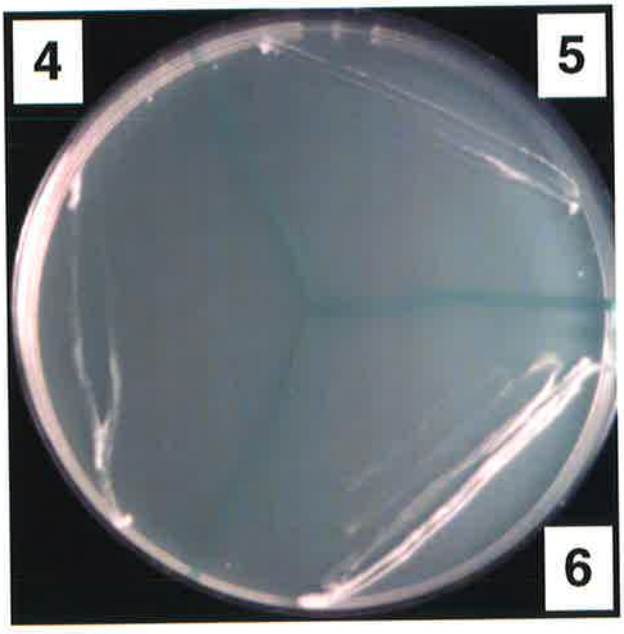

A
S1pBs-SK
EcoRI
EcoRI
5' NcoI
3' SacI
PCR
digest with
NcoI/SacI
NcoI
SacI
ESP1pBs-KS
KpnI
ATG
NcoI
SacI
digest with
NcoI/SacI
KpnI
ATG
NcoI
SacI
ligate
pYES2
NcoI
KpnI SacI
Gal1
digest with
KpnI/SacI
KpnI SacI
Gal1
KpnI
ATG
NcoI
SacI
digest with
KpnI/SacI
KpnI
ATG
NcoI
SacI
ligate
KpnI
ATG
NcoI
SacI
Gal1
ESP1/NcoI-K7pYES2
B
S1pBs-SK
EcoRI
EcoRI
5' AatII
3' SacI
PCR
digest with
AatII/SacI
AatII
SacI
ESP1pYES2
ATG
AatII
SacI
Gal1
digest with
AatII/SacI
AatII
SacI
ATG
AatII
SacI
Gal1
ligate
ATG
AatII
SacI
Gal1
ESP1/AatII-K7pYES2

Figure 5.15: Growth of *ESP1*^{ts} strain, expressing ESP1, K7 and ESP1-K7 fusion proteins, at the restrictive temperature.

Photographs of galactose yeast plates streaked with the *ESP1*^{ts} strain alone (#1) and the *ESP1*^{ts} strain transformed with the yeast expression vector, pYES2 (#2), ESP1pYES2 (#3), ESP1/*Nco*I-K7pYES2 (#4), ESP1/*Aat*II-K7pYES2 (#5) and K7pYES2 (#6). Plates were grown at the restrictive temperature of 37°C for three days.


1
3
2


4
5
6

and ESP1/*Nco*I-K7pYES2 were not capable of supporting growth at 37°C (Figure 5.15), indicating that in this assay the C-terminal region of K7 was not able to substitute for the *ESP1* C-terminus and rescue the *ESP1* mutant phenotype.

5.4 DISCUSSION

5.4.1 Interaction between K7 and mPTTG may Negatively Regulate K7 Activity

The prediction that K7 could play a role during mitosis was postulated in chapter 4 and was supported by several experiments reported in this chapter. K7 was shown to interact with mPTTG by immunoprecipitation, suggesting functional homology with KIAA0165. Similar to Cut1 and Esp1, K7 may therefore be regulated by a destruction box-containing protein. Progression though G2/M was inhibited by overexpression of mPTTG in Cos-1 cells, with 69.1% of cells expressing mPTTG residing in the G2/M population, consistent with a negative regulatory role for PTTG on K7 activity. Saturation of the cell with excess mPTTG protein could prevent efficient ubiquitination of mPTTG by the APC and impede degradation. Thus, residual mPTTG protein bound to K7 may prevent the activation of cohesin subunit cleavage. However, the mitotic block observed in Cos-1 cells overexpressing mPTTG could also reflect saturation of the APC machinery followed by inefficient ubiquitination and degradation of cyclin B, a protein which must be degraded for cells to exit mitosis.

Although overexpression of K7 in Cos-1 cells resulted in a 7-8% increase in the proportion of cells in the G2/M cell cycle phases following analysis of the entire cell population, compared with overexpression of EGFP alone, transfected cells within the culture were not enriched for cells in the G2/M phases. This suggested that transient overexpression of K7 did not have a significant effect on cell cycle progression through G2/M. This was not unexpected given that regulation of K7 activity appeared to occur via the mPTTG protein, which may be present in sufficient quantities to prevent any observable effect of K7 overexpression in this assay.

The increased proportion of cells in the G2/M phases within the entire population of cells overexpressing pEGFP-K7 may reflect high level expression of a large protein (approximately 250 kDa) which can slow growth rates as the cells' metabolic resources become directed toward production of the fusion protein (FuGENE instruction manual, Roche). As cells in culture reach

confluence and become surrounded by other cells, they become quiescent and blocked with a 2n DNA content, a phenomenon termed contact inhibition. Slowing the growth rate by protein overexpression in transfected cells could result in less confluent cell cultures, and as a result, less 2n-containing cells (G1 cells) in the entire population.

5.4.2 Analysis of K7 and mPTTG protein localisations

Although the K7 protein sequence contained consensus sequences for a bipartite nuclear localisation sequence (4.2.2), nuclear localisation was not observed. Instead the K7 protein displayed two distinct localisations. In approximately 70% of transfected cells, K7 was uniformly distributed throughout the cytoplasm, and excluded from the nucleus. In the remaining 30% of transfected cells, K7 displayed a punctate peri-nuclear distribution characteristic of centrosome localisation. Co-staining with γ-tubulin, a centrosomal marker, confirmed this localisation. This distribution was consistent with the sub-cellular localisation of Cut1 to the cytoplasm in interphase cells and Cut1 and Esp1 to the spindle pole body in mitotic cells in *S. pombe* and *S. cerevisiae* respectively (Funabiki *et al.*, 1996a; Kumada *et al.*, 1998; Ciosk *et al.*, 1998), and placed K7 in a location which could enable this protein to perform a function associated with cohesin subunit cleavage.

mPTTG-FLAG protein localisation studies demonstrated diffuse nuclear and cytoplasmic staining. Flow cytometric analysis of these cells demonstrated that the majority of these cells were present in the G2/M cell cycle phases. However, there were no observable condensed chromosomes or metaphase arrested cells, suggesting that the G2/M block may occur in G2 or in early mitosis, prior to chromosome condensation. This was consistent with studies in *Xenopus* which demonstrated loss of 95% of 14S cohesin complex during early mitosis (Losada *et al.*, 1998), implying that the stage at which the K7/mPTTG complex may influence cohesin complex dissociation may occur earlier in mammals than in yeast.

mPTTG protein localisation was redistributed to centrosomes in Cos-1 cells upon co-transfection with pEGFP-K7. This was reminiscent of the situation in fission and budding yeast whereby association of Cut1 and Esp1 with Cut2 and Pds1 respectively was essential for the correct sub-cellular localisation of Cut1 and Esp1 to the spindle pole body and mitotic spindle (Kumada *et al.*, 1998; Ciosk *et al.*, 1998).

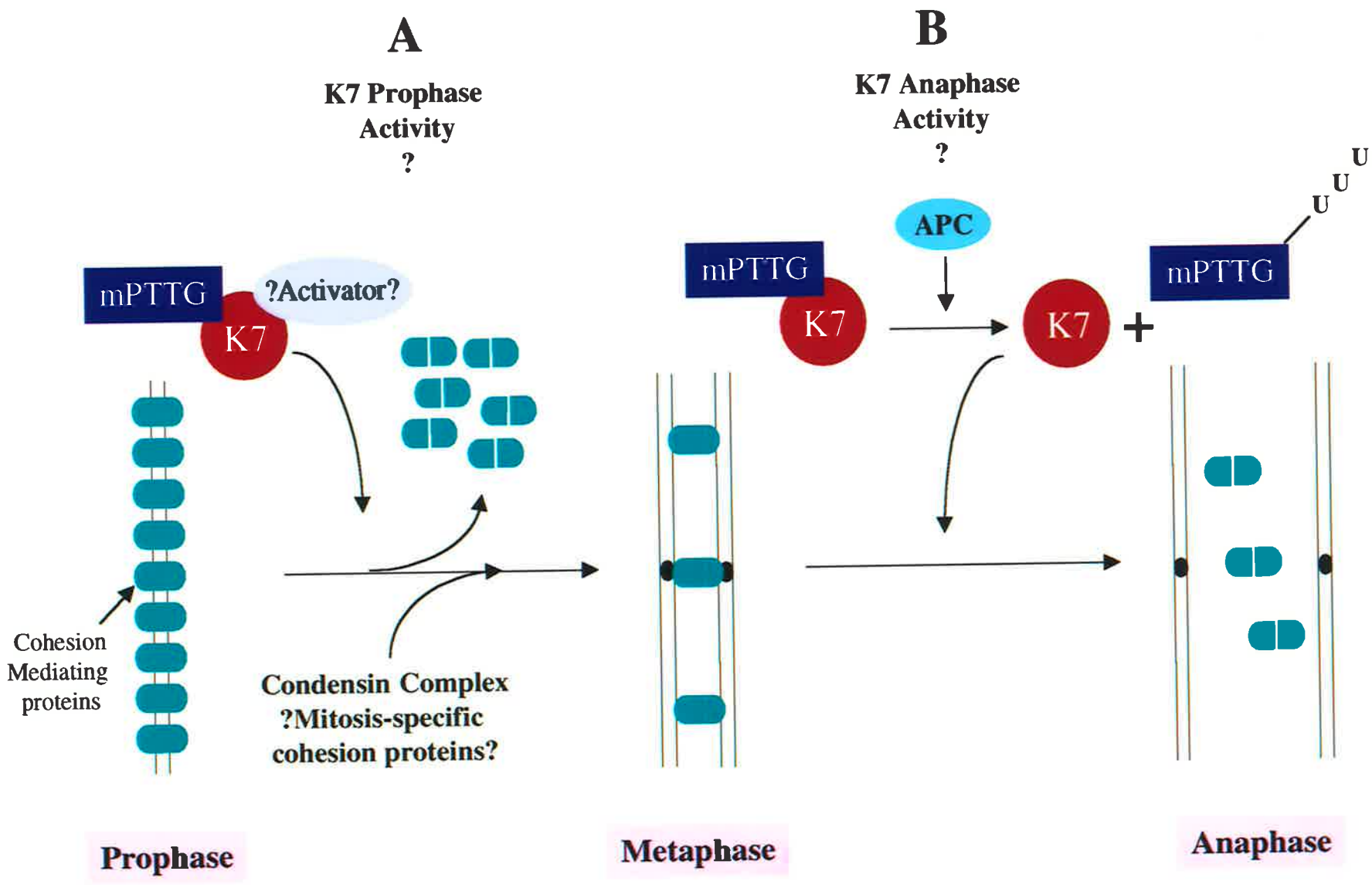
5.4.3 A Role for K7/mPTTG during Early Mitosis in Mammals

Evidence presented in this chapter suggested a role for the K7/mPTTG complex during early mitosis in mammalian cells, indicating that a number of the proteins involved in mediating sister chromatid cohesion and dissociation in yeast and mammals may be conserved. However, the mechanism which controls cohesin complex dissociation is not well understood in higher organisms, and appears to be a more complex process than that described in yeast. In yeast, the cohesin complex has been postulated to mediate both arm and centromere cohesion, thereby acting as the sole regulator of chromosome cohesion during mitosis (Hirano, 1999). Furthermore, the mechanisms of sister chromatid cohesion and chromosome condensation have been postulated to be linked in yeast via the Scc1 protein (Guacci *et al.*, 1997). In contrast, although the 14S cohesin complex appears to be a pre-requisite for correct assembly and architecture required for maintaining sister chromatid association during mitosis, 95% of the complex is lost at the onset of mitosis during prophase in *Xenopus,* prior to association of the 13S condensin complex which mediates chromosome condensation in vertebrates (Losada *et al.*, 1998) (Figure 5.16 A). Therefore, sister chromatid cohesion during mitosis in higher eukaryotes could be mediated by residual cohesin complex subunits (Losada *et al.*, 1998) or by mitosis-specific factors that are distinct from the cohesin subunits identified in yeast (Hirano, 1999). A candidate for a mitosis-specific factor is MEI-S322, a *Drosophila* protein which is recruited to centromeres in prophase of mitosis and becomes displaced during the metaphase to anaphase transition (Moore *et al.*, 1998). However, although *mei-S322* is essential for cohesion between sister chromatid during meiosis, *mei-S332* does not appear essential for mediating sister chromatid associations during mitotic divisions (Kerrebrock *et al.*, 1992; Kerrebrock *et al.*, 1995; Moore *et al.*, 1998), suggesting that other unidentified proteins could mediate centromere cohesion during mitosis in metazoans.

Given the added complexity which is emerging from work in metazoans, the role of cohesin dissociating proteins in mammals analogous to Esp1 and Cut1, such as K7, remains unresolved. K7 could mediate dissociation of the 14S cohesin complex from chromatin at the onset of mitosis, suggesting that K7 functions during early mitosis in mammalian cells (Figure 5.16 A). However, if K7 functions during early mitosis, it is not clear how it might escape negative regulation by mPTTG, as PTTG is not degraded until late mitosis in human cells (Zou *et*

Figure 5.16: Potential action of K7 in mitosis

Schematic representation of a two-step model for the possible roles of K7 during the consecutive loss of cohesion mediating proteins in mammalian cells. 95% of the 14S cohesin complex is lost at the onset of mitosis, during prophase, prior to association of the 13S condesin complex which mediates chromosome condensation (Losada *et al.*, 1998). K7 could act during this stage to mediate cleavage of the cohesin subunits. How K7 might overcome mPTTG regulation to become activated at this stage is not known. Sister chromatid separation occurs at the metaphase to anaphase transition of mitosis, via removal of residual cohesin complex subunits or removal of other proteins which specifically mediate sister chromatid association during mitosis. By analogy with Cut1/Esp1 function (Funabiki *et al.*, 1996a; Uhlmann *et al.*, 1999) K7 could function during this stage to result in cleavage of these proteins and loss of sister chromatid association. mPTTG negative regulation which has been proposed to be present during this stage of mitosis in vertebrate cells could be overcome by degradation of the protein late in mitosis following ubiquitination by the anaphase promoting complex (APC) (Zou *et al.*, 1999).


A
K7 Prophase Activity ?
mPTTG
?Activator?
K7
Cohesion Mediating proteins
Condensin Complex ?Mitosis-specific cohesion proteins?
Prophase
Metaphase
B
K7 Anaphase Activity ?
APC
mPTTG
K7
K7
+
mPTTG
U
U
U
Anaphase

al., 1999). Studies in yeast have implicated Esp1 in mediating cleavage of the Scc1 cohesin complex subunit at the metaphase to anaphase transition (Uhlmann *et al.*, 1999). Furthermore, exogenous addition of a non-degradable mutant PTTG protein to *Xenopus* oocytes synchronised in metaphase inhibited chromosome segregation (Zou *et al.*, 1999). Therefore, K7 could also possibly function during the metaphase to anaphase transition and cleave residual cohesin complex subunits or mitosis-specific proteins to induce chromosome segregation (Figure 5.16 B). mPTTG negative regulation during this stage could be overcome by degradation of the protein late in mitosis following ubiquitination by the anaphase promoting complex (APC) (Zou *et al.*, 1999). However, a late mitotic function could not be determined in the assay used as induction of cell cycle arrest during early mitosis precluded analysis of later mitotic defects.

Therefore, additional regulatory mechanisms which control K7 and/or mPTTG appear likely in mammalian cells. Investigation of the timing of K7 and mPTTG activity during mitosis in mammalian cells and characterisation of the additional proteins involved in sister chromatid cohesion will be required in order to begin to understand the mechanism of cohesin dissociation and chromosome segregation in mammalian cells.

5.4.4 *K7* was unable to Induce Chromosome Segregation in Yeast

Substitution of two C-terminal regions of *K7* for corresponding *ESP1* C-terminal regions did not result in complementation of the $ESP1^{ts}$ mutant phenotype when these constructs were expressed from a high copy number plasmid. This indicated that although substantial sequence homology is present between the two proteins in the C-terminal region, and that aspects of the chromosome segregation mechanism may be conserved between yeast and mammals, the K7 C-terminal region was unable to participate in the analogous pathway in *S. cerevisiae*. The components involved in mediating sister chromatid association may not be sufficiently conserved between yeast and mammals to allow the C-terminal region of K7 to participate in this pathway in *S. cerevisiae*. Furthermore, even though it has been proposed that the conserved C-terminal region of Esp1 may provide the active site of a cysteine endopeptidase which cleaves the cohesin subunit, Scc1 (Uhlmann *et al.*, 1999; Nasmyth *et al.*, 2000), the Scc1 cleavage site identified in *S. cerevisiae* (Uhlmann *et al.*, 1999) is not conserved in the vertebrate Scc1 protein (Glotzer,

1999). Therefore, it is not likely that the C-terminal region of K7 could recognise and cleave Scc1 in *S. cerevisiae*.

CHAPTER 6:

GENERAL DISCUSSION

6.1 INTRODUCTION

The work presented in this thesis was of two parts, firstly it described the use of marker gene expression to define distinct pluripotent cell sub-populations within the *Oct4*$^{+}$ pool of pre- and early post-implantation mouse development. Secondly, one of the marker genes, *K7*, was investigated in more detail for possible function, and a putative role for this gene during mitosis was identified.

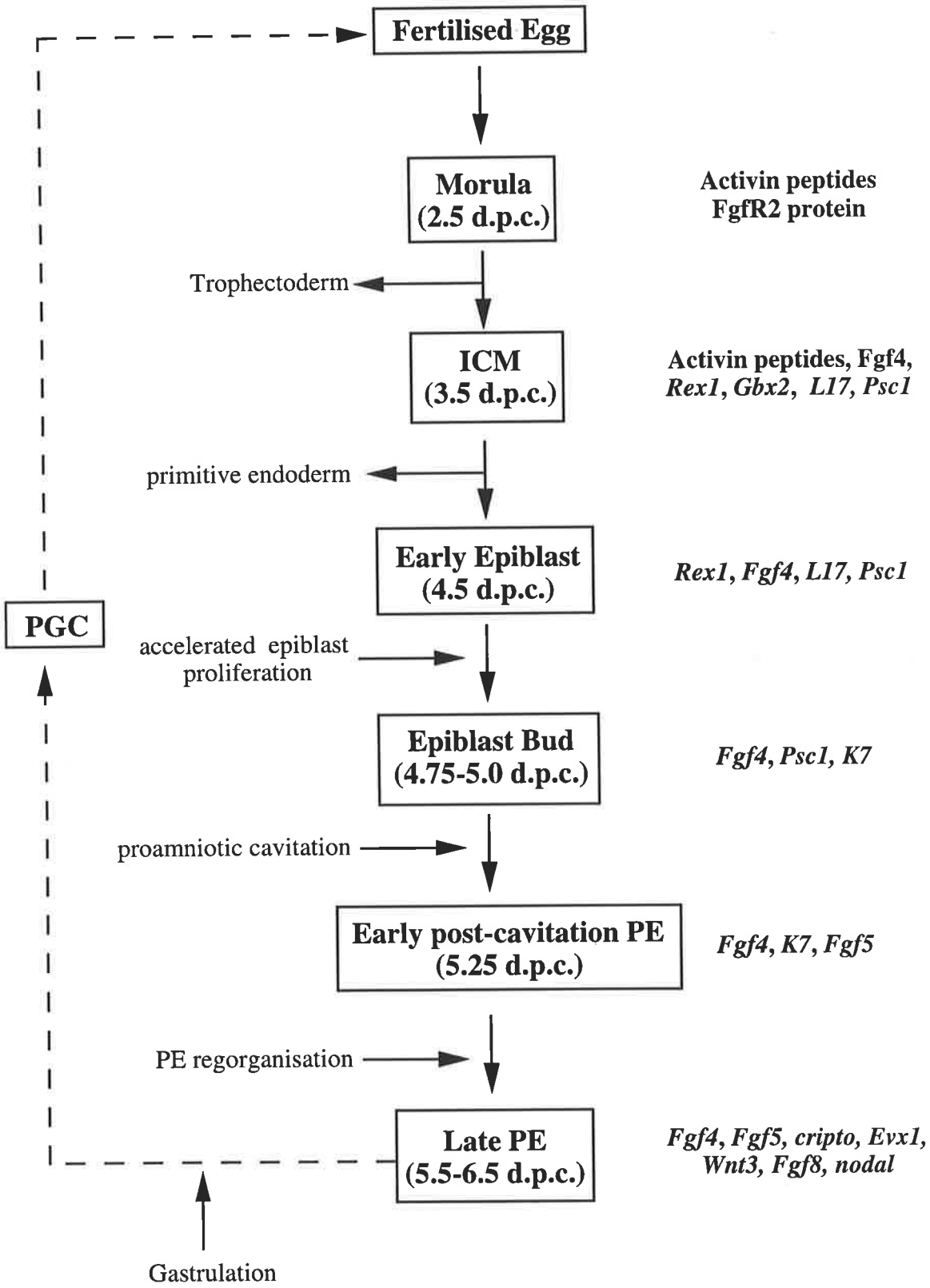
6.2 PLURIPOTENT CELL SUB-POPULATIONS DURING EARLY MOUSE EMBRYOGENESIS

The precise expression mapping of marker genes, isolated during a differential display PCR screen of ES and EPL cell RNA (*L17*, *Psc1* and *K7*), and two markers previously described in the literature (*Rex1* and *Fgf5*), revealed the presence of pluripotent cell sub-populations within the *Oct4*$^{+}$ pool during early mouse embryogenesis. Integration of this gene expression data with expression data of other reported genes and gene products, as determined by whole mount *in situ* hybridisation or whole mount immunohistochemical staining respectively, provided a molecular description of cell progression in the pluripotent lineages (Figure 6.1). As the totipotent morula develops into the blastocyst, the ICM lineage formed can be distinguished from the morula by the loss of FGFR2 protein, which becomes detectable in nascent trophectoderm (Haffner-Krausz *et al.*, 1998), and induction of *Fgf4* transcripts and protein within the ICM at 3.5 d.p.c. (Niswander and Martin, 1992; Rappolee *et al.*, 1994). *Fgf4* expression persists within later pluripotent cell populations until gastrulation (Niswander and Martin, 1992). Progression of the ICM to early epiblast at approximately 4.5 d.p.c. is marked by downregulation of activin peptides within the epiblast, while expression of *Rex1*, *L17, Psc1* and *Fgf4* persists (Albano *et al.*, 1993). The 4.75-5.0 d.p.c. epiblast population has lost *Rex1* and *L17* expression, while expression of *Psc1* is maintained, and *K7* expression is upregulated. Early post-cavitation primitive ectoderm does not express *Psc1* but expresses *K7* and *Fgf5*. Later stage primitive ectoderm (5.75-6.5 d.p.c.) is distinguishable from early primitive ectoderm by the downregulation of *K7* and maintenance of *Fgf5* expression (Haub and Goldfaúb, 1991; Hébert *et al.*, 1991). Regionalised gene expression also becomes apparent in the primitive ectoderm from this stage as *Evx1*, *Wnt3*, *Fgf8* and *nodal* become expressed in the proximal posterior primitive

Figure 6.1: Subclassification of *in vivo* pluripotent cell populations by gene expression

Morphological stages of early mouse pluripotent cell development, as defined in this thesis (3.2.1), are aligned with approximate timing of embryonic stages in days *post coitum* (d.p.c.) and embryological events. Gene expression markers that are restricted during pluripotent cell development, as demonstrated by *in situ* hybridisation, are indicated in italics. Gene expression markers that are restricted during pluripotent cell development, as demonstrated by immunohistochemistry of gene products, are not italicised.

ICM = inner cell mass, PE = primitive ectoderm


Fertilised Egg
Morula
(2.5 d.p.c.)
Activin peptides
FgfR2 protein
Trophectoderm
ICM
(3.5 d.p.c.)
Activin peptides, Fgf4,
Rex1, Gbx2, L17, Psc1
primitive endoderm
Early Epiblast
(4.5 d.p.c.)
Rex1, Fgf4, L17, Psc1
PGC
accelerated epiblast
proliferation
Epiblast Bud
(4.75-5.0 d.p.c.)
Fgf4, Psc1, K7
proamniotic cavitation
Early post-cavitation PE
(5.25 d.p.c.)
Fgf4, K7, Fgf5
PE regorganisation
Late PE
(5.5-6.5 d.p.c.)
Fgf4, Fgf5, cripto, Evx1,
Wnt3, Fgf8, nodal
Gastrulation

ectoderm in an increasingly distal fashion (Dush and Martin, 1992; Crossley and Martin, 1995; Varlet *et al.*, 1997; Liu *et al.*, 1999).

The presence of distinct patterns of gene expression within pluripotent cell sub-populations highlights the complexity of the developmental progression of pluripotent cells throughout early mouse development. These gene expression changes appear to correlate with key developmental events which occur during this stage of development, including blastocyst formation at 3.5 d.p.c., increased cellular proliferation, proposed in this thesis to begin around day 4.75, proamniotic cavitation, which begins at 5.0 d.p.c., and primitive ectoderm reorganisation into an epithelial layer at approximately day 5.5-5.75 of embryogenesis. Therefore, developmental progression of pluripotent cells occurs as a series of regulated stages. Gene expression changes *in vivo* correlated closely with gene expression *in vitro,* during the ES to EPL cell transition. This validated the differentiation of EPL cells from ES cells and their subsequent developmental progression in culture, as a model for progression of ICM to primitive ectoderm *in vivo*. Furthermore, comparison of the *in vitro* and *in vivo* expression data allowed a more precise definition of ES and EPL cells. Therefore, ES cells most likely represent epiblast cells prior to day 4.75 of embryogenesis. EPL cells cultured for 4 days in the presence of LIF or 2 days in the absence of LIF are likely to represent pluripotent cell subtypes equivalent to 4.75 d.p.c. epiblast tissue. EPL cells cultured for 4 days in the absence of LIF or 6 days in the presence of LIF are most closely related to pluripotent cells present at day 5.0 of embryogenesis, while EPL cells grown for 6 days in the absence of LIF or 8 days in the presence of LIF appear related to primitive ectoderm cells at day 5.25.

6.2.1 Applications of Marker Genes to the Analysis of Pluripotent Cell Development

6.2.1.1 Analysis of pluripotent cell development in the mouse

(i) Functional analysis of pluripotent cell developmental progression *in vitro*:

L17 and *Psc1* were expressed in ES cells and early passage EPL cells, but expression was downregulated upon progression to later passage EPL cells. This indicated that these genes may have functional roles during the maintenance/self-renewal ability of ES cells or early passage EPL cells, or in the inhibition of differentiation of ES cells or early passage EPL cells. *K7*

displayed a reciprocal expression pattern to that of *L17* and *Psc1*, with expression upregulating in later passage EPL cells, grown for 4-6 days in culture in the presence of LIF. This was indicative of a potential requirement for *K7* in the progression of early passage EPL cells to later passage EPL cells. Recently, the level of Oct4 protein has been demonstrated to govern the self-renewal capacity and differentiation potential of ES cells (Niwa *et al.*, 2000). A less than two fold increase in *Oct4* expression from baseline levels present in ES cells, which maintains the self-renewal capacity of ES cells, results in differentiation into primitive endoderm or mesoderm, while repression of *Oct4* induces loss of pluripotency and de-differentiation into trophectoderm. Therefore, the functional roles of *L17*, *Psc1* and *K7 in vitro* could be analysed by the creation of null mutant ES cells (Shearwin-Whyatt *et al.*, 1999), or via the generation of ES cell lines overexpressing *L17*, *Psc1* and *K7*. Mutation of these genes in ES cells may allow functional analysis of *L17*, *Psc1* and *K7* during the ES to EPL cell transition and allow investigation into possible functions for these genes in self-renewal and differentiation capabilities of pluripotent cells *in vitro*. However, since these genes are proposed to play a role in ES cell maintenance and/or self-renewal, double knockout ES cells for these genes may not be viable.

Alterations in pluripotent cell progression following either null mutation or overexpression of *L17*, *Psc1* or *K7* in ES cells could be assayed during the ES to EPL cell transition using the marker genes analysed in chapter 3 of this thesis. Self-renewal/proliferation capacity of ES and EPL cells could be analysed using standard proliferation based assays such as the CellTitre 96 nonradioactive cell proliferation assay kit available from Promega or by 5'-bromo-2'-deoxyuridine (BrdU) incorporation. The differentiation capacity of the double knockout ES cells or overexpressing ES cell lines could be assessed using differentiation in monolayer culture or as embryoid bodies. Defects in forming extra-embryonic tissues such as primitive endoderm and derivatives may be possible, as expression of *L17*, *Psc1* and *K7* are normally downregulated in these tissues. Defects in forming primitive ectoderm, which may delay mesoderm induction, if the threshold number of cells required for gastrulation to commence is not produced, may also be encountered. Differentiation dysfunctions could be assayed using marker genes specific for primitive endoderm and derivatives (*Hex* for primitive endoderm, Thomas *et al.*, 1998; *SPARC* for parietal endoderm, Holland *et al.*, 1987; *alphafetoprotein* for visceral endoderm, Dziadek and Adamson, 1978), primitive ectoderm (*Fgf5*, section 3.2.1, Haub

and Goldfarb, 1991; Hébert *et al.*, 1991) and mesoderm (*brachyury*, Herrmann, 1991). Defects in the capability of pluripotent cells to give rise to certain post-gastrulation tissues, such as mesoderm derivatives, including cardiac muscle and blood cells, or ectoderm derivatives, including neurons, could be assayed using differentiation regimes determined in our laboratory, followed by marker gene analysis for the tissue type affected (Remiszweski, 1999; Lake *et al.*, 2000; J. Rathjen, unpublished).

(ii) Functional analysis of pluripotent cell development *in vivo*:

The effect of null mutation of *L17*, *Psc1* and *K7 in vivo* could be determined by standard gene targeting techniques (Shearwin-Whyatt *et al.*, 1999). Analysis of phenotypes in embryos lacking *L17*, *Psc1* or *K7* could reveal critical embryological stages at which these genes are required. Furthermore, ectopic expression of constructs containing *L17*, *Psc1* or *K7* microinjected into fertilised eggs and analysis of developmental phenotypes could provide additional information regarding gene function and complement the null mutational approaches.

L17 was expressed in pluripotent ICM cells of pre-implantation blastocysts, and in the epiblast population following primitive endoderm differentiation, but was downregulated shortly after. *L17* expression was also downregulated in the primitive endoderm. Mutation of the *L17* gene may therefore produce dysfunction in pluripotent cell maintenance, pluripotent cell proliferation, or primitive endoderm differentiation. *Psc1* was expressed in the pluripotent ICM cells of pre-implantation blastocysts, and in the epiblast population until the onset of proamniotic cavitation, but was not expressed in the primitive endoderm. Therefore, pluripotent cell defects that might be expected in *Psc1* mutant embryos could include altered pluripotent cell maintenance or proliferation, dysfunctions in primitive endoderm or primitive ectoderm differentiation, or proamniotic cavitation. *K7* was expressed in the pluripotent cells at the epiblast bud stage and in the newly cavitated primitive ectoderm, but not following primitive ectoderm reorganisation. Targeted disruption or overexpression of this gene might affect the onset of accelerated cellular proliferation, proamniotic cavitation, or primitive ectoderm formation.

Mutant embryos would be analysed initially by morphology to determine defects in forming certain tissues such as primitive endoderm and derivatives, primitive ectoderm or mesoderm, by harvesting embryos at differing stages of pre- and post-implantation development.

Defects would be analysed further using marker genes specific for the affected cell type(s), as described in 6.2.1.1i, or by BrdU incorporation to determine any proliferative defects (Hayashi *et al.*, 1988).

(iii) A molecular approach to the identification of signaling networks that regulate pluripotent cell progression:

Isolation of gene promoters could provide valuable insight into the control of pluripotent cell developmental progression. Promoter regions which regulate *L17*, *Psc1* and *K7* expression could be isolated from genomic clones by standard molecular methodologies. Generation of deletion constructs of promoter regions cloned upstream of a β-galactosidase reporter gene, followed by transfection into ES cells and expression analysis during the ES to EPL cell transition *in vitro*, could be used to determine the sequences functionally important for gene expression and regulation during pluripotent cell progression. To confirm that the promoter regions isolated are sufficient to recapitulate the gene expression observed *in vivo*, transgenic mice expressing β-galactosidase under the control of promoter sequences identified *in vitro* during the ES to EPL cell transition could be produced. Potential transcription factor binding sites could be identified within the promoter regions responsible for ES or EPL specific gene expression, mutated and assayed for effects on expression *in vitro*, using the ES to EPL cell model, and *in vivo*, using transgenic mice.

These analyses may allow the identification of regulatory hierarchies that control pluripotent cell developmental progression *in vitro*. For example, *L17* demonstrated an expression pattern indistinguishable from that of *Rex1*. The *Rex1* promoter contains an octomer sequence which mediates both activation and repression of the promoter depending on the levels of Oct4 protein present, and a novel element, termed Rox1, which reduces *Rex1* expression when mutated (Ben-Shushan *et al.*, 1998). Isolation and functional analysis of the *L17* gene promoter could determine whether *L17* and *Rex1* are regulated coordinately by Oct4, or other transcription factor complexes. This would place *L17*, a member of the CP2 transcription factor family (Rodda, 1998; Scherer, 1999), in a transcriptional regulatory hierarchy containing *Oct4* and *Rex1*. The determination of transcription factors which regulate the expression of *L17*, *Psc1* or *K7* may also allow the signal transduction pathways which control gene regulation to be

deduced, and begin to unravel the signaling networks which control pluripotent cell developmental progression.

6.2.1.2 Relationships between pluripotent cell populations in vivo and in vitro

The relationships between *in vitro* and *in vivo* pluripotent cell populations of mouse and humans, including ES, EC and EG cells, in many instances are not well understood. Expression analysis of marker genes described in this thesis may determine the relationships, if any, between mouse ES, EC and EG cell lines and human ES, EC and EG cell lines. If human ES, EC or EG cell lines are closely related to mouse ES or EPL cells, they may also have a similar developmental potential to mouse ES or EPL cells, and provide candidates for application of lineage specific differentiation protocols generated for mouse ES and EPL cells (Remiszweski, 1999; Lake *et al.*, 2000; J. Rathjen, unpublished). Differentiation of human pluripotent cells *in vitro* into certain post-gastrulation tissues such as neurons may be important for cell based therapeutics (reviewed in Rathjen *et al.*, 1998). Determining the relationship of human EC cell lines with *in vitro* and *in vivo* pluripotent cell populations may also allow the origin of germ cell tumours to be resolved.

Isolation of pluripotent cell lines from species other than mouse would allow application of ES cell technology and precise genome manipulation to other mammalian species, including commercial livestock. The timing of isolation appears important for the isolation of ES cell lines from mice (Brook and Gardner, 1997), therefore, isolation of homologous sequences of marker genes from the species of interest and mapping marker gene expression in the particular species may allow cross-species equivalents of ES or EPL cell to be determined *in vivo*. This thesis suggests that using *Rex1* or *L17* expression may enhance isolation procedures for ES cell equivalents from other mammalian species. However, identification of a marker which distinguishes ICM (3.5 d.p.c.) from early epiblast (4.5 d.p.c.) may facilitate isolation of ES cell-like pluripotent cells, which are most closely related to the 4.5 d.p.c. epiblast (Brook and Gardner, 1997), from other mammalian species.

6.3 ANALYSIS AND POTENTIAL FUNCTIONS OF THE *K7* GENE

Work presented in this thesis determined that *K7* encoded a gene with considerable identity at both the DNA and protein level to a human gene of unknown function, termed *KIAA0165*. Expression of both genes in the adult was restricted to the testis, haematopoietic lineages and gastrointestinal tract, suggesting that these two genes were homologues. The deduced amino acid sequences of K7 and KIAA0165 displayed similarity in their C-terminal regions to proteins from three fungal organisms; Cut1, from *S. pombe*, Esp1, from *S. cerevisiae*, and bimB, from *A. nidulans*. Temperature sensitive mutation of each gene results in almost identical phenotypes, characterised by uncoupling of DNA replication from mitosis such that DNA synthesis occurs in the absence of chromosome segregation (Baum *et al.*, 1988; Uzawa *et al.*, 1990; McGrew *et al.*, 1992; May *et al.*, 1992). Although the three fungal proteins carry out analogous functions, they too only share sequence identity in their C-termini.

The mechanism by which Cut1 and Esp1 mediate chromosome segregation has been investigated in *S. pombe* and *S. cerevisiae* respectively. Both proteins are held in an inactive state by interaction with species specific negative regulatory proteins (Cut2 and Pds1 respectively) which contain destruction box sequences that target these proteins for degradation at the metaphase to anaphase transition, following ubiquitination by the anaphase promoting complex (APC) (Funabiki *et al.*, 1996a and b; Ciosk *et al.*, 1998). Following loss of Pds1 inhibition, Esp1 has been shown to induce cleavage of the cohesin subunit, Scc1, allowing chromosome segregation to occur (Uhlmann *et al.*, 1999). The identification of invariant histidine and cysteine residues in the C-terminal region of Esp1, Cut1, bimB, KIAA0165 and K7 which are flanked by amino acids which share some homology to the active site of a cysteine endopeptidase suggests that the C-terminal region may represent a catalytic domain of a cysteine protease (Nasmyth *et al.*, 2000).

Similarity between the K7 C-terminal amino acid sequence and Cut1 and Esp1 C-terminal protein sequence was supported by functional investigation, which was consistent with a possible role for K7 during early mitosis. K7 was shown to interact with mouse PTTG, a destruction box-containing protein which had been demonstrated to inhibit chromosome segregation following addition of a non-destructable mutant PTTG protein in *Xenopus* (Zou *et al.*, 1999)

mPTTG overexpression in mammalian cells was found to inhibit cell cycle progression during G2/M, presumably as a result of failure of the APC to ubiquitinate excess mPTTG protein, which remains associated with K7, precluding its action of inducing cleavage of cohesin subunits. However, protein localisation studies demonstrated nuclear and cytoplasmically localised mPTTG, which displayed centrosomal association following co-expression of K7. No metaphase arrested cells were observed, suggesting that the mPTTG/K7 complex may function during G2 or early mitosis in mammals.

However, the actual role of K7 during mitosis was not determined. By analogy with Esp1 function, it can be proposed that the K7 protein could contain protease activity and cleave proteins which mediate sister chromatid associations in higher organisms. In vertebrates 95% of the 14S cohesin complex is lost at the onset of mitosis (Losada *et al.*, 1998), suggesting that either the residual 5% of the cohesin complex which remains associated with chromatin mediates mitotic cohesion until anaphase or that mitosis-specific proteins maintain cohesion during mitosis in higher eukaryotes. Therefore, K7 could possibly function at two different stages of mitosis; during prophase to induce 14S cohesin complex dissociation from chromatin, or by analogy with Esp1, at the metaphase to anaphase transition when either the remainder of the cohesin complex, or mitosis-specific factors not part of the 14S cohesin complex, become cleaved to allow chromosome segregation to occur (Figure 5.16). Determination of the timing of K7 activity and PTTG inhibition, and when and how they associate with chromatin will be required in order to understand how proteins such as K7 and PTTG may influence cohesin complex dissociation during mitosis.

A role for K7 during early mitosis was also consistent with the expression pattern of this gene in the embryo and adult. *K7* was detected in pluripotent cells *in vitro* and during early post-implantation mouse development. *K7* expression was upregulated *in vivo* at the stage when accelerated proliferation has been proposed (3.4.2). RNAse protection analysis during later development indicated that *K7* transcripts were developmentally regulated and expressed in all tissues analysed at 16.5 d.p.c, although in a differentially expressed manner. This was consistent with a role for this gene in cell division as proliferation within the developing embryo is required for growth of the foetus. However, *K7* expression in adult tissues was more tightly regulated, with expression detected only in the testis, stomach and the haematopoietic tissues of the bone

marrow, thymus and spleen, but not the mesenteric lymph nodes. These tissues all contain stem cells and/or rapidly proliferating progenitor cells which are required for tissue regeneration. *K7* expression within stem cell-containing tissues suggests a role for this gene in stem cell renewal or accelerated progenitor cell proliferation, consistent with the proposed role for this gene during cell division. An alternate function for *K7* in the testis, the site of highest *K7* expression, can also be proposed. In addition to a potential mitotic role in self renewal of stem cells (type-A spermatogonia) or accelerated proliferation of the intermediate type-B spermatogonia population within the testis, K7 could potentially function during cohesin complex cleavage during meiosis. Investigation of *K7* expression at a cellular level (see 6.3.1.5) could determine the role of this gene in testis function.

6.3.1 Future Analysis of *K7* Function

6.3.1.1 Analysis of the sub-cellular localisation of K7 and mPTTG during the cell cycle

K7 and mPTTG protein localisation during the cell cycle, particularly during the phases of mitosis, has not been investigated. Generation of anti-K7 antibodies and anti-mPTTG antibodies and assaying for protein expression using immunohistochemistry in a synchronous cell population which endogenously expresses the two proteins could allow visualisation of protein localisation during the cell cycle. Cell synchronisations would be carried out using nocodozole to block cells in mitosis and following release from this block, cells would be re-blocked in G1 phase using aphidicolin, for synchronous release and analysis. Immunohistochemical analysis of cells at different time points following release from the G1 phase block, using anti-K7 or anti-mPTTG antibodies, and co-staining with a DNA stain such as Hoechst-33258, would allow visualisation of protein localisation during the phases of mitosis.

Generation of anti-K7 antibodies was attempted during this work using two different regions of the protein which were found to be poorly immunogenic (see section 2.5), suggesting that generation of anti-K7 antibodies may be problematic. If this is the case an alternate method could be employed to analyse K7 protein localisation during the cell cycle. Generation of a mouse cell line, such as NIH 3T3 cells stably overexpressing EGFP-K7 would allow direct visualisation of EGFP-K7 localisation in real time during a 24 hour period, using confocal laser scanning

fluorescent microscopy. Co-staining with a live DNA stain such as Hoechst-33342 would be used to identify the mitotic phase of the cell.

Such protein localisation analyses could determine when these proteins are likely to be functioning during mitosis, and distinguish between possible functional roles during cohesin subunit dissociation at prophase or anaphase.

6.3.1.2 Identification of proteins that interact with K7

Identification of proteins which interact with K7 and the kinetics of the associations may provide valuable information regarding the role of K7 during mitosis. K7 interacting proteins may include proteins involved in localisation of K7 to the centrosome, or proteins which mediate sister chromatid cohesion in mammalian cells. Gel filtration analysis of protein extracts from EPL cells which endogenously express K7, or gel filtration analysis of protein extracts of NIH 3T3 cells overexpressing EGFP-K7, followed by western analysis of gel filtration fractions using anti-K7 antibodies or anti-GFP antibodies, could determine whether K7 is part of a complex, if K7 is detected in gel filtration fractions larger than the size of the native protein. The approximate size of the complex could also be determined.

Synchronisation of either cell line as described in 6.3.1.1, followed by gel filtration analysis of extracts corresponding to different stages of the cell cycle and western blotting of fractions as described above, could determine whether the K7-containing complex changes constituents during mitosis, providing additional insight into the timing of K7 action. Analysis of association and disassociation of mPTTG, using anti-mPTTG antibodies may provide a positive control for this approach, as work carried out in synchronised HeLa cells suggests that PTTG is proteolysed coincident with cyclin B late in mitosis and therefore should be lost at this time from the complex (Zou *et al.*, 1999).

Further purification of the gel filtration fractions which contains the K7 protein complex may allow the proteins within the complex to be identified. Purification using affinity chromatography or immunoprecipitation with anti-K7 or anti-GFP antibodies, electrophoresis by SDS-PAGE and silver staining could allow visualisation of K7 complex components. Protein bands could be excised and analysed by mass spectroscopy and/or N-terminal sequencing to

determine the identity of the proteins within the K7 complex (reviewed in Küster and Mann, 1998).

Yeast two hybrid technology (Clontech) could also be used to identify K7 interacting proteins or verify the interactions identified using gel filtration analysis of K7 complexes. Using the yeast two hybrid approach K7 could be used as a "bait" for detection of interacting proteins using a cDNA library constructed from ES cells, to determine interacting proteins in pluripotent cells, or testis, to determine interacting proteins which may be involved in meiotic K7 function. Isolation and sequence analysis of the cDNAs encoding proteins which demonstrate an interaction with K7 in the yeast two-hybrid screen can be used to determine the identity of the interacting protein. Novel interacting proteins could be further characterised by cDNA cloning and sequence homology to other known proteins or protein motifs in order to identify possible functions for the interaction with K7. Sub-cellular localisation and chromatin association studies of K7-interacting proteins identified using the yeast two hybrid system could also determine potential functions for these proteins in sister chromatid cohesion. Yeast two hybrid analysis using different regions of K7 as bait, for example the C-terminal region which is conserved between mammals and yeast, with mPTTG and other interacting proteins identified during this analysis, could also determine which regions of K7 are important for specific protein-protein interactions.

6.3.1.3 Determining the role of K7 in dissociation of proteins from chromatin in an in vitro assay system

Establishment of an *in vitro* assay system which recapitulates cohesin-chromatin associations would greatly aid the investigation of a direct role for K7 in mediating protein dissociation from chromatin during mitosis in mammals. Crude preparations of chromatin and associated proteins (Coleman *et al.*, 1996), isolated from cell lines (eg. NIH 3T3 cells) arrested in a mitosis by the addition of nocodozole could allow direct investigation of the effect of K7 on dissociation of proteins from chromatin, by the addition of protein extracts overexpressing (eg. NIH 3T3 cell line or ES cell line overexpressing K7) or lacking (eg. K7 double knockout ES cell line) K7, and western blot analysis of proteins which dissociate from chromatin. Detection of proteins dissociating from chromatin would require generation of antibodies against proteins which have the potential to be cleaved in response to the addition of K7.

Scc1 could be investigated initially as a candidate for cleavage following addition of K7 overexpressing cell extracts to the chromatin preparation, as cleavage of this protein has been demonstrated in response to addition of *ESP1* overexpressing extracts in *S. cerevisiae*. However, the cleavage site utilised in Scc1 of *S. cerevisiae* is not conserved in vertebrate Scc1 (Glotzer, 1999). This suggests that either a different cleavage site may be utilised in vertebrate Scc1 or that a different protein may be targeted for cleavage in higher eukaryotes. Therefore, other proteins may need to be investigated using the *in vitro* assay to identify substrates cleaved in response to the addition of K7 expressing extracts. These may include K7 interacting proteins, isolated from experiments suggested in 6.3.1.3, which demonstrate chromatin association in sub-cellular and cell fractionation studies.

Recently, the C-terminal region conserved between K7, KIAA0165, Esp1, Cut1 and bimB has been suggested to contain a putative catalytic site of a cysteine endopeptidase (Nasmyth *et al.*, 2000). Therefore, analysis of the contribution of the invariant histidine and cysteine residues of K7, which are proposed to represent the active site of the protease, could be investigated by Quickchange site directed mutagenesis (Stratagene) of the histidine residue at amino acid position 1999 to alanine and/or the cysteine residue at amino acid position 2025 to serine, overexpressing the mutant protein and adding these extracts to the cleavage assay. This would determine if these residues were specifically required for mediating cleavage of proteins from chromatin.

6.3.1.4 Analysis of K7 function in pluripotent cells in vitro

A direct role for K7 during mitosis in mammals could be determined using mutational analysis of the gene *in vitro* as described earlier in section 6.2.1.1. Targeted removal of both copies of the K7 gene from ES cells could allow direct assessment of the effect of K7 in a pluripotent cell population. Loss of K7 function in ES cells may demonstrate proliferative defects which could be measured using standard proliferation based assays such as the CellTitre 96 nonradioactive cell proliferation assay kit available from Promega or by BrdU incorporation. Flow cytometric analysis of double knockout ES cell lines may also determine which stage of the cell cycle was affected. Defects in chromosome movements could be monitored by the positions

of mitotic chromosomes, detected by DNA specific fluorescent dyes, such as Hoechst-33258 or propidium iodide.

Since a role for K7 in cell division is proposed, *K7* loss of function mutant cells may not be viable. To assess the requirement for *K7* in pluripotent cells, non-degradable mPTTG mutants could be constructed by mutating the destruction box sequence from RKAL to AKAA (Zou *et al.*, 1999). Overexpression of non-destructible mPTTG mutants should prevent mPTTG proteolysis during mitosis. As a result, inhibition of K7 activity should remain throughout the cell cycle and the effect of this dominant negative mutation could be investigated during mitosis, and the timing of K7/mPTTG function determined. Stable transfection of a mPTTG destruction box mutant ($mPTTG^{dm}$) expression construct into pluripotent cell lines, such as ES cells or P19 cells, a mouse EC cell line which displays characteristics of primitive ectoderm (McBurney and Rogers, 1982), using an inducible expression system to express $mPTTG^{dm}$ such as the "Tet on" system (Clontech), would circumvent cell viability problems associated with constitutive overexpression of $mPTTG^{dm}$. Cell proliferation assays and flow cytometric analysis could be used to assay for defects in cell division following induction of $PTTG^{dm}$ expression in ES cells or P19 cells. Furthermore, chromosome movements during mitosis could be monitored using DNA specific fluorescent dyes, such as Hoechst-33258 or propidium iodide, following induction of $mPTTG^{dm}$ expression to identify the timing and type of mitotic dysfunction which may be induced.

6.3.1.5 Analysis of K7 expression in vivo

RNAse protection analysis indicated *K7* expression in the testis, bone marrow, thymus, spleen, and stomach. Testis, bone marrow and the gastrointestinal tract are regenerative tissues which contain stem cells. In order to determine if *K7* is expressed within the stem cell niches or early progenitor cells, *in situ* hybridisation using *K7* specific riboprobes on tissue sections from these mouse tissues could be carried out. The morphological architecture of the testis and gastrointestinal tract is well documented, and the physical locations of the stem cell niches and transiently amplifying populations have been defined (reviewed in de Rooij and van Dissel-Emiliani, 1997; Wright, 1997). To determine whether haematopoietic stem cells or immature but committed lineages express *K7*, mouse bone marrow could be sorted by flow cytometry using antibodies specific for proteins and cell surface markers expressed by the different haematopoietic

lineages (Paul, 1999), followed by RNA isolation and RNAse protection or RT-PCR analysis. The determination of the cell lineages which express *K7* could greatly aid the analysis of mutant phenotypes generated by gene targeting studies (6.3.1.6). RNAse protection analysis and *in situ* hybridisation analysis of other stem cell based tissues such as the keratinocyte stem cells of the skin and hair could also be carried out to provide further insight into the role of K7 in stem cell niches.

An alternate function for *K7* in spermatogenesis can also be proposed as differentiation of type-B spermatogonia into a primary spermatocytes is followed by meiosis in the testis. *K7* could also function to induce cohesin complex dissociation during meiosis. In the mouse, stem cells known as type-A spermatogonia, are first detected 3-7 days after birth and lie on the basement membrane surrounding the seminiferous tubules (Malkov *et al.*, 1998). Primary spermatocytes are observable from approximately postnatal day 14 and located towards the lumen of the seminiferous tubule, and differentiation into mature spermatocytes (spermatids) occurs at the luminal surface by postnatal day 28. Therefore, northern blot analysis of mouse testis RNA isolated at various postnatal stages, in concert with *in situ* hybridisation of mouse testis sections at various stages of the spermatogenic cycle, could identify the *K7* expressing cells. This may allow investigation of whether *K7* functions in stem cell maintenance, in the transient population which divide rapidly and/or during meiosis of spermatocytes.

6.3.1.6 Analysis of K7 function in pluripotent cells in vivo

Creation of a conditional mutant *K7* mouse (reviewed in Shearwin-Whyatt *et al.*, 1999) would allow analysis of not only the effects of loss of *K7* gene expression on pluripotent cells of early mouse development, but analysis of K7 function in specific tissues, such as the testis, gastrointestinal tract and haematopoietic lineages. The phenotype of *K7* conditional mutants could be analysed in a number of ways. Firstly, phenotypes of conditional *K7* mutants in adult mice would be analysed initially by assessment of morphological characteristics of tissues which express *K7*, using basic histology. Defects specific for the testis may include abnormal seminiferous tubule morphology, lacking characteristics of spermatogenesis (de Rooij and van Dissel-Emiliani, 1997). The effect of this defect on mouse fertility could then be assessed using timed matings. Gastrointestinal tract dysfunction could be assessed by morphological analysis of

the intestinal crypt for villus atrophy (Wright, 1997), and measurement of body weight, as villus architecture is likely to contribute directly to weight loss due to decreased nutritional absorption (Chen *et al.*, 1990). Defects in the organ size, architecture and cellularity of tissues of the haematopoietic system which express *K7,* such as bone marrow, spleen and thymus could suggest whether targeted removal of the *K7* gene affects the formation, architecture or presence of certain cell lineages in these tissues.

As *K7* expression within these tissues was proposed to affect stem cells or rapidly proliferating progenitor cells, investigation of stem cell or early progenitor cell defects, such as their proliferative capacity could be analysed by examining the incorporation of BrdU into DNA during S phase as described previously by Hayashi *et al.* (1988). The morphology of the testis and gastrointestinal tract has been studied extensively and the sites of stem cell niches and progenitor cells are well documented (de Rooij and van Dissel-Emiliani, 1997; Wright, 1997) such that histological analysis of tissue morphology could determine if stem cells and progenitor cell populations were present and in the correct cellular architecture within the tissue. Any defects observed could be examined further using expression analysis of marker genes specific for stem cells, progenitor cell lineages or differentiated progeny. Flow cytometry of haematopoietic lineages using antibodies directed to cell surface antigens specific for particular lineages, such as stem cells or progenitor cells of red blood cells, lymphocytes, granulocytes and monocytes (Paul, 1999), could determine not only if stem cells and progenitor cells have decreased numbers in *K7* mutants, but also if the differentiation of any of these lineages from the haematopoietic stem cells is compromised following mutation of *K7*. Furthermore, *in vitro* haematopoietic colony forming unit assays (reviewed in Robb, 1997) could be used to determine if early progenitor cell compartments are compromised by loss of *K7* function. Abnormal chromosome segregation has been implicated in genomic instability or aneuploidy, which underlies many cancers (Lengauer *et al.*, 1998). Therefore, analysis of metaphase spreads for chromosome fusions by abnormal G-banding patterns and FISH analysis using probes directed against telomere repeats would determine if cytogenetic abnormalities such as chromosome instability exist following targeted removal of the *K7* gene.

BIBLIOGRAPHY

Abbas, A. K., Lichtman, A. H., and Pober, J. S. (1991) Cellular and molecular immunology, W.B. Saunders: Philadelphia.

Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A., and Brûlet, P. (1995) Forebrain and midbrain regions are deleted in *Otx2-/-* mutants due to a defective anterior neuroectoderm specification during gastrulation. *Development,* **121**: 3279-3290.

Albano, R. M., Groome, N., and Smith, J. C. (1993) Activins are expressed in preimplantation mouse embryos and in ES and EC cells and are regulated on their differentiation. *Development,* **117**: 711-23.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) Molecular biology of the cell, Garland publishing: New York.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.,* **215**: 403-410.

Ambrosetti, D., Basilico, C., and Dailey, L. (1997) Synergistic activation of the Fibroblast Growth Factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.,* **17**: 6321-6329.

Amon, A. (1999) The spindle checkpoint. *Curr. Opin. Gen. Dev.,* **9**: 69-75.

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J. K., and Lonai, P. (1998) Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pre-gastrulation mammalian development. *Proc. Natl. Acad. Sci. USA,* **95**: 5082-5087.

Baum, P., Yip, C., Goetsch, L., and Byers, B. (1988) A yeast gene essential for regulation of spindle pole duplication. *Mol. Cell. Biol.,* **8**: 5386-5397.

Becker, D. L., Leclerc-David, C., and Warner, A. (1992) The relationship of gap junctions and compaction in the preimplantation mouse embryo. *Development,* **Supp.**: 113-118.

Beddington, R. S. P. (1982) An autographic analysis of tissue potency in different regions of the embryonic ectoderm during gastrulation in the mouse. *J. Embryol. exp. Morphol.,* **69**: 265-285.

Beddington, R. S. P. (1983) The origin of foetal tissues during gastrulation in the rodent. *In:* Development in Mammals (M. H. Johnson, ed.), Elsevier: Amsterdam, pg. 1-32.

Beddington, R. S. P., and Robertson, E. J. (1989) An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development,* **105**: 733-737.

Beddington, R. S. P., and Robertson, E. J. (1998) Anterior patterning in mouse. *TIG,* **14**: 277-284.

Beddington, R. S. P., and Robertson, E. J. (1999) Axis development and early asymmetry in mammals. *Cell,* **96**: 195-209.

Belo, J. A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M., and De Robertis, E. M. (1997) *Cerberus-like* is a secreted factor with neutralising activity expressed in the anterior visceral endoderm of the mouse gastrula. *Mech. Dev.,* **68**: 45-57.

Ben-Shushan, E., Thompson, J. R., Gudas, L. J., and Bergman, Y. (1998) *Rex-1*, a gene encoding a transcription factor expressed in the early embryo, is regulated via Oct-3/4 and Oct-6 binding to an Octomer site and a novel protein, Rox-1, binding to an adjacent site. *Mol. Cell. Biol.,* **18**: 1866-1878.

Bernard, O., Burkitt, V., Webb, G. C., Bottema, C. D., Nicholl, J., Sutherland, G. R., and Matthew, P. (1996) Structure and chromosome localisation of the genomic locus encoding the *Kiz1* LIM-kinase gene. *Genomics,* **35**: 593-596.

Bettess, M. (1993) Gene expression in murine pluripotential stem cells. *Honours Thesis.* Department of Biochemistry, University of Adelaide.

Bhatt, H., Brunet, L. J., and Stewart, C. L. (1992) Uterine expression of leukaemia inhibitory factor coincides with the onset of blastocyst implantation. *Proc. Natl. Acad. Sci. USA,* **88**: 11408-11412.

Blum, M., Gaunt, S. J., Cho, K. W. Y., Steinbeisser, H., Blumberg, B., Bittner, D., and De Robertis, E. M. (1992) Gastrulation in the mouse: the role of the homeobox gene *goosecoid. Cell,* **69**: 1097-1106.

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

Brady, G., Billia, F., Knox, J., Hoang, T., Kirsh, I. R., Voura, E. B., Hawley, R. G., Cumming, R., Buchwald, M., Siminovitch, K., Miyamoto, N., Boehmelt, G., and Iscove, N. N. (1995) Analysis of gene expression in a complex differentiation hierachy by global amplification of cDNA from single cells. *Curr. Biol.,* **5**: 909-922.

Brandon, E. P., Idzera, R. L., and McKnight, G. S. (1995a) Targeting the mouse genome: a compendium of knockouts (part I). *Curr. Biol.,* **5**: 625-634.

Brandon, E. P., Idzera, R. L., and McKnight, G. S. (1995b) Targeting the mouse genome: a compendium of knockouts (part II). *Curr. Biol.,* **5**: 758-765.

Brandon, E. P., Idzera, R. L., and McKnight, G. S. (1995c) Targeting the mouse genome: a compendium of knockouts (part III). *Curr. Biol.,* **5**: 873-881.

Brook, F. A., and Gardner, R. L. (1997) The origin and efficient derivation of embryonic stem cells in the mouse. *Proc. Natl. Acad. Sci. USA,* **94**: 5709-5712.

Buehr, M., and McLaren, A. (1974) Size regulation in chimaeric mouse embryos. *J. Embryol. exp. Morphol.*, **31**: 229-234.

Cabin, D. E., and Reeves, R. H. (1998) Mouse Chromosome 16. *Mammalian Genome,* **8** (special issue): S307-S319.

Celano, P., Vertino, P. M., and Casero Jr., R. A. (1993) Isolation of polyadenylated RNA from cultured cells and intact tissue. *Biotechniques,* **15**: 27-28.

Chai, N., Patel, Y., Jacobson, K., McMahon, J., McMahon, A., and Rappolee, D. A. (1998) FGF is an essential regulator of the fifth cell division in preimplantation mouse embryos. *Dev. Biol.*, **198**: 105-115.

Chapman, C., Remiszewski, J. L., Webb, G. C., Schulz, T. C., Bottema, C. D. K., and Rathjen, P. D. (1997) The mouse homeobox gene, *Gbx2*: Genomic organisation and expression in pluripotent cells *in vitro* and *in vivo*. *Genomics,* **46**: 223-233.

Chen, J.-M., Rawlings, N. D., Stevens, R. A. E., and Barrett, A. J. (1998) Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. *FEBS Lett.*, **441**: 361-365.

Chen, T. S., Currier, G. J., and Wabner, C. L. (1990) Intestinal transport during the life span of the mouse. *J. Gerontol.*, **45**: 129-133.

Chen, W. S., Manova, K., Weinstein, D. C., Duncan, S. A., Plump, A. S., Prezioso, V. R., Bachavarova, R. F., and Darnell Jr., J. E. (1994) Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes & Dev.*, **8**: 2466-2477.

Chisholm, J. C., Johnson, M. H., Warren, P. D., Fleming, T. P., and Pickering, S. J. (1985) Developmental variability within and between mouse expanding blastocysts and their ICMs. *J. Embryol. exp. Morphol.*, **86**: 311-336.

Chomczynski, P., and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.,* **162**: 156-159.

Ciosk, R., Zachariae, W., Michaelis, C., Shevchenko, A., Mann, M., and Nasmyth, K. (1998) An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. *Cell,* **93**: 1067-1076.

Ciruna, B. G., Schwartz, L., Harpal, K., Yamaguchi, T. P., and Rossant, J. (1997) Chimeric analysis of *fibroblast growth factor receptor-1 (Fgfr1)* function: a role for FGFR1 in morphogenetic movement through the primitive streak. *Development,* **124**: 2829-2841.

Coleman, T. R., Carpenter, P. B., and Dunphy, W. G. (1996) The *Xenopus* Cdc6 protein is essential for the initiation of a single round of DNA replication in cell free extracts. *Cell,* **87**: 53-63.

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., and Robertson, E. J. (1994) A primary requirement for *nodal* in the formation and maintenance of the primitive streak in the mouse. *Development,* **120**: 1919-1928.

Conover, J. C., Ip, N. Y., Poueymirou, W. T., Bates, B., Goldfarb, M. P., DeChiara, T. M., and Yancopoulos, G. D. (1993) Ciliary neurotrophic factor maintains the pluripotentiality of embryonic stem cells. *Development,* **119**: 559-65.

Conquet, F., Peyrieras, N., Tiret, L., and Brûlet, P. (1992) Inhibited gastrulation in mouse embryos overexpressing the leukemia inhibitory factor. *Proc. Natl. Acad. Sci. USA,* **89**: 8195-8199.

Coucouvanis, E., and Martin, G. R. (1995) Signals for death and survival: a two-step mechanism for cavitation in the vertebrate embryo. *Cell,* **83**: 279-287.

Coucouvanis, E., and Martin, G. R. (1999) BMP signaling plays a role in visceral endoderm differentiation and cavitation in the early moue embryo. *Development,* **126**: 535-546.

Crossley, P., and Martin, G. (1995) The mouse *Fgf8* gene encodes a family of polypeptides that is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development,* **121**: 439-451.

Damjanov, I., Damjanov, A., and Solter, D. (1987) Production of teratocarcinomas from embryo's transplanted to extra-uterine sites. *In:* Teratocarcinomas and Embryonic Stem Cells, A Practical Approach (E.J. Robertson, ed.), IRL Press: Oxford, pg. 1-18.

Dardik, A., Smith, R. M., and Schultz, R. M. (1992) Colocalisation of transforming growth factor-alpha and functional epidermal growth factor receptor (EGFR) to the inner cell mass and preferential localisation of the EGFR on the basolateral surface of the trophectoderm in the mouse blastocyst. *Dev. Biol.,* **154**: 396-409.

de Rooij, D. G., and van Dissel-Emiliani, F. M. F. (1997) Regulation of proliferation and differentiation of stem cells in the male germ line. *In:* Stem cells (C. S. Potten, ed.), Academic press: London, pg. 283-313.

DeGregori, J., Russ, A., von Melchner, H., Priyaranjan, P., Jenkins, N. A., Copeland, N. G., and Ruley, H. G. (1994) A murine homolog of the yeast RNA1 gene is required for postimplantation development. *Genes & Dev.,* **8**: 265-276.

Deng, C.-X., Wynshaw-Boris, A., Shen, M. M., Daugherty, C., and Ornitz, D. M. (1994) Murine FGFR-1 is required for postimplantation growth and axial organisation. *Genes & Dev.,* **8**: 3045-3057.

Ding, J., Yang, L., Yan, Y., Chen, A., Desai, N., Wynshaw-Boris, A., and Shen, M. M. (1998) *Cripto* is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature,* **395**: 702-707.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985) The *in vitro* development of blastocyst derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. exp. Morphol.*, **87**: 27-45.

Duncan, S. A., Manova, K., Chen, W., Hoodless, P., Wienstein, D. C., Bacharova, R. F., and Darnell Jr, J. E. (1994) Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proc. Natl. Acad. Sci. USA,* **91**: 7598-7602.

Duncan, S. A., Nagy, A., and Chan, W. (1997) Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of *HNF-4-/-* embryos. *Development,* **124**: 279-287.

Dush, M. K., and Martin, G. R. (1992) Analysis of Mouse Evx Genes: *Evx-1* Displays Graded Expression in the Primitive Streak. *Dev. Biol.*, **151**: 273-287.

Dziadek , M. A., and Adamson, E. (1978) Localisation and synthesis of alphafetoprotein in post-implantation mouse embryos. *J. Embryol. exp. Morphol.*, **43**: 298-313.

Dziadek, M. (1979) Cell differentiation in isolated inner cell masses of mouse blastocysts *in vitro:* onset of specific gene expression. *J. Embryol. exp. Morphol.*, **53**: 367-379.

Edwards, D. R., Parfett, C. L., and Denhardt, D. T. (1985) Transcriptional regulation of two serum-induced RNAs in mouse fibroblasts: equivalence of one species to B2 repetitive elements. *Mol. Cell. Biol.*, **5**: 3280-8.

Ema, M., Suzuki, M., Morita, M., Hirose, K., Sogawa, K., Matsuda, Y., Gotoh, O., Sajjoh, Y., Fujii, H., Hamada, H., and Fujii-Kuriyama, Y. (1996) cDNA cloning of a murine homologue of *Drosophila* single-minded, its mRNA expression in mouse development and chromosome localisation. *Biochem. Biophys. Res. Comm.*, **218**: 588-594.

Evans, E. P. (1989) Standard ideogram. *In:* Genetic variants of the laboratory mouse (M. F. Lyon and A. G. Searle, eds.), Oxford University Press: Oxford, pg. 576-577.

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

Evans, M. J., and Kaufman, M. H. (1983) Pluripotential cells grown directly from normal mouse embryos. *Cancer Surv.,* **2**: 185-207.

Fässler, R., and Meyer, M. (1995) Consequences of lack of β1 integrin gene expression in mice. *Genes & Dev.,* **9**: 1896-1908.

Feldman, B., Poueymirou, W., Papionannou, V. E., DeChiara, T. M., and Goldfarb, M. (1995) Requirement of FGF-4 for postimplantation mouse development. *Science,* **267**: 246-249.

Festing, M. F. W. (1992) Origins and characteristics of inbred strains of mice. *Mouse Genome,* **90**: 231-315.

Fischer, D., and Eisenberg, D. (1996) Protein fold recognition using sequence derived predictions. *Protein Sci.,* **5**: 947-955.

Fleming, T. P., Javed, Q., and Hay, M. (1992) Epithelial differentiation and intercellular junction formation in the mouse early embryo. *Development,* **Supp.**: 105-12.

Fleming, T. P., and Pickering, S. J. (1985) Maturation and polarisation of the endocytotic system in outside blastomeres during mouse preimplantation development. *J. Embryol. exp. Morphol.,* **89**: 175-208.

Fraidenraich, D., Lang, R., and Basilico, C. (1998) Distinct regulatory elements govern *Fgf4* gene expression in the mouse blastocyst, myotomes and developing limb. *Dev. Biol.,* **204**: 197-209.

Funabiki, H., Kumada, K., and Yanagida, M. (1996a) Fission yeast Cut1 and Cut2 are essential for sister chromatid separation, concentrate along the metaphase spindle, and form large complexes. *EMBO J.,* **15**: 6617-6628.

Funabiki, H., Yamano, H., Kumada, K., Nagao, K., Hunt, T., and Yanagida, M. (1996b) Cut2 proteolysis required for sister-chromatid separation in fission yeast. *Nature,* **381**: 438-441.

Funabiki, H., Yamano, H., Nagao, K., Tanaka, H., Yasuda, H., Hunt, T., and Yanagida, M. (1997) Fission yeast Cut2 required for anaphase has two destruction boxes. *EMBO J.,* **16**: 5977-5987.

Gardner, R., and Papaioannou, V. E. (1975) Differentiation in the Trophectoderm and Inner Cell Mass. *In:* The Early Development of Mammals (M. Balls and A. E. Wild, eds.), Cambridge University Press: London, pg. 107-132.

Gardner, R. L. (1971) Manipulations on the blastocyst. *Advan. Biosci.,* **6**: 279-296.

Gardner, R. L. (1983) Origin and differentiation of extraembryonic tissues of the mouse. *Int. Rev. exp. Pathol.,* **24**: 63-133.

Gardner, R. L. (1985) Regeneration of endoderm from primitive ectoderm in the mouse embryo: fact or artefact? *J. Embryol. exp. Morphol.,* **88**: 303-326.

Gardner, R. L. (1997) The early blastocyst is bilaterally symmetrical and its axis of asymmetry is aligned with the animal-vegetal axis of the zygote in the mouse. *Development,* **124**: 289-301.

Gardner, R. L., and Cockroft, D. L. (1998) Complete dissipation of coherent clonal growth occurs before gastrulation in mouse epiblast. *Development,* **125**: 2397-2402.

Gardner, R. L., and Johnson, M. H. (1972) An investigation of inner cell mass and trophoblast tissues following their isolation from the mouse blastocyst. *J. Embryol. exp. Morphol.,* **28**: 279-312.

Gardner, R. L., Papaioannou, V. E., and Barton, S. C. (1973) Origin of the ectoplacental cone and secondary giant cells in mouse blastocysts reconstituted from isolated trophoblast and inner cell mass. *J. Embryol. exp. Morphol.,* **30**: 561-572.

Gardner, R. L., and Rossant, J. (1979) Investigation of the fate of 4.5 day post-coitum mouse inner cell mass cells by blastocyst injection. *J. Embryol. exp. Morphol.,* **52**: 141-152.

Gavrieli, Y., Sherman, Y., and Ben-Sasson, S. A. (1992) Identification of programmed cell death *in situ* via specific labelling of nuclear DNA fragmentation. *J. Cell Biol.,* **118**: 493-501.

Gearing, D. P., and Bruce, G. (1992) Oncostatin M binds the high-affinity leukemia inhibitory factor receptor. *New Biol.,* **4**: 61-65.

Ginsburg, M., Snow, M. L. H., and McLaren, A. (1990) Primordial germ cells in the mouse embryo during gastrulation. *Development,* **110**: 521-528.

Glotzer, M. (1999) Chromosome segregation: Samurai separation of siamese sisters. *Curr. Biol.,* **9**: 531-534.

Gong, J., Ardelt, B., Traganos, F., and Darzynkiewicz, Z. (1994) Unscheduled expression of cyclin B1 and cyclin E in several leukemic and solid tumour lines. *Cancer Res.,* **54**: 4285-4288.

Guacci, V., Koshland, D., and Strunnikov, A. (1997) A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of *MCD1* in *S. cerevisiae*. *Cell,* **91**: 47-57.

Haffner-Krausz, R., Gorivodshy, M., Chen, Y., and Lonai, P. (1999) Expression of *Fgfr2* in the early mouse embryo indicates its involvement in preimplantation development. *Mech. Dev.*, **85**: 167-172.

Hahnel, A. C., Rappolee, D. A., Millan, J. L., Manes, T., Ziomek, C. A., Theodosiou, N. G., Werb, Z., Pederson, R. A., and Schultz, G. A. (1990) Two alkaline phosphatase genes are expressed during early development in the mouse embryo. *Development,* **110**: 555-564.

Hakem, R., de la Pompa, J. L., Sirad, C., Mo, R., Woo, M., Hakem, A., Potter, J., Reitmair, A., Billia, F., Firpo, E., Hui, C. C., Roberts, J., Rossant, J., and Mak, T. W. (1996) The tumor suppressor gene *Brca1* is required for embryonic proliferation in the mouse. *Cell,* **85**: 1009-1023.

Harrison, S. M., Dunwoodie, S. L., Arkell, R. M., Lehrach, H., and Beddington, R. S. (1995) Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo. *Development,* **121**: 2479-2489.

Harvey, M. B., Leco, K. J., Arcellana-Panlilio, M. Y., Zhang, X., Edwards, D. R., and Schultz, G. A. (1995) Proteinase expression in the early mouse embryo is regulated by leukaemia inhibitory factor and epidermal growth factor. *Development,* **121**: 1005-1014.

Haub, O., and Goldfarb, M. (1991) Expression of the *fibroblast growth factor-5* gene in the mouse embryo. *Development,* **112**: 397-406.

Hayashi, Y., Koike, M., Matsuname, M., and Hoshino, T. (1988) Effects of fixation time and enzymic digestion on immunohistochemical demonstration of bromodeoxyuridine in formalin-fixed, paraffin-embedded tissue. *J. Histochem. Cytochem.*, **36**: 511-514.

Hébert, J. M., Boyle, M., and Martin, G. R. (1991) mRNA localisation studies suggest that murine *FGF-5* plays a role in gastrulation. *Development,* **112**: 407-15.

Hébert, J. M., Rosenquist, T., Gotz, J., and Martin, G. R. (1994) FGF5 as a regulator of the hair growth cycle: evidence from targeted and spontaneous mutations. *Cell,* **78**: 1017-1025.

Heldin, C.-H. (1995) Dimerisation of cell surface receptors in signal transduction. *Cell,* **80**: 213-223.

Herrmann, B. G. (1991) Expression pattern of the *Brachyury* gene in whole-mount T^{Wis}/T^{Wis} mutant embryos. *Development,* **113**: 913-917.

Hess, E. J. (1996) Identification of the *weaver* mouse mutation: The end of the beginning. *Neuron,* **16**: 1073-1076.

Hirano, T. (1999) SMC-mediated chromosome mechanics: a conserved scheme from bacteria to vertebrates. *Genes & Dev.,* **13**: 11-19.

Hogan, B., Beddington, R., Constantini, F., and Lacy, E. (1994) Manipulating the Mouse Embryo: A Laboratory Manual (2nd Edition), Cold Spring Harbour Laboratory Press: Cold Spring Harbour, NY.

Hogan, B., Constantini, F., and Lacy, E. (1986) Manipulating the Mouse Embryo: A Laboratory Manual (1st Edition), Cold Spring Harbour Laboratory Press: Cold Spring Harbour, NY.

Hogan, B., and Tilly, R. (1977) *In vitro* culture and differentiation of normal mouse blastocysts. *Nature,* **265**: 626-629.

Hogan, L. M., Barlow, D. P., and Tilly, R. (1983) F9 Teratocarcinoma cells as a model for the differentiation of parietal and visceral endoderm in the mouse embryo. *Cancer Surv.,* **2**: 115-140.

Holland, P. W. H., Harper, S. J., McVey, J. H., and Hogan, B. J. (1987) *In vivo* expression of mRNA for the Ca^{2+}-binding protein *SPARC* (Ostenonectin) revealed by *in situ* hybridisation. *J. Cell Biol.,* **105**: 473-482.

Holloway, S. L., Glotzer, M., King, R. W., and Murray, A. W. (1993) Anaphase is initiated by proteolysis rather than by the inactivation of maturation-promoting factor. *Cell,* **73**: 1393-1402.

Hosler, B. A., LaRosa, G. J., Grippo, J. F., and Gudas, L. J. (1989) Expression of *REX-1*, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. *Mol. Cell. Biol.,* **9**: 5623-9.

Hosler, B. A., Rogers, M. B., Kozak, C. A., and Gudas, L. J. (1993) An octamer motif contributes to the expression of the retinoic acid-regulated zinc finger gene Rex-1 (Zfp-42) in F9 teratocarcinoma cells. *Mol. Cell. Biol.,* **13**: 2919-28.

Ikuta, K., Kina, T., MacNeil, U., Uchida, N., Peault, B., Chien, Y. H., and Weissman, I. L. (1990) A developmental switch in thymic lymphocyte maturation potential occurs at the level of haematopoietic stem cells. *Cell,* **62**: 863-874.

Jackson, B. W., Grund, C., Winter, S., Franke, W. W., and Illmensee, K. (1981) Formation of cytoskeletal elements during mouse embryogenesis. Epithelial differentiation and intermediate-sized filaments in early postimplantation embryos. *Differentiation,* **20**: 203-216.

Johansson, B. M., and Wiles, M. V. (1995) Evidence for involvement of activin A and bone morphogenetic protein-4 in mammalian mesoderm and haemopoietic development. *Mol. Cell. Biol.,* **15**: 141-151.

Johnson, L. V., Calarco, P. G., and Siebert, M. L. (1977) Alkaline phosphatase activity in the preimplantation mouse embryo. *J. Embryol. exp. Morphol.,* **40**: 83-89.

Johnson, M. H., and Ziomek, C. A. (1983) Cell interactions influence the fate of mouse blastomeres undergoing the transition from the 16- to 32-cell stage. *Dev. Biol.*, **95**: 211-218.

Jones-Villeneuve, E. M. V., McBurney, M. W., Rogers, K. A., and Kallins, V. I. (1982) Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. *J. Cell Biol.*, **94**: 253-262.

Kamei, M., Webb, G. C., Heydon, K., Hendry, I. A., Young, I. G., and Campbell, H. D. (2000) *Sohl*, the mouse homologue of the *Drosophila melanogaster small lobes* gene: Organisation, chromosomal mapping and localisation of gene product to the olfactory bulb. *Genomics,* **64**: 82-89.

Kaufman, M. H. (1992) The atlas of mouse development, Academic Press: London.

Kavanagh, S. J. (1998) The pluripotent stem cell marker Psc1 localises to nuclear speckles. *Honours Thesis*. Department of Biochemistry, University of Adelaide.

Keller, G., Kennedy, M., Papayannopoulou, T., and Wiles, M. V. (1993) Haematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell. Biol.*, **13**: 473-486.

Kerrebrock, A. W., Miyazaki, W. Y., Birnby, D., and Orr-Weaver, T. L. (1992) The *Drosophila mei-S332* gene promotes sister chromatid cohesion in meiosis following kinetochore differentiation. *Genetics,* **130**: 827-841.

Kerrebrock, A. W., Moore, D. P., Wu, J. S., and Orr-Weaver, T. L. (1995) MEI-S332, a Drosophila protein required for sister chromatid cohesion, can localise to meiotic centrosome regions. *Cell,* **83**: 247-256.

Kidder, G. M. (1992) The genetic program for preimplantation development. *Dev. Genet.*, **13**: 319-25.

Krieg, P. A., and Melton, D. A. (1987) *In vitro* RNA synthesis with SP6 RNA polymerase. *Meth. Enzymol.,* **155**: 397-415.

Kumada, K., Nakamura, T., Nagao, K., Funabiki, H., Nakagawa, T., and Yanagida, M. (1998) Cut1 is loaded onto the spindle by binding to Cut2 and promotes anaphase spindle movement upon Cut2 proteolysis. *Curr. Biol.,* **8**: 633-641.

Küster, B., and Mann, M. (1998) Identifying proteins and post-translational modifications by mass spectrometry. *Curr. Opin. Struc. Biol.,* **8**: 393-400.

Labosky, P. A., Barlow, D. P., and Hogan, B. L. (1994) Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of *insulin-like growth factor 2 receptor (Igf2r)* gene compared with embryonic stem (ES) cell lines. *Development,* **120**: 3197-3204.

Lake, J.-A. (1996) Differentiation of pluripotent murine embryonic stem cells. *PhD. Thesis.* Department of Biochemistry, University of Adelaide.

Lake, J.-A., Rathjen, J., Remiszewski, J. L., and Rathjen, P. D. (2000) Reversible programming of pluripotent cell differentiation. *J. Cell Sci.,* **113**: 555-566.

Lane, P. W. (1964) Weaver, wv. *Mouse News Lett.,* **30**: 32.

Latham, K. E., Garrels, J. I., Chang, C., and Solter, D. (1991) Quantitative analysis of protein synthesis in mouse embryos. Extensive reprogramming at the one- and two-cell stages. *Development,* **112**: 921-933.

Lawson, K. A., and Hage, W. J. (1994) Clonal analysis of the origin of primordial germ cells in the mouse. In: Germline Development.*Ciba Found. Symp.* Vol. 182, Wiley: Chichester, pg. 68-91.

Lawson, K. A., Menses, J. J., and Pedersen, R. A. (1991) Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development,* **113**: 891-911.

Lee, S., Gilula, N. B., and Warner, A. E. (1987) Gap junctional communication and compaction during preimplantation stages of mouse development. *Cell,* **51**: 851-860.

Lemieux, N., Dutrillaux, B., and Viegas-Pequignot, E. (1992) A simple method for simultaneous R- or G-banding and fluorescence *in situ* hybridisation of a single-copy gene. *Cytogenet. Cell Genet.,* **59**: 311-312.

Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature,* **396**: 643-649.

Lewis, N. E., and Rossant, J. (1982) Mechanism of size regulation in mouse embryo aggregates. *J. Embryol. exp. Morphol.,* **72**: 169-181.

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

Lim, D.-S., and Hasty, P. (1996) A mutation in mouse *rad51* results in an early embryonic lethal that is suppressed by a mutation in *p53*. *Mol. Cell. Biol.,* **16**: 7133-7143.

Lints, T. J., Parsons, L. M., Hartley, L., Lyons, I., and Harvey, R. P. (1993) *Nkx-2.5*: a novel murine homeobox gene expressed in the early heart progenitor cells and their myogenic descendants. *Development,* **119**: 419-431.

Liu, C.-Y., Flesken-Nikitin, A., Li, S., Zeng, Y., and Lee, W.-H. (1996) Inactivation of the mouse *Brca1* gene leads to failure in the morphogenesis of the egg cylinder in early postimplantation development. *Genes & Dev.,* **10**: 1835-1843.

Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R., and Bradley, A. (1999) Requirement for *Wnt3* in vertebrate axis formation. *Nature Genetics,* **22**: 361-365.

Losada, A., Hirano, M., and Hirano, T. (1998) Identification of *Xenopus* SMC protein complexes required for sister chromatid cohesion. *Genes & Dev.*, **12**: 1986-1997.

Ludwig, T., Chapman, D. L., Papaioannou, V. E., and Efstratiadis, A. (1997) Targeted mutations of breast cancer susceptibility gene homologs in mice: lethal phenotypes of *Brca1*, *Brca2*, *Brca1/Brca2*, *Brca1/p53*, and *Brca2/p53* nullizygous embryos. *Genes & Dev.*, **11**: 1226-1241.

Lyons, I., Parsons, L. M., Hartley, L., Ruili, L., Andrews, J. E., Robb, L., and Harvey, R. P. (1995) Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene *Nkx-2.5*. *Genes & Dev.*, **9**: 1654-1666.

Malkov, M., Fisher, Y., and Don, J. (1998) Developmental schedule of the postnatal rat testis determined by flow cytometry. *Biol. Reprod.*, **59**: 84-92.

Manejwala, F. M., Cragoe, E. J., and Schultz, R. M. (1989) Blastocoel expansion in the preimplantation mouse embryo: role of extracellular sodium and chloride and possible apical routes of their entry. *Dev. Biol.*, **133**: 210-220.

Mantalenakis, S. J., and Ketchel, M. M. (1966) Frequency and extent of delayed implantation in lactating rats and mice. *J. Reprod. Fertil.*, **12**: 391-394.

Martin, G. R. (1980) Teratocarcinomas and Mammalian Embryogenesis. *Science,* **209**: 768-776.

Martin, G. R. (1981) Isolation of a pluripotential cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA,* **72**: 1441-1445.

Martin, G. R., and Evans, M. J. (1974) The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. *Cell,* **2**: 163-172.

Martin, G. R., and Evans, M. J. (1975a) Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies *in vitro*. *Proc. Natl. Acad. Sci. USA,* **72**: 1441-1445.

Martin, G. R., and Evans, M. J. (1975b) Multiple differentiation of clonal teratocarcinoma stem cells following embryoid body formation in vitro. *Cell,* **6**: 467-474.

Martin, G. R., Smith, S., and Epstien, C. J. (1978) Protein synthesis patterns in teratocarcinoma stem cells and mouse embryos at early stages of development. *Dev. Biol.,* **66**: 8-16.

Massague, J. (1996) TGF beta signaling: receptors, transducers and Mad proteins. *Cell,* **85**: 947-950.

Matsui, Y., Zsebo, K., and Hogan, B. M. L. (1992) Derivation of pluripotent embryonic stem cells from murine primordial germ cells in culture. *Cell,* **70**: 841-847.

May, G. S., McGoldrick, C. A., Holt, C. L., and Denison, S. H. (1992) The *bimB3* mutation of aspergillus nidulans uncouples DNA replication from the completion of mitosis. *J. Biol. Chem.,* **267**: 15737-15743.

McBurney, M. W., Jones-Villeneuve, E. M. V., Edwards, M. K., and Anderson, P. J. (1982) Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. *Nature,* **299**: 165-167.

McBurney, M. W., and Rogers, B. J. (1982) Isolation of male embryonal carcinoma cells and their chromosome replication patterns. *Dev. Biol.,* **89**: 503-508.

McGrew, J. T., Goetsch, L., Byers, B., and Baum, P. (1992) Requirement for ESP1 in the nuclear division of Saccharomyces cerevisiae. *Mol. Biol. Cell,* **3**: 1443-1454.

Melton, D. A. (1991) Pattern formation during animal development. *Science,* **252**: 234-241.

Metcalf, D. (1989) The molecular control of cell division, differentiation, commitment and maturation in haemopoietic cells. *Nature,* **339**: 27-30.

Metcalf, D. (1991) The colony stimulating factors: discovery to clinical use. *Philosophical Transcripts of the Royal Society of London,* **333**: 147-173.

Michaelis, C., Ciosk, R., and Nasmyth, K. (1997) Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell,* **91**: 35-45.

Mishina, Y., Suzuki, A., Ueno, N., and Behringer, R. R. (1996) *Bmpr* encodes a type 1 bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes & Dev.,* **9**: 3027-3037.

Monk, M., and McLaren, A. (1981) X-chromosome activity in foetal germ cells of the mouse. *J. Embryol. exp. Morphol.,* **63**: 75-84.

Moore, D. P., Page, A. W., Tang, T. T. L., Kerrebrock, A. W., and Orr-Weaver, T. L. (1998) The cohesion protein MEI-S332 localises to condensed meiotic and mitotic centromeres until sister chromatids separate. *J. Cell Biol.,* **140**: 1003-1012.

Moreau, J.-F., Donaldson, D. D., Bennett, F., Witek-Giannotti, J., Clark, S. C., and Wong, G. G. (1988) Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. *Nature,* **336**: 690-692.

Morrisey, E. E., Tang, Z., Sigrist, K., Lu, M. M., Jiang, F., Ip, H. S., and Parmacek, M. S. (1998) GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes & Dev.,* **12**: 3579-3590.

Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M., I., C., and Smith, A. G. (1994) Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression. *Proc. Natl. Acad. Sci. USA,* **91**: 4303-4307.

Mummery, C. L., Feyen, A., Freund, E., and Shen, S. (1990) Characteristics of embryonic stem cell differentiation: a comparison with two embryonal carcinoma cell lines. *Cell. Diff & Dev.*, **30**: 195-206.

Mummery, C. L., van Achterberg, T. A. E., van den Eijnden Raaij, A. J. M., van Haaster, L., de Laat, S. W., and Piersma, A. H. (1991) Visceral endoderm like cells induce differentiation of murine P19 embryonal carcinoma cells. *Differentiation,* **46**: 51-60.

Nagase, T., Seki, N., Ishikawa, K., Tanaka, A., and Nomura, N. (1996) Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res.*, **3**: 17-24.

Nasmyth, K., Peters, J.-M., and Uhlmann, F. (2000) Splitting the chromosome: Cutting the ties that bind sister chromatids. *Science,* **288**: 1379-1384.

Nesbitt, M. N., and Francke, U. (1973) A system of nomenclature for banding patterns of mouse chromosomes. *Chromosoma,* **41**: 145-158.

Nichols, J., and Gardner, R. L. (1984) Heterogeneous differentiation of external cells in individual isolated early mouse inner cell masses in culture. *J. Embryol. exp. Morphol.*, **80**: 225-240.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell,* **95**: 379-91.

Niswander, L., and Martin, G. R. (1992) *Fgf-4* expression during gastrulation, myogenesis, limb and tooth development in the mouse. *Development,* **114**: 755-68.

Niwa, H., Miyazaki, J., and Smith, A. G. (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Gen.*, **24**: 372-376.

Oegema, K., Wiese, C., Martin, O. C., Milligan, R. A., Iwamatsu, A., Mitchison, T. J., and Zheng, Y. (1999) Characterisation of two related *Drosophila* γ-tubulin complexes that differ in their ability to nucleate microtubules. *J. Cell Biol.,* **144**: 721-733.

Orr-Urtreger, A., Givol, D., Yayon, A., Yarden, Y., and Lonai, P. (1991) Developmental expression of two murine fibroblast growth factor receptors, *flg* and *bek. Development,* **113**: 1419-1434.

Palmieri, S. L., Peter, W., Hess, H., and Schöler, H. R. (1994) Oct4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Dev. Biol.,* **166**: 259-267.

Papaioannou, V. E., McBurney, M. W., Gardner, R. L., and Evans, M. J. (1975) Fate of teratocarcinoma cells injected into early mouse embryos. *Nature,* **258**: 70-73.

Papaioannou, V. E., and Rossant, J. (1983) Effects of the embryonic environment on proliferation and differentiation of embryonal carcinoma cells. *Cancer Surv.*, **2**: 165-181.

Parameswaran, M., and Tam, P. P. L. (1995) Regionalisation of cell fate and morphogenetic movement of the mesoderm during mouse gastrulation. *Dev. Genet.,* **17**: 16-28.

Patil, N., Cox, D. R., Bhat, D., Faham, M., Myers, R. M., and Peterson, A. S. (1995) A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nat. Genet.,* **11**: 126-129.

Paul, W. (1999) Compartmentalisation of the immune system. *In:* Fundamental Immunology, Lippincott Raven: New York.

Pedersen, R. A., Spindle, A. I., and Wiley, L. M. (1977) Regeneration of endoderm by ectoderm isolated from mouse blastocysts. *Nature,* **270**: 435-437.

Pei, L., and Melmed, S. (1997) Isolation and characterisation of a pituitary tumor-transforming gene (PTTG). *Mol. Endocrinol.,* **11**: 433-441.

Pennica, D., Shaw, K. J., Swanson, T. A., Moore, M. W., Shelton, D. L., Zioncheck, K. A., Rosenthal, A., Taga, T., Paoni, N. F., and Wood, W. I. (1995) Cardiotrophin-1: Biological activities and binding to the leukemia inhibitory factor receptor/gp130 signalling complex. *J. Biol. Chem.,* **270**: 10915-10922.

Pera, M. F., Blasco Lafita, M. J., and Mills, J. (1987) Cultured stem-cells from human testicular teratomas: the nature of human embryonal carcinoma, and its comparison with two types of yolk-sac carcinoma. *Int. J. Cancer,* **40**: 334-343.

Piquet-Pellorce, C., Grey, L., Mereau, A., and Heath, J. K. (1994) Are LIF and related cytokines functionally equivalent? *Exp. Cell Res.,* **213**: 340-7.

Poirier, F., Chan, C.-T. J., Timmons, P. M., Robertson, E., Evans, M. J., and Rigby, P. W. J. (1991) The murine *H19* gene is activated during embryonic stem cell differentiation *in vitro* and at the time of implantation in the developing embryo. *Development,* **120**: 1105-1114.

Power, M.-A., and Tam, P. P. L. (1993) Onset of gastrulation, morphogenesis and somitogenesis in mouse embryos displaying compensatory growth. *Anat. Embryol.,* **187**: 493-504.

Quinlan, G. A., Williams, E. A., Tan, S. S., and Tam, P. P. L. (1995) Neuroectodermal fate of epiblast cells in the distal region of the mouse egg cylinder: implication for body plan organisation during early embryogenesis. *Development,* **121**: 87-98.

Rands, G. F. (1986) Size regulation in the mouse embryo. The development of quadruple aggregates. *J. Embryol. exp. Morphol.,* **94**: 139-148.

Rappolee, D. A., Basillico, C., Patel, Y., and Werb, Z. (1994) Expression and function of FGF-4 in peri-implantation in mouse embryos. *Development,* **120**: 2259-2269.

Rathjen, J., Lake, J., Bettess, M. D., Washington, J. M., Chapman, G., and Rathjen, P. D. (1999) Formation of a primitive ectoderm like cell population, EPL cells, from ES cells in response to biologically derived factors. *J. Cell Sci.,* **112**: 601-612.

Rathjen, P. D., Lake, J., Whyatt, L. M., Bettess, M. D., and Rathjen, J. (1998) Properties and uses of embryonic stem cells: prospects for application to human biology and gene therapy. *Reprod. Fertil. Dev.,* **10**: 31-47.

Rathjen, P. D., Toth, S., Willis, A., Heath, J. K., and Smith, A. G. (1990a) Differentiation inhibiting activity/leukaemia inhibitory factor (DIA/LIF) is produced in matrix-associated and diffusible forms generated by alternate promoter usage. *Cell,* **62**: 1105-1114.

Rathjen, P. D., Nichols, J., Toth, S., Edwards, D. R., Heath, J. K., and Smith, A. G. (1990b) Developmentally programmed induction of differentiating inhibitory activity and the control of stem cell populations. *Genes & Dev.,* **4**: 2308-2318.

Remiszewski, J. (1999) The role of *Gbx2* in murine embryonic development. *PhD. Thesis.* Department of Biochemistry, University of Adelaide.

Resnick, J. L., Bixler, L. S., Cheng, L., and Donovan, P. J. (1992) Long-term proliferation of mouse primordial germ cells in culture. *Nature,* **359**: 550-551.

Riethmacher, D., Brinkmann, V., and Birchmeier, C. (1995) A targeted mutation in the mouse E-cadherin gene results in defective perimplantation development. *Proc. Natl. Acad. Sci. USA,* **92**: 855-859.

Robb, L. (1997) Haematopoiesis: Origin pinned down at last. *Curr. Biol.,* **7**: R10-R12.

Robbins, J., Dilworth, S. M., Laskey, R. A., and Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence. *Cell,* **64**: 615-623.

Robertson, E. J. (1987) Embryo-derived stem cell lines. *In:* Teratocarcinomas and embryonic stem cells: a practical approach (E. J. Robertson, ed.), IRL press: Oxford, pg. 71-112.

Rodda, S. J. (1998) Isolation and characterisation of mouse *L17*: a marker of the pluripotent epiblast. *Honours Thesis*. Department of Biochemistry, University of Adelaide.

Rogers, M. B., Hosler, B. A., and Gudas, L. J. (1991) Specific expression of a retinoic acid-regulated, zinc-finger gene, *Rex-1*, in preimplantation embryos, trophoblast and spermatocytes. *Development,* **113**: 815-24.

Rosen, B., and Beddington, R. S. P. (1993) Whole-mount *in situ* hybridisation in the mouse embryo: gene expression in three dimensions. *TIG,* **9**: 162-167.

Rosenquist, T. A., and Martin, G. R. (1995) Visceral endoderm-1 (VE-1): an antigen marker that distinguishes anterior from posterior embryonic visceral endoderm in the early post-implantation mouse embryo. *Mech Dev,* **49**: 117-21.

Rosner, M. H., Vigano, A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W. J., and Staudt, L. M. (1990) A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature,* **345**: 686-692.

Rossant, J. (1975) Investigation of the determinative state of the mouse inner cell mass. II. The fate of isolated inner cell masses transferred to the oviduct. *J. Embryol. exp. Morphol.,* **33**: 991-1001.

Rossant, J. (1977) Cell commitment in early rodent development. *In:* Development in Mammals (M. H. Johnson, ed.), Elsevier: Amsterdam. pg. 119-150.

Rossant, J. (1993) Immortal Germ Cells? *Curr. Biol.,* **3**: 47-49.

Rossant, J., and Ofer, L. (1977) Properties of extra-embryonic ectoderm isolated from postimplantation mouse embryos. *J. Embryol. exp. Morphol.,* **39**: 183-194.

Rossant, J., and Tamura-Lis, W. (1979) Potential of isolated mouse inner cell masses to form trophectoderm derivatives *in vivo*. *Dev. Biol.,* **70**: 255-61.

Rossant, J., and Vijh, K. M. (1980) Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. *Dev. Biol.,* **76**: 475-482.

Rudnicki, M. A., and McBurney, M. W. (1987) Cell culture methods and induction of differentiation of embryonal carcinoma cell lines. *In:* Teratocarcinomas and embryonic stem cells (E. J. Robertson, ed.), IRL Press: Oxford, pg. 19-49.

Russell, C. (1999) A role for *Psc1* in nucleocytoplasmic trafficking. *Honours Thesis*. Department of Biochemistry, University of Adelaide.

Salas, P. J. I. (1999) Insoluble γ-tubulin-containing structures are anchored to the apical network of intermediate filaments in polarised CACO-2 epithelial cells. *J. Cell Biol.,* **146**: 645-657.

Scherer, M. A. (1999) Molecular characterisation of L17: a novel developmentally regulated member of the CP2-like family of transcription factors. *Honours Thesis*. Department of Biochemistry, University of Adelaide.

Schiebel, E. (2000) γ-tubulin complexes: binding to the centrosome, regulation and microtubule nucleation. *Curr. Opin. Cell Biol.,* **12**: 113-118.

Schöler, H. R., Dressler, G. R., Balling, R., Rohdewohld, H., and Gruss, P. (1990a) Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *EMBO J,* **9**: 2185-95.

Schöler, H. R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990b) New type of POU domain in germ line-specific protein Oct-4. *Nature,* **344**: 435-9.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. USA,* **95**: 5857-5864.

Schultz, R. M. (1986) Molecular aspects of mammalian oocyte growth and maturation. *In:* Experimental Approaches to Mammalian Embryonic Development (J. Rossant and R. A. Pedersen, eds.), Cambridge University Press: London pg. 195-237.

Schulz, T. C. (1996) A system for the isolation of markers for subpopulations of murine pluripotent cells. *PhD. Thesis*. Department of Biochemistry, University of Adelaide.

Sefton, M., Johnson, M. H., and Clayton, L. (1992) Synthesis and phosphorylation of uvomorulin during mouse early development. *Development,* **115**: 313-318.

Shamblott, M. J., Axelman, J., Wang, S., Bugg, E. M., Littlefield, J. W., Donovan, P. J., Blumenthal, P. D., Huggins, G. R., and Gearhart, J. D. (1998) Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc. Natl. Acad. Sci. USA,* **95**: 13726-13731.

Sharan, S. K., Morimatsu, M., Albrecht, U., Lim, D., Regel, E., Dinh, C., Sands, A., Eichele, G., Hasty, P., and Bradley, A. (1997) Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking *Brca2*. *Nature,* **386**: 804-810.

Shearwin-Whyatt, L. M., Remiszewski, J., and Rathjen, P. D. (1999) Gene targeting. *In:* Encyclopedia of Molecular Biology (T. E. Creighton, ed.), John Wiley and Sons: New York, pg. 987-993.

Shen, M. M., and Leder, P. (1992) Leukemia inhibitory factor is expressed by the preimplantation uterus and selectively blocks primitive ectoderm formation *in vitro. Proc. Natl. Acad. Sci. USA,* **89**: 8240-8244.

Simeone, A., Acampora, D., Mallamaci, A., Stornaiuolo, A., D'Apice, M., Nigro, R., and Boncinelli, E. (1993) A vertebrate gene related to *orthodenticle* contains a homeobox of the *bicoid* class and demarcates anterior neuroectoderm in the gastrulating mouse embryo. *EMBO J.*, **12**: 2735-2747.

Smith, A. G. (1991) Culture and Differentiation of Embryonic Stem Cells. *J. Tiss. Cult. Meth.*, **13**: 89-94.

Smith, A. G. (1992) Mouse embryo stem cells: Their identification, propagation and manipulation. *Sem. Cell Biol.*, **3**: 383-399.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., and Rogers, D. (1988) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature,* **336**: 688-692.

Snow, M. H. L. (1976) Embryo growth during the immediate postimplantation period. *In:* Embryogenesis in mammals (K. Elliot and M. O'Connor, ed.), Elsevier: Amsterdam, pg. 53-70.

Snow, M. H. L. (1977) Gastrulation in the mouse: Growth and regionalisation of the epiblast. *J. Embryol. exp. Morphol.*, **42**: 293-303.

Snow, M. H. L., and Tam, P. L. L. (1979) Is compensatory growth a complicating factor in mouse teratology? *Nature,* **279**: 557-559.

Solter, D., and Knowles, B. (1978) Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *Proc. Natl. Acad. Sci. USA,* **75**: 5565-5569.

Spyropoulos, D. D., and Cappechi, M. R. (1994) Targeted disruption of the even-skipped gene, *evx1*, causes early post-implantation lethality of the mouse conceptus. *Genes & Dev.*, **8**: 1949-1961.

Stephens, L. E., Sutherland, A. E., Klimanskaya, I. V., Andrieux, A., Meneses, J., Pedersen, R. A., and Damsky, C. H. (1995) Deletion of β1 integrins in mice results in inner cell mass failure and peri-implantation lethality. *Genes & Dev.,* **9**: 1883-1895.

Stewart, C. L., Kaspar, P., Brunet, L. J., Bhatt, H., Gadi, I., Kontgen, F., and Abbondanzo, S. J. (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature,* **359**: 76-79.

Stewart, T. A., and Mintz, B. (1981) Successive generations of mice produced from an established culture line of euploid teratocarcinoma cells. *Proc. Natl. Acad. Sci. USA,* **78**: 6314-6318.

Strickland, S., and Mahdavi, V. (1978) The induction of differentiation in teratocarcinoma stem cells by retinoic acid. *Cell,* **15**: 393-355.

Strickland, S., Smith, K. K., and Marotti, K. R. (1980) Hormonal induction of differentiation in teratocarcinoma stem cells: generation of parietal endoderm by retinoic acid and dibutyryl cAMP. *Cell,* **21**: 347-355.

Strübing, C., Ahnert-Hilger, G., Shan, J., Wiedenmann, B., Hescheler, J., and Wobus, A. M. (1995) Differentiation of pluripotent embryonic stem cells into the neuronal lineage *in vitro* gives rise to mature inhibitory and excitatory neurons. *Mech. Dev.,* : 275-287.

Sutherland, A. E., Calarco, P. G., and Damsky, C. H. (1993) Developmental regulation of integrin expression at the time of implantation in the mouse embryo. *Development,* **119**: 1175-1186.

Suzuki, A., del la Pompa, J. L., Hakem, R., Elia, A., Yoshida, R., Mo, R., Nishina, H., Chuang, T., Wakeham, A., Itie, A., Koo, W., Billia, P., Ho, A., Fukumoto, M., Hui, C.-C., and Mak, T. W. (1997) *Brca2* is required for embryonic cellular proliferation in the mouse. *Genes & Dev.,* **11**: 1242-1252.

Suzuki, A., Thies, R. S., Yamaji, N., Song, J. J., Wozley, J. M., Murakami, K., and Ueno, N. (1994) A truncated bone morphogenic protein receptor affects dorsal-ventral patterning in the *Xenopus* embryo. *Proc. Natl. Acad. Sci. USA,* **91**: 10255-10259.

Takahashi, M., Shibata, H., Shimakawa, M., Miyamoto, M., Mukai, H., and Ono, Y. (1999) Characterisation of a novel giant scaffolding protein, CG-NAP, that anchors multiple signaling enzymes to centrosome and the golgi apparatus. *J. Biol. Chem.,* **274**: 17267-17274.

Takeda, K., Noguchi, K., Shi, W., Tanaka, T., Matsumoto, M., Yoshida, N., Kishimoto, T., and Akira, S. (1997) Targeted disruption of the mouse *Stat3* gene leads to early embryonic lethality. *Proc. Natl. Acad. Sci. USA,* **94**: 3801-3804.

Tam, P. P. L. (1988) The post-implantation development of mitomycin C-treated mouse blastocysts. *Teratology,* **37**: 205-212.

Tam, P. P. L. (1989) Regionalisation in the mouse embryonic ectoderm: allocation of prospective ectodermal tissues during gastrulation. *Development,* **107**: 55-67.

Tam, P. P. L., and Beddington, R. S. P. (1987) The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development,* **99**: 109-126.

Tam, P. P. L., and Beddington, R. S. P. (1992) Establishment and organisation of germ layers in the gastrulating mouse embryo. *In*: Postimplantation development in the mouse. *Ciba Found. Symp.,* Vol. 165, Wiley: Chichester, pg. 27-49.

Tam, P. P. L., and Behringer, R. R. (1997) Mouse gastrulation: the formation of a mammalian body plan. *Mech. Dev.,* **68**: 3-25.

Tam, P. P. L., Parameswaran, M., Kinder, S. L., and Weinberger, R. P. (1997) The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. *Development,* **124**: 1631-1642.

Tam, P. P. L., and Snow, M. H. L. (1981) Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J. Embryol. exp. Morphol.,* **64**: 133-147.

Tam, P. P. L., and Zhou, S. X. (1996) The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev. Biol.,* **178**: 124-132.

Tanaka, S., Kunath, T., Hadjantonakis, A.-K., Nagy, A., and Rossant, J. (1998) Promotion of trophoblast stem cell proliferation by FGF4. *Science,* **282**: 2072-2075.

Tarkowski, A. K., and Wroblewska, J. (1967) Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J. Embryol. exp. Morphol.,* **18**: 155-180.

Tatusova, T. A., and Madden, T. L. (1999) BLAST 2 sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.,* **174**: 247-250.

Taylor, K. D., and Pilkó, L. (1987) Patterns of mRNA prevalence and expression of B1 and B2 transcripts in early mouse embryos. *Development,* **101**: 877-892.

Telford, N. A., Watson, A. J., and Schultz, G. A. (1990) Transcription from maternal to embryonic control in early mammalian development: A comparison of several species. *Mol. Reprod. Dev.,* **26**: 90-100.

Thomas, P. Q., and Beddington, R. S. P. (1996) Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. *Curr. Biol.,* **6**: 1487-1496.

Thomas, P. Q., Brown, A., and Beddington, R. S. P. (1998) Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development,* **125**: 85-94.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positive-specific gap penalties and weight matrix choice. *Nuc. Acids Res.*, **22**: 4673-4680.

Thomson, J. A., Itskovitzeldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998) Embryonic stem cell lines derived from human blastocysts. *Science,* **282**: 1145-1147.

Threadgill, D. W., Dlugosz, A. A., Hansen, L. A., Tennenbaum, T., Lichti, U., Yee, D., LaMantia, C., Mourton, T., Herrup, K., Harris, R. C., Barnard, J. A., Yuspa, S. H., Coffey, R. J., and Magnuson, T. (1995) Targeted disruption of mouse EGF Receptor: effect of genetic background on mutant phenotype. *Science,* **269**: 230-234.

Tóth, A., Ciosk, R., Uhlmann, F., Galova, M., Schleiffer, A., and Nasmyth, K. (1999) Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes & Dev.*, **13**: 320-333.

Uhlmann, F., Lottspeich, F., and Nasmyth, K. (1999) Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature,* **400**: 37-42.

Uzawa, S., Samejima, I., Hirano, T., Tanaka, K., and Yanagida, M. (1990) The fission yeast cut1^{+} gene regulates spindle pole body duplication and has homology to the budding yeast ESP1 gene. *Cell,* **62**: 913-925.

van den Eijnden-van Raaij, A. J. M., van Achterberg, T. A. E., van der Kruijssen, C. M. M., Piersma, A. H., Huylebroeck, D., de Laat, S. W., and Mummery, C. L. (1991) Differentiation of aggregated murine P19 embryonal carcinoma cells is induced by a novel visceral endoderm-specific FGF-like factor and inhibited by Activin A. *Mech. Dev.*, **33**: 157-166.

Varlet, I., Collingnon, J., and Robertson, E. J. (1997) *nodal* expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. *Development,* **124**: 1033-1044.

Vestweber, D., Gossler, A., Boller, K., and Kemler, R. (1987) Expression and distribution of cell adhesion molecule Uvomorulin in mouse preimplantation embryos. *Dev. Biol.,* **124**: 451-456.

Voyle, R. B. (1999) Mechanisms of intracellular and extracellular cytokine production from the human leukaemia inhibitory factor gene. *PhD. Thesis*. Department of Biochemistry, University of Adelaide.

Wiley, L. M. (1987) Trophectoderm: the first epithelium to develop in the mammalian embryo. *Scanning Microsc.,* **2**: 417-426.

Wilkinson, D. G., Bhatt, S., and Herrmann, B. G. (1990) Expression pattern of the mouse *T* gene and its role in mesoderm formation. *Nature,* **343**: 657-659.

Williams, R. L., Hilton, D. J., Pease, S., Wilson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988) Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature,* **336**: 684-692.

Wilson, R. *et al.* (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans. *Nature,* **368**: 32-8.

Winnier, G., Blessing, M., Labosky, P. A., and Hogan, B. L. M. (1995) Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes & Dev.,* **9**: 2105-2116.

Wright, N. A. (1997) Stem cell repertoire in the intestine. *In:* Stem cells (C. S. Potten, ed.), Academic press: London, pg. 315-330.

Yamaguchi, T. P., Conlon, R. A., and Rossant, J. (1992) Expression of the fibroblast growth factor receptor FGF-R1/*flg* during gastrulation and segmentation in the mouse embryo. *Dev. Biol.,* **152**: 75-88.

Yamaguchi, T. P., Harpal, K., Henkemeyer, M., and Rossant, J. (1994) *fgfr-1* is required for embryonic growth and mesodermal patterning during mouse gastrulation. *Genes & Dev.,* **8**: 3032-3044.

Yamamoto, A., Guacci, V., and Koshland, D. (1996) Pds1p is required for faithful execution of anaphase in the yeast, Saccharomyces cerevisiae. *J. Cell Biol.,* **133**: 85-97.

Yang, X., Li, C., Xu, X., and Deng, C. (1998) The tumor suppressor SMAD4/DPC4 is essential for epiblast proliferation and mesoderm induction in mice. *Proc. Natl. Acad. Sci. USA,* **95**: 3667-3672.

Yeom, Y. I., Ha, H. S., Balling, R., Schöler, H. R., and Artzt, K. (1991) Structure, expression and chromosomal location of the Oct-4 gene. *Mech. Dev.,* **35**: 171-9.

Yoshinaga, K., and Adams, C. E. (1966) Delayed implantation in the spayed, progesterone treated adult mouse. *J. Reprod. Fertil.,* **qw**: 593-595.

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995) Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes & Dev.,* **9**: 2635-2645.

Zhang, X., Horwitz, G. A., Prezant, T. R., Valentini, A., Nakashima, M., Bronstein, M. D., and Melmed, S. (1999) Structure, expression and function of human pituitary tumor-transforming gene (PTTG). *Mol. Endocrinol.,* **13**: 156-166.

Zheng, Y., Wong, M. L., Alberts, B., and Mitchison, T. (1995) Nucleation of microtubule assembly by a γ-tubulin-containing ring complex. *Nature,* **378**: 578-583.

Ziomek, C. A., Johnson, M. H., and Handyside, A. H. (1982) The developmental potential of mouse 16-cell blastomeres. *J. Exp. Zool.*, **221**: 345-355.

Zou, H., McGarry, T. J., Bernal, T., and Kirschner, M. W. (1999) Identification of a vertebrate sister-chromatid separation inhibitor involved in transformation and tumorogenesis. *Science,* **285**: 418-422.

Erratum

Chapter 1: Introduction:

Page 3, paragraph 1: final citation should read: Gilbert & Solter, 1985; Taylor & Pilko, 1987; Kidder, 1992)

Page 17, paragraph 2. The following sentence should be added at the end of the paragraph: Finally, the homeobox gene *Hex* also exhibits regionalised expression within visceral endoderm (at 5.5 d.p.c.) being expressed in the distal tip of the egg cylinder. Lineage analysis has demonstrated that these cells move unilaterally to assume an anterior position while continuing to express *Hex*. The primitive streak forms on the opposite side of the egg cylinder approximately 24 hours after the initial anterior movement, suggesting that *Hex* expression is the earliest molecular anteroposterior asymmetry in the mouse embryo (Thomas *et al.* 1998).

Chapter 2: Materials & Methods

Page 69: add a final paragraph: For peri-implantation in situ hybridisations, 8 to 10 embryos were routinely used in embryonic *in situ* hybridisation experiments. The exceptions were (i) 3.5 d.p.c. blastocysts in which 15-20 blastocysts were used to ensure at least 10 embryos were recovered from processing baskets for staining, and (ii) 5.0 d.p.c. embryos, which by nature are technically difficult to dissect and if insufficient numbers were obtaining during a dissection, then 3-5 embryos were analysed in baskets also containing 5.25 d.p.c. embryos. All *in situ* hybridisation experiments were repeated, often several times in the cases in which variable expression patterns were observed.

Chapter 3:

Page 82, paragraph 2, third sentence should read: In addition to the pluripotent cell expression, *Fgf5* transcripts were detected at 4.75 d.p.c. in the primitive endoderm in 8 out of 21 embryos.

Page 83, paragraph 1, third sentence should read: *L17* expression appeared to be downregulated within pluripotent cells at around 4.75 d.p.c. as expression of this transcript was variable at this stage, with 9 out of 22 embryos expressing *L17*.

Page 84, paragraph 1, second sentence should read: *Psc1* transcripts were also expressed throughout the epiblast cells at 4.75 d.p.c. and detectable at 5.0 d.p.c. in the epiblast bud, immediately prior to proamniotic cavity formation.

Page 84, paragraph 2, fourth sentence should read: Variable *K7* expression was observed at 4.75 d.p.c. with 11 out of 20 embryos being positive for *K7* expression.

Figure 3.10 legend: Panel A) Rex1 expression; B) L17 expression; C) Psc1 expression; D) FGF 5 expression

Chapter 5:

Page 121, paragraph 2, fourth sentence should read: The first EGFP-K7 localisation pattern, observed in 75% of cells (69 out of 92 cells analysed) is illustrated by the fluorescing cell on the left of Figure 5.5E.

Page 121, paragraph 2, sixth sentence should read: The second EGFP-K7 localisation pattern, observed in 25% of cells (23 out of 92 cells anlysed), and illustrated by the fluorescent cell on the right hand side of Figure 5.5E, was associated with a punctuate distribution, concentrated in the peri-nuclear region.

Bibliography:

Gilbert S.F. and Solter D. (1985) Onset of paternal and maternal Gpi-1 expression in preimplantation mouse embryos. *Dev Biol.* **109**:515-7.

Thomas P.Q., Brown A, Beddington R.S. (1998). *Hex*: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development.* **25**:85-94.