



A System for the Isolation of Markers for Subpopulations of Murine Pluripotent Cells

A thesis submitted to the University of Adelaide
for the degree of Doctor of Philosophy

by

Thomas Carl Schulz, B.Sc. (Hons)

Department of Biochemistry
University of Adelaide
Adelaide, South Australia

November 1996

TABLE OF CONTENTS

| | |
|---|----|
| THESIS SUMMARY | i |
| STATEMENT | iv |
| ACKNOWLEDGMENTS | v |
| CHAPTER 1: GENERAL INTRODUCTION | |
| 1.1 Development..... | 1 |
| 1.1.1 The Study of Mammalian Development | 1 |
| 1.1.2 The Mouse as a Model for Mammalian Development..... | 1 |
| 1.2 Preimplantation Development | 2 |
| 1.2.1 Fertilisation and Cleavage Stage Development..... | 2 |
| 1.2.2 Compaction and Polarisation..... | 3 |
| 1.2.3 Blastulation and Implantation | 3 |
| 1.3 Peri-implantation Development | 4 |
| 1.4 Postimplantation Development..... | 5 |
| 1.4.1 Proamniotic Cavity and Primitive Ectoderm Formation | 5 |
| 1.4.2 Gastrulation..... | 6 |
| 1.4.3 Partitioning of the Primordial Germ Cell Lineage | 6 |
| 1.5 Pluripotent Cell Populations in Early Mouse Development..... | 7 |
| 1.5.1 The Heterogeneous Nature of Murine Pluripotent Cell Populations..... | 8 |
| 1.5.2 Gene Expression in Murine Pluripotent Cells..... | 10 |
| 1.6 Embryonic Stem (ES) Cells as an <i>in vitro</i> Model of Early Development..... | 12 |
| 1.6.1 Isolation of ES cells..... | 12 |
| 1.6.2 The Relationship of ES Cells to Embryonic Pluripotent Cells: a Model of the Inner Cell Mass | 13 |
| 1.6.3 Maintenance of ES Cells: the Undifferentiated State | 14 |
| 1.6.4 Differentiation of ES Cells and <i>in vitro</i> Models of Development ... | 14 |
| 1.6.5 Spontaneous and Chemical Induction of ES Cell Differentiation... | 15 |
| 1.6.6 Differentiation of ES Cells as Embryoid Bodies..... | 15 |
| 1.6.7 Expansion of the Pluripotent Cell Model of Early Development: Lineage-specific Differentiation of ES Cells to X Cells in Response to Soluble Biological Factors | 16 |
| 1.6.8 X Cells as a Model for Primitive Ectoderm | 17 |
| 1.6.9 The Significance of X Cell Formation for <i>in vitro</i> Models of Pluripotent Cell Development..... | 18 |

| | | |
|-------|--|----|
| 1.7 | Generation and Application of Pluripotent Cell Specific Markers..... | 19 |
| 1.7.1 | Mutational Approaches to Define Markers for Pluripotent Cell Types | 21 |
| 1.7.2 | Generating Pluripotent Cell Markers by Homology Based Approaches..... | 21 |
| 1.7.3 | Cloning of Pluripotent Cell Markers from Embryonic Tissue | 22 |
| 1.7.4 | Cloning of Pluripotent Cell Markers from Cell Lines <i>in vitro</i> | 23 |
| 1.7.5 | Identifying Pluripotent Cell Markers by Gene Traps..... | 23 |
| 1.7.6 | Cloning of Pluripotent Cell Markers by Differential Display Polymerase Chain Reaction | 24 |
| 1.8 | Aims and Approaches..... | 26 |

CHAPTER 2: MATERIALS AND METHODS

| | | |
|--------|--|----|
| 2.1 | Abbreviations | 28 |
| 2.2 | Materials | 29 |
| 2.2.1 | Chemicals and Reagents..... | 29 |
| 2.2.2 | Radiochemicals | 29 |
| 2.2.3 | Kits..... | 29 |
| 2.2.4 | Enzymes..... | 30 |
| 2.2.5 | Buffers and Solutions..... | 30 |
| 2.2.6 | Plasmid Vectors | 31 |
| 2.2.7 | Riboprobe Templates | 31 |
| 2.2.8 | Oligonucleotides..... | 32 |
| 2.2.9 | Bacterial Strains | 33 |
| 2.2.10 | Bacterial Growth Media | 34 |
| 2.2.11 | Tissue Culture Cell Lines and Media..... | 34 |
| 2.2.12 | DNA Markers..... | 35 |
| 2.2.13 | Miscellaneous Materials | 35 |
| 2.3 | Molecular Methods | 35 |
| 2.3.1 | Restriction Endonuclease Digestion of DNA..... | 35 |
| 2.3.2 | Agarose Gel Electrophoresis..... | 35 |
| 2.3.3 | Polyacrylamide Gel Electrophoresis | 35 |
| 2.3.4 | Purification of Linear DNA Fragments | 36 |
| 2.3.5 | Agarose Minigels for DNA Southern Blot | 36 |
| 2.3.6 | Preparation of Vector DNA | 36 |
| 2.3.7 | DNA Ligation Reactions..... | 37 |
| 2.3.8 | Preparation of Competent Cells..... | 37 |
| 2.3.9 | Transformation of Competent Bacterial Cells..... | 37 |
| 2.3.10 | Rapid Small Scale Preparation of DNA (Mini-prep) | 37 |

| | | |
|--------|---|----|
| 2.3.11 | Rapid Large Scale Preparation of DNA (50 ml Midi-prep)..... | 38 |
| 2.3.12 | Generation of Sequencing Subclones..... | 38 |
| 2.3.13 | Double Stranded Sequencing of Plasmid DNA..... | 39 |
| 2.3.14 | Library Screening..... | 39 |
| 2.3.15 | Lambda Zapping..... | 39 |
| 2.3.16 | Reverse Transcription..... | 40 |
| 2.3.17 | Reverse Transcription PCR (RT/PCR)..... | 40 |
| 2.3.18 | Differential Display PCR..... | 40 |
| 2.3.19 | Rapid Amplification of cDNA Ends (RACE/PCR)..... | 41 |
| 2.3.20 | Isolation of Cytoplasmic RNA from Cultured Cells..... | 41 |
| 2.3.21 | Isolation of RNA from Tissue Samples..... | 42 |
| 2.3.22 | Isolation of (A) _n ⁺ RNA..... | 42 |
| 2.3.23 | RNA Gels for Northern Blot Analysis..... | 42 |
| 2.3.24 | Northern Blot Transfer..... | 43 |
| 2.3.25 | Radioactive DNA Probes..... | 43 |
| 2.3.26 | Radioactive RNA Probes..... | 43 |
| 2.3.27 | Hybridisation of Radioactive Probes to Nylon and Nitrocellulose Filters..... | 43 |
| 2.3.28 | Ribonuclease Protection Analysis..... | 44 |
| 2.3.29 | Digoxigenin (DIG) Labelled RNA Probes..... | 44 |
| 2.3.30 | <i>in situ</i> Hybridisation of Cell Monolayers..... | 44 |
| 2.3.31 | Embryonic <i>in situ</i> Hybridisation..... | 45 |
| 2.3.32 | Mouse Embryo Powder..... | 46 |
| 2.3.33 | Antibody Preblocking..... | 46 |
| 2.3.34 | Containment Facilities..... | 46 |
| 2.3.35 | Phosphorimager Analysis, Autoradiograph Scanning and Image Manipulation..... | 47 |
| 2.3.36 | Sequencing Software and Database Searches..... | 47 |
| 2.4 | Tissue Culture Methods..... | 47 |
| 2.4.1 | Maintenance of ES Cells..... | 47 |
| 2.4.2 | Differentiation of ES Cells..... | 47 |
| 2.5 | Embryological Methods..... | 49 |
| 2.5.1 | Mouse Strains and Matings..... | 49 |
| 2.5.2 | Embryo Processing Baskets..... | 49 |
| 2.5.3 | Embryo Isolation and Dissection..... | 49 |
| 2.5.4 | Embryo Fixation and Dehydration..... | 50 |
| 2.5.5 | Animal Manipulations..... | 50 |

CHAPTER 3: A METHOD FOR THE IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES DURING PLURIPOTENT CELL PROGRESSION *IN VITRO*

| | | |
|----------|---|----|
| 3.1 | Introduction | 51 |
| 3.1.1 | <i>In vitro</i> Models of the Inner Cell Mass and Primitive Ectoderm | 51 |
| 3.1.2 | Differential Display Polymerase Chain Reaction and Gene Class Enrichment..... | 52 |
| 3.1.3 | Homeobox Genes and Expression in Pluripotent Cell Subpopulations | 52 |
| 3.1.4 | Chapter Summary | 54 |
| 3.2 | Modification and Trial of Differential Display PCR Conditions | 54 |
| 3.2.1 | Differential Display PCR Modifications | 54 |
| 3.2.1i | Amplification Profiles With 3'-hp | 55 |
| 3.2.1ii | Assessment of Template Preparation Methods | 55 |
| 3.2.1iii | 3'-hp Display and Detection of Differential Gene Expression | 56 |
| 3.2.2 | 3'-hp DDPCR Reamplification and Cloning Procedures | 58 |
| 3.2.3 | Sequence Analysis of 3'-hp DDPCR Products | 59 |
| 3.2.4 | Discussion of 3'-hp DDPCR Modifications..... | 60 |
| 3.3 | DDPCR Screen for ES Cell Restricted Pluripotent Cell Markers | 62 |
| 3.3.1 | Identification of Early Pluripotent Cell Specific Markers..... | 63 |
| 3.3.2 | Expression Profile of ES and Early Pluripotent Cell Markers | 63 |
| 3.3.3 | Summary of 3'-hp DDPCR Screen for Early Pluripotent Cell Markers..... | 65 |
| 3.4 | Isolation and Cloning of ES and Early Pluripotent Cell Markers | 66 |
| 3.4.1 | Validation of A03/360 and B04/400 Expression Patterns | 67 |
| 3.4.2 | Sequence of A03/360 and B04/400 Markers..... | 69 |
| 3.4.3 | <i>In situ</i> Analysis of A03/360 and B04/400 Expression in ES Cell Monolayers | 70 |
| 3.5 | Discussion..... | 71 |

CHAPTER 4: EXPRESSION ANALYSIS OF EARLY PLURIPOTENT CELL MARKERS *IN VITRO* AND DURING EARLY EMBRYOGENESIS

| | | |
|-------|---|----|
| 4.1 | Introduction | 75 |
| 4.2 | Expression of A03/360 and B04/400 <i>in vitro</i> | 76 |
| 4.2.1 | <i>In situ</i> Hybridisation Analysis of A03/360 and B04/400 Expression During ES Cell Differentiation..... | 77 |

| | | |
|--------|--|----|
| 4.2.2 | <i>In situ</i> Hybridisation Analysis of A03/360 and B04/400 Expression During the ES to X Cell Transition..... | 78 |
| 4.3 | Analysis of A03/360 and B04/400 Expression During Early Mouse Embryogenesis..... | 79 |
| 4.3.1 | Whole Mount <i>in situ</i> Hybridisation Analysis of <i>Oct-4</i> Expression During Early Mouse Development | 80 |
| 4.3.1i | Specificity of Embryonic <i>in situ</i> Hybridisation Analysis..... | 81 |
| 4.3.2 | <i>In situ</i> Hybridisation Analysis of <i>Icm1</i> (A03/360) Expression During Early Mouse Embryogenesis..... | 82 |
| 4.3.3 | <i>In situ</i> Hybridisation Analysis of <i>Psc1</i> (B04/400) Expression During Early Mouse Embryogenesis..... | 83 |
| 4.4 | Discussion..... | 86 |
| 4.4.1 | The Expression of <i>Icm1</i> and <i>Psc1</i> in Pluripotent Cell Subpopulations <i>in vitro</i> | 86 |
| 4.4.2 | <i>Icm1</i> and <i>Psc1</i> Expression and Pluripotent Cell Subpopulations in the Mouse Embryo..... | 88 |
| 4.4.3 | Correlation of <i>Icm1</i> and <i>Psc1</i> Expression to Embryonic Development | 90 |
| 4.4.4 | Embryonic Equivalents of ES Cells and X Cells | 92 |
| 4.4.5 | Nomenclature of Pluripotent Cell Subpopulations | 93 |
| 4.4.6 | Potential for Identifying Pluripotent Cell Subpopulations..... | 93 |

CHAPTER 5: ISOLATION OF *Psc1* cDNA CLONES AND ANALYSIS OF *Psc1* EXPRESSION DURING LATER DEVELOPMENT

| | | |
|-------|--|-----|
| 5.1 | Introduction | 95 |
| 5.2 | Isolation and Sequence Analysis of <i>Psc1</i> cDNA Clones..... | 96 |
| 5.2.1 | Isolation of <i>Psc1</i> cDNA Clones | 96 |
| 5.2.2 | Sequence Analysis of <i>Psc1</i> cDNA Clones | 97 |
| 5.2.3 | Similarity of <i>Psc1</i> with CLEB0336.3 10 and HFBDS04 | 98 |
| 5.3 | <i>Psc1</i> Expression in Foetal and Adult Murine Tissues..... | 99 |
| 5.4 | Discussion..... | 100 |
| 5.4.1 | Analysis of <i>Psc1</i> Transcripts..... | 100 |
| 5.4.2 | Cellular Functions of <i>Psc1</i> and Nuclear Speckle Domains..... | 100 |
| 5.4.4 | Function of <i>Psc1</i> During Development | 102 |
| 5.4.5 | Future Analysis of the Cellular Function of <i>Psc1</i> | 103 |

CHAPTER 6: FINAL DISCUSSION

| | | |
|---------|---|-----|
| 6.1 | A System for the Isolation of Markers for Subpopulations of Murine Pluripotent Cells..... | 104 |
| 6.2 | Pluripotent Cell Subpopulations or “States” During Early Embryogenesis | 104 |
| 6.2.1 | Identification of Pluripotent Cell Subpopulations or “States” During Early Embryogenesis | 105 |
| 6.3 | Future Work and Application of <i>Icm1</i> and <i>Psc1</i> to the Analysis of Pluripotent Cell Development..... | 105 |
| 6.3.1 | Analysis of Mutations that Affect Pluripotent Cell Development ... | 106 |
| 6.3.2 | Identifying Developmental Signals During Embryogenesis..... | 106 |
| 6.3.3 | Analysis of Pluripotent Cell Development..... | 107 |
| 6.3.3i | Functional Analysis of Pluripotent Cell Development <i>in vitro</i> | 107 |
| 6.3.3ii | Mutational Analysis of <i>Icm1</i> and <i>Psc1</i> <i>in vivo</i> and Gene Function During Embryogenesis | 108 |
| 6.3.4 | Isolation and Application of <i>Icm1</i> and <i>Psc1</i> Promoters..... | 110 |

| | |
|--|------------|
| APPENDIX: AN UNUSUAL ARRANGEMENT OF 13 ZINC FINGERS IN THE VERTEBRATE GENE Z13..... | 111 |
|--|------------|

| | |
|-------------------------|------------|
| REFERENCES | 117 |
|-------------------------|------------|

THESIS SUMMARY

The coordinated regulation of pluripotent cell development is critical for the generation of extraembryonic tissues, the differentiated lineages of the embryo, and establishment of the basic body plan during mouse embryogenesis. Accumulating evidence points to considerable heterogeneity within the developing pluripotent cell pool. The generation of specific markers will be critical for the identification and analysis of the implied pluripotent cell subpopulations.

Mouse embryonic stem (ES) cells are a pluripotent cell type derived from the blastocyst inner cell mass and provide a system to investigate pluripotent cell biology *in vitro*. ES cells develop to a distinct pluripotent cell type *in vitro*, termed X cells, in response to MedII conditioned medium. Pluripotent cell types present during the development of ES cells to X cells are a model for the inner cell mass to primitive ectoderm transition *in vivo*, and a source of differentially expressed genes that could be exploited to identify subpopulations of pluripotent cells during early embryogenesis.

The general aim of this thesis was to develop methods for the identification of markers for pluripotent cell subpopulations in the developing mouse embryo.

A screen for ES cell markers was carried out, to identify transcripts that were differentially expressed between ES cells and X cells, to define the embryological equivalents of ES cells, and to investigate pluripotent cell heterogeneity during early development. A modified differential display polymerase chain reaction (DDPCR) system identified nine transcripts that were restricted to ES cells and early pluripotent cell types within the ES to X cell transition. Of these, two novel cDNA markers, A03/360 (*Icm1*) and B04/400 (*Psc1*), were isolated and characterised.

DDPCR analysis identified two types of X cells, an "early" X cell that was closely related to ES cells, and a "late" X cell type with distinctive gene expression. This analysis demonstrated that multiple pluripotent cell subpopulations exist within the ES cell to X cell transition.

In situ hybridisation analysis demonstrated that A03/360 (*Icm1*) and B04/400 (*Psc1*) exhibited distinct but overlapping expression profiles during early embryogenesis, subdividing the pool of *Oct-4*⁺ pluripotent cells. Inner cell mass 1 (*Icm1*, A03/360) was expressed in pluripotent cells during preimplantation development, in the morula and inner cell mass. *Icm1* expression was downregulated during cellular differentiation to trophectoderm and primitive endoderm and prior to the formation of primitive ectoderm. Expression of *Icm1* therefore identified a pluripotent cell sub-type present during preimplantation development, suggesting potential roles of *Icm1* during the differentiation of trophectoderm and primitive endoderm, or in the maintenance of pluripotency in the inner cell mass.

Peri-implantation stem cell 1 (*Psc1*, B04/400) was expressed in the late stage inner cell mass, in inner cell mass derivatives during peri-implantation development, and in the embryonic ectoderm prior to proamniotic cavitation. *Psc1* expression therefore identified a pluripotent cell subpopulation present during peri-implantation development, suggesting potential roles of *Psc1* in the differentiation of primitive endoderm, proliferation of stem cells, or proamniotic cavitation. *Psc1* was also expressed in the extraembryonic ectoplacental cone, which indicated potential roles during early placental development.

Icm1 and *Psc1* expression revealed the presence of overlapping subpopulations within the pluripotent cell pool, which highlighted the complexity of pluripotent cell development and regulation. The expression of *Psc1* refined the definition of the embryonic equivalents of ES cells and “early” X cells as pluripotent cells present from approximately 4.0/4.5 days post-coitum (d.p.c.) to 5.0 d.p.c. and the embryonic equivalents of “late” X cells as the primitive ectoderm from approximately 5.25 d.p.c. This verified the *in vitro* model of pluripotent cell development and demonstrated the potential of this system for the identification of pluripotent cell subpopulations from the inner cell mass to the primitive ectoderm stages of embryogenesis.

Psc1 was selected for additional analysis, and cDNA clones spanning a 3.5 kb *Psc1* sequence were isolated. The 1005 residue *Psc1* open reading frame contained three regions of similarity to the predicted *C. elegans* protein CLEB0336.3 10 and one region to the

human expressed sequence tag HFBDS04, suggesting potential novel protein domains. The presence of potential nuclear and subnuclear localisation sequences suggested that the *Psc1* protein could be localised nuclear “speckle” regions, subnuclear domains that contain pre-mRNA splicing machinery and splicing regulators.

RNase protection analysis demonstrated that *Psc1* was differentially regulated between individual tissues at 16.5 d.p.c. and in the adult. High level *Psc1* expression was detected in embryonic lung and brain, in adult lung, and in the placenta. This suggested that *Psc1* activity could be a component of a recurring developmental function, required at multiple sites during embryogenesis and in the adult, and confirmed a potential role for *Psc1* during placental development.

The approaches described in this thesis demonstrate the potential to identify and characterise molecular heterogeneity within the developing pluripotent cell pool *in vivo*, via the controlled progression and analysis of pluripotent cells *in vitro*.

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 to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED.

DATE ...1/12/96.....

ACKNOWLEDGMENTS

I would like to thank Professor Peter. D. Rathjen for allowing me to undertake a Ph.D. in the Department of Biochemistry.

I would like to thank the following people, who have all been a part of this thesis and of my life in the department. Firstly, to Julian Wells, for his enthusiasm and vision. To Pete, for fostering a hacker in the rough and saving me from a lonely existence. Thankyou for your supervision, advice and for critically reading the thesis. Thanks to Ian Lyons for supervision, careers counselling and tips on photography.

Thankyou to all the co-habitants of my roving Ph.D., in labs 20, 20A, 110, 111, 108, 104, and 102, there are too many of you to thank individually. To all the past, residual and commercialised members of the Wells group, thanks for putting up with an expatriate. I couldn't just leave. Thanks especially to Renate, Kym and Leslie.

Thanks to everyone in the Rathjen lab for a great couple of years. In particular: Gav for discussions and learning how not to spill label; Mikey for not letting truth get in the way of a great story; Linda for chuckling to herself at the mention of "baskets"; Bryan for harbinger and swishes from the back corner; Joy for trying to show me an embryo you can't see; Julie for her new leaf; and tangentially to that, Rog.

I would specifically like to thank Paul Thomas, Michael Bettess, Julie-Anne Lake and Gavin Chapman for the odd RNA sample; Joy Rathjen for dissection advice; O&G BresaGen-ites, Joe Wrin, Ian Lyons and Renate Faast for the spare blastos.

Thankyou Mum for all your encouragement. Thankyou Dad for the space, I understand your sacrifice. To Mum and Dad Myers, thanks for adopting me as one of your own. Thanks Paul, Sonia, Greg, Jenny, Ros and Craig for being there. Thanks Andrew for calling me "Herr Doctor Schulz" for the last five years.

To my wife Jo-anne. Thankyou for being my best friend, for understanding and supporting me though every stage without question. *"Now darling come with me, help me share my load. From the dark and lonely street, to the bright side of the road."* Van Morrison.

Finally, thankyou to God, who has set my path and blessed my work. *Knowledge puffs up, but love builds up. 1 Cor. 8:1*

During the course of this work I was supported by an Australian Postgraduate Research Award (Industry).

CHAPTER 1

General Introduction



CHAPTER 1

GENERAL INTRODUCTION

1.1 Development

1.1.1 The Study of Mammalian Development

The development of complex organisms from a fertilised egg has fascinated biologists for centuries. Understanding of development has largely come from studies of invertebrates such as *Drosophila melanogaster* and lower vertebrates such as *Xenopus laevis*, which are amenable to genetic analysis and structural manipulation respectively. This has enabled the investigation of developmental processes such as embryonic induction, organising centres, cell fate decisions, differentiation and cell migration.

The study of mammalian development is particularly interesting, as we ourselves are mammals and information gleaned from model organisms sheds light on human embryology. This information also aids the understanding of the molecular basis of genetic and developmental disease conditions, providing potential therapeutic approaches. Furthermore, increased understanding of mammalian development may enable the development of useful approaches to modify livestock or agricultural production.

The analysis of mammalian development has benefited extensively from the study of lower organisms. However, early mammalian development is dependent on zygotic genome activation, rather than maternal cytoplasmic contribution and development takes place within the uterine environment. This demonstrates that fundamental developmental differences exist in mammals that will not be resolved by the analysis of lower organisms alone.

1.1.2 The Mouse as a Model for Mammalian Development

The mouse is small, has large litters with a short gestation time and is easily maintained in breeding colonies, factors that make the mouse a suitable model for the investigation of mammalian development. The ability to inbreed *Mus musculus* has led to the generation of over 400 inbred strains for research (Festing 1992) and the identification of at least 1000

genes (Copeland *et al.* 1993). Furthermore, the morphological changes that occur throughout murine embryogenesis have been thoroughly detailed by histological and electron microscopic analysis (Kaufman 1992). Finally, the capacity to culture and manipulate murine cells and preimplantation embryos has enabled the isolation of pluripotent stem cell lines *in vitro* (Evans and Kaufman 1981, Martin 1981). These cells are capable of contributing to embryonic tissues when genetically altered and reintroduced into the preimplantation embryo, and provide the opportunity to study gene function in a developmental context by gene targeting (reviewed in Smith 1992). Pluripotent cell culture also enables the elucidation of early embryonic signals and developmental gene function through the analysis of cellular events such as proliferation and differentiation. These features make the mouse model a powerful and effective system to analyse and decipher the molecular controls of mammalian development.

Detailed reviews of murine development are described in Hogan *et al.* 1994, Kaufman 1992, Rossant and Pedersen 1988. A summary of cell types and developmental lineages formed during murine embryogenesis is shown in figure 1.1.

1.2 Preimplantation Development

1.2.1 Fertilisation and Cleavage Stage Development

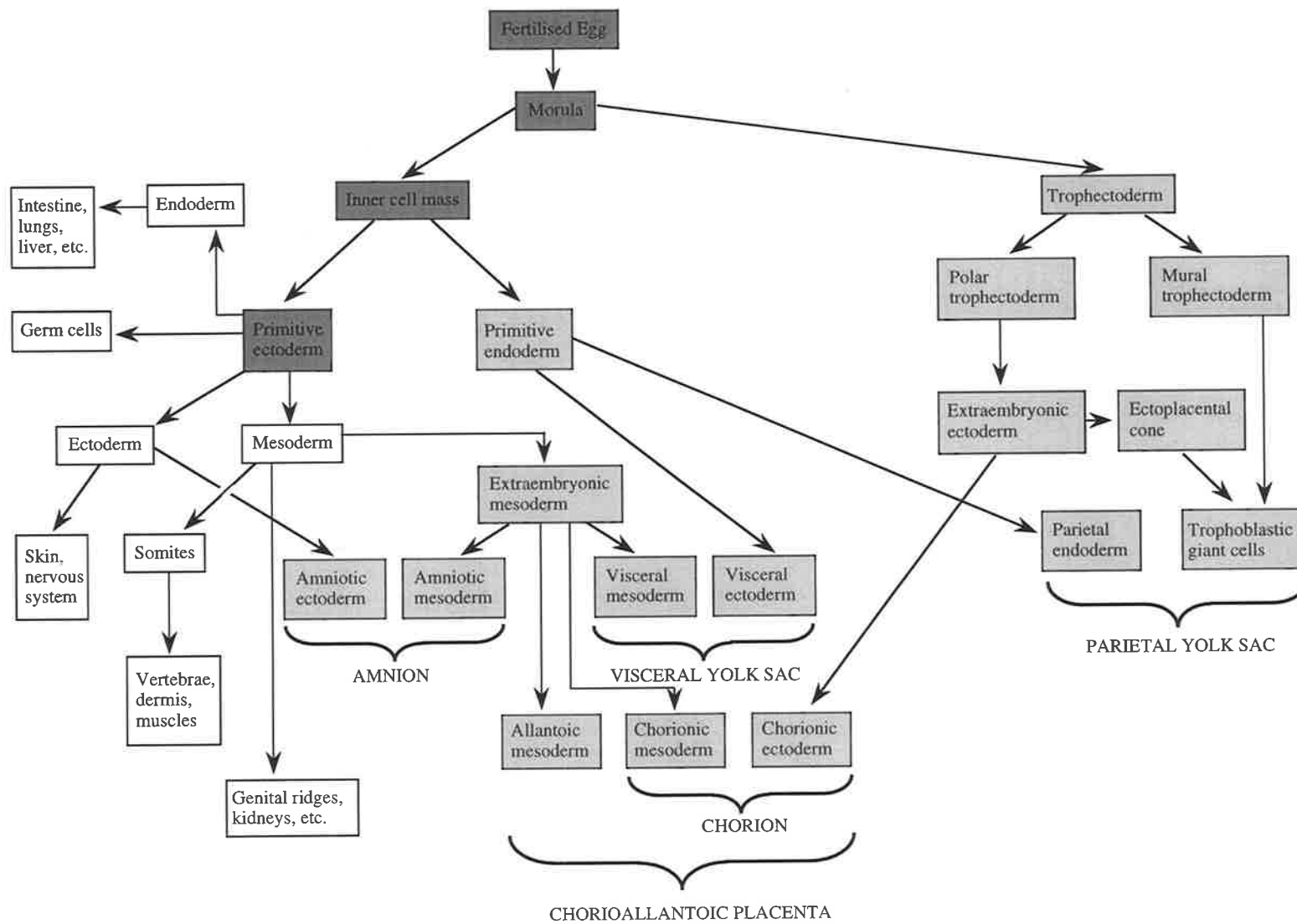
The genesis of the embryonic mouse begins with fertilisation, the fusion of the oocyte and sperm gametes. The fertilised oocyte subsequently initiates the second meiotic division and nuclear membranes enclose the maternal and paternal haploid chromosomes. Following DNA replication and migration of pronuclei towards the centre of the cell, the pronuclear membranes break down and chromosomes align on the spindle, resulting in the first cleavage division approximately 17-20 hours post-insemination.

Rapid degradation of maternally loaded oocyte mRNA occurs during the two-cell stage, being replaced by zygotic genome activation and transcription (Kidder 1992, Latham *et al.* 1991). Transcriptional activation signals the transfer of developmental control to the embryonic genome.

Figure 1.1 Summary of the Tissue Lineages of Murine Embryogenesis.

Darkly shaded boxes indicate tissues which will give rise to the embryo proper and extraembryonic tissues. The open boxes indicate tissues of the embryo proper, while lightly shaded boxes represent extraembryonic tissues.

Adapted from Hogan *et al.* (1994).



The embryonic blastomeres cleave again twice, approximately once every 20 hours, resulting in the formation of an embryo with eight loosely associated cells. Individual blastomeres are totipotent up to the early eight-cell stage, with normal development resulting from single blastomeres of two- and four-cell stage embryos (reviewed in Pedersen 1988).

1.2.2 Compaction and Polarisation

As cleavage progresses beyond the 8-cell stage there is a restriction in blastomere potential that results in the first embryonic differentiation event, the formation of trophoctoderm. Changes that result from the process of compaction, which begins at approximately 60 hours post-fertilisation, are associated with trophoctoderm differentiation. The blastomeres of the 8-cell stage embryo flatten, reducing intercellular space and maximising cell-cell contacts. Tight junctions form at outer surfaces (Fleming *et al.* 1992), creating blastomere polarity, and an increase in cell-cell communication is mediated by the establishment of gap junctions (Becker *et al.* 1992). Blastomere polarisation generates distinct apical and basolateral membrane, cytoskeletal and cytoplasmic domains.

Compaction therefore leads to a close association between blastomeres and ensures that at the next cleavage, to the 16-cell morula, distinct internal and external cells are generated for the first time. A cleavage plane horizontal to the axis of a polarised and compacted blastomere results in apical (outer) and basal (inner) cells. Exclusive features of the polarised phenotype are inherited by each daughter cell. Apical cells characteristically form the precursors of the trophoctodermal lineage whereas basal cells give rise to the inner cell mass (reviewed in Gardner 1983, Pedersen 1988).

1.2.3 Blastulation and Implantation

Cells of the 16- and 32-cell morula that inherit the apical domain phenotype continue to differentiate, forming a developmentally restricted functional epithelium, the trophoctoderm. By the late 32-cell stage the trophoctoderm features a basement membrane, mature cell-cell contacts and junctions, cytoplasmic reorganisation and membrane polarity (reviewed in Fleming *et al.* 1992). The trophoctoderm forms an impermeability barrier and is a precursor of extraembryonic tissues, such as chorion and placenta. As trophoctoderm differentiates

there is a concurrent inter-cellular accumulation of fluid, creating the blastocoel. Blastocoel cavitation is due to a passive flow of water, countering sodium influx generated by a Na^+ , K^+ -ATPase present on the trophoctoderm basal membrane (reviewed in Biggers *et al.* 1988).

Some internal morula cells become precursors of the inner cell mass (ICM), a small clump of 20-40 nonpolarised cells abutting the polar trophoctoderm at one end of the blastocoel. The ICM contains the pluripotent cells that give rise to the embryo proper and becomes more prominent with the gradual expansion of the blastocoel.

At 4.0-4.5 days post-coitum (d.p.c.) the second differentiation event of murine embryogenesis occurs. An epithelial layer of primitive endoderm cells differentiates from the ICM cells that line the blastocoel. Primitive endoderm is restricted in developmental potential, generating the extraembryonic parietal and visceral endoderm of the yolk sac (reviewed in Gardner 1983).

Just prior to implantation, a mural trophoctoderm protease and uterine enzymes digest the zona pellucida, breaching the glycoprotein matrix. The blastocyst undergoes a series of rhythmic contractions and hatches from the zona (Kaufman 1992). Hatched embryos enlarge to a characteristic ovoid morphology, termed the fully expanded blastocyst. The ICM of the fully expanded blastocyst has a flattened morphology and is termed the epiblast. The epiblast has a reduced capacity to differentiate to primitive endoderm (Beddington 1983).

Murine preimplantation development, from the four cell stage to the fully expanded blastocyst, is summarised in figure 1.2.

1.3 Peri-implantation Development

The surface of the uterine epithelium becomes receptive to implantation and at approximately 4.5 d.p.c. the blastocyst attaches to the anti-mesometrial uterine wall, via the mural trophoctoderm. Blastocyst attachment is consolidated by the decidual reaction, the stimulation of uterine stroma to deposit a spongy tissue mass around the embryo.

Implantation coincides with key developmental changes to cell types of the blastocyst (Hogan *et al.* 1994). The epiblast, which consists of approximately 20 cells at implantation, undergoes rapid proliferation and generates approximately 8000 epiblast cells by 7.5 d.p.c. Mural trophoctoderm ceases cell division but continues DNA replication, generating primary

Figure 1.2 Murine Preimplantation Development.

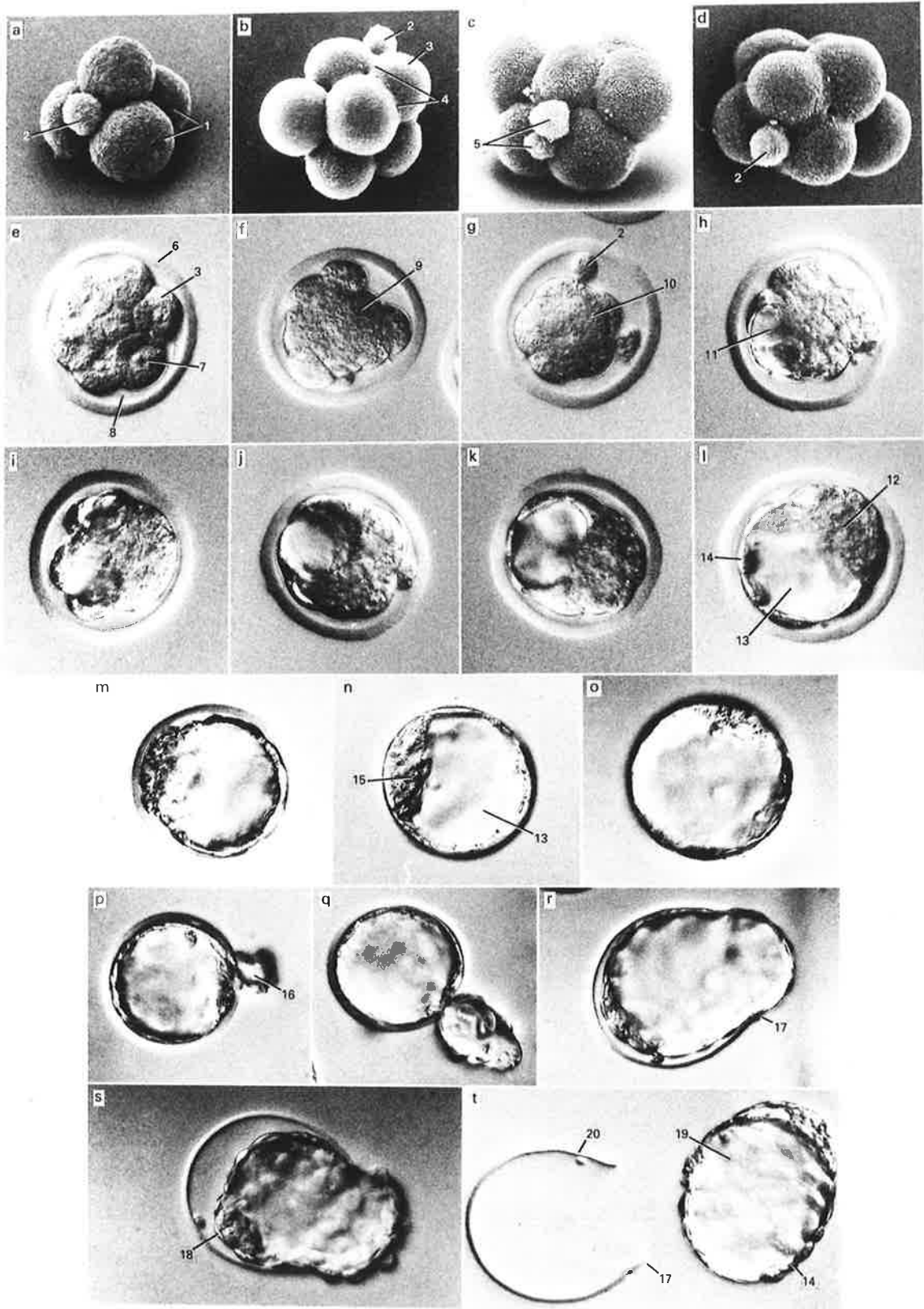
(a-d) 4- to 8-cell stage embryos of pre- and early compaction stages. (a) 4-cell stage, (b) 8-cell stage (very early compaction), (c-d) 8-cell embryos (early compaction).

(e-l) Progression of 8-cell embryo to blastocyst stage. (e) 8-cell embryo, (f) early morula, (g) fully compacted morula, (h-k) early blastulation, (l) mid stage blastocyst.

(m-t) Blastocyst "hatching". (m-o) gradual increase in blastocyst volume and thinning of the zona pellucida, (p-s) breach in zona pellucida and escape of trophectoderm, (t) fully expanded blastocyst with spherical/ovoid form.

1. blastomeres of 4-cell stage embryo
2. second polar body
3. blastomere of 8-cell stage embryo
4. early compaction (individual blastomere boundaries no longer distinct)
5. division products of the second polar body
6. zona pellucida
7. nucleus of 8-cell stage embryo
8. perivitelline space
9. early morula stage embryo (indistinct blastomere boundaries)
10. fully compacted morula stage embryo
11. accumulation of blastocoelic fluid (early cavity formation)
12. inner cell mass
13. blastocoel
14. mural trophectoderm cell
15. inner cell mass
16. trophectoderm escaping through breach in zona pellucida
17. boundary of breach in zona pellucida
18. polar trophectoderm cell
19. fully expanded blastocyst
20. empty zona pellucida

Reproduced from Kaufman (1992).



trophoblastic giant cells with polytene chromosomes. Polar trophoblasts proliferate rapidly, migrating around the embryo and penetrating into the endometrium, eventually forming the majority of the placenta. A projection of polar trophoblasts forces the epiblast to protrude into the blastocoel. As the epiblast proliferates, primitive endoderm differentiates into two distinct cell types. Firstly, endoderm that remains in contact with the epiblast and extraembryonic ectoderm becomes organised into an epithelial layer of visceral endoderm. Secondly, primitive endoderm cells that migrate onto the blastocoel surface of the trophoblasts become parietal endoderm. By 5.0 d.p.c. the epiblast consists of a solid mass of apolar pluripotent cells, the “epiblast bud”, surrounded by a basement membrane and visceral endoderm.

Peri-implantation and early postimplantation stages of murine development are summarised in figure 1.3.

1.4 Postimplantation Development

1.4.1 Proamniotic Cavity and Primitive Ectoderm Formation

The process of proamniotic cavitation converts the solid bud of the 5.0 d.p.c. epiblast to the primitive ectoderm, a pseudostratified columnar epithelium (Fig. 1.3). By 5.5 d.p.c. a small central cavity caused by the programmed cell death of central epiblast cells has occurred. At 6.0 d.p.c. cavitation has extended throughout the epiblast and extraembryonic ectoderm and the pluripotent cells are organised into a cup-shaped epithelial monolayer of primitive ectoderm.

A two step mechanism for proamniotic cavitation has been proposed by Coucouvanis and Martin (1995). A temporal “death signal” expressed by visceral endoderm induces apoptosis in epiblast cells. Survival of an outer monolayer of ectoderm is mediated by contact with the basement membrane and an associated “survival signal”.

There are several distinguishing features between the pluripotent cells of the ICM and primitive ectoderm. Firstly, primitive ectoderm cells are polarised, with a subapical concentration of cytokeratins (Jackson *et al.* 1981) and a continuous basal lamina (Leivo *et al.* 1980) and secondly, the primitive ectoderm exhibits a higher rate of cell division than the

Figure 1.3 Peri-implantation and Early Postimplantation Murine Development.

(A) Schematic representation of peri-implantation development, the upper row depicts sagittal sections and the lower row depicts transverse sections through the planes indicated by dotted lines. Pluripotent cells are shown in blue. (E 4.0) Blastocyst stage embryo prior to implantation, containing an outer trophoctodermal layer surrounding the fluid-filled blastocoel and an inner cell mass (ICM). Primitive endoderm has formed by delamination and differentiation of ICM cells facing the blastocoel. Some primitive endoderm cells migrate away from the ICM and form the parietal endoderm. Primitive endoderm cells that remain in contact with the ICM form the visceral endoderm. (E 5.0) The ICM (epiblast) proliferates to form a solid embryonic ectodermal “bud”, surrounded by visceral endoderm. (E 6.0) The proamniotic cavity has formed and is lined by primitive endoderm.

(B) Sagittal section of a 4.5 d.p.c. implanting blastocyst

10. dilated blood vessel
11. endometrial layer
12. inner cell mass
13. visceral endoderm
14. mural trophoctoderm cells
15. blastocoel

(C) Sagittal section of a 5.0-5.5 d.p.c. embryo (“epiblast bud” stage)

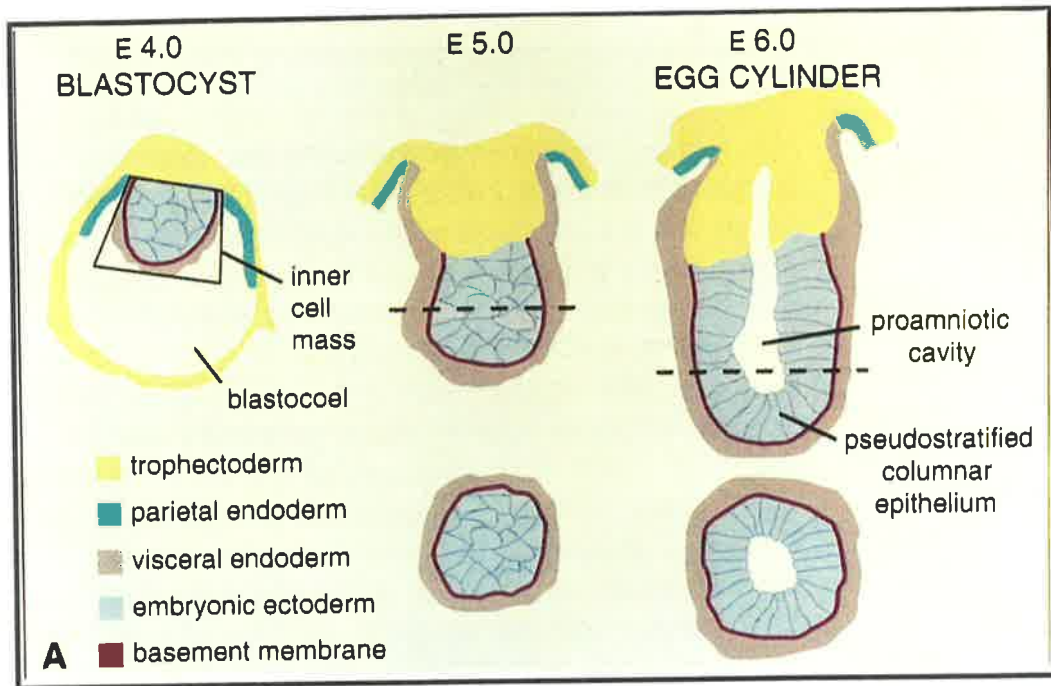
- ec. embryonic ectoderm
- ex. extraembryonic ectoderm
- pe. parietal endoderm
- ve. visceral endoderm

(D) Sagittal section of a 5.5 d.p.c. embryo

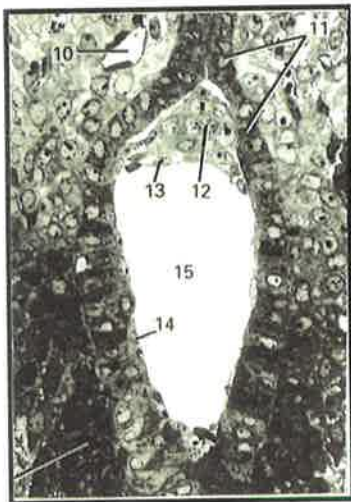
11. ectoplacental cone
12. extraembryonic ectoderm
13. visceral layer of extraembryonic endoderm
14. approximate line of demarcation between embryonic and extraembryonic endoderm
15. proamniotic cavity
16. visceral layer of embryonic endoderm
17. embryonic ectoderm cells (primitive ectoderm)
19. parietal layer of extraembryonic endoderm
20. Reichert’s membrane

(A) Reproduced from Coucouvanis and Martin (1995), (B) and (D) adapted from Kaufman (1992), (C) reproduced from Hébert *et al.* (1991).

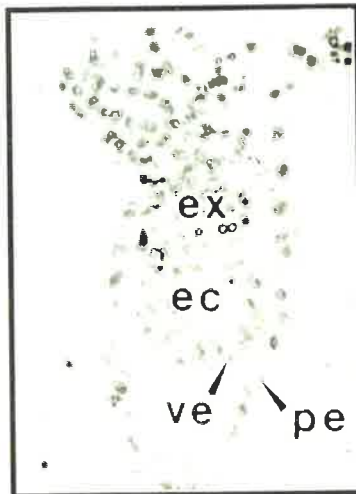
A



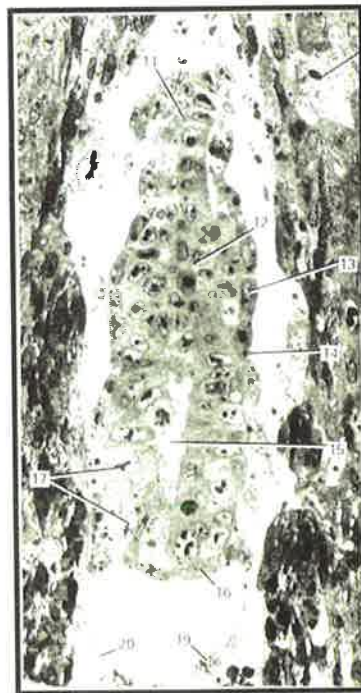
B



C



D



ICM (Snow 1977). Thirdly, the ICM and primitive ectoderm exhibit distinct gene expression (1.5.2).

1.4.2 Gastrulation

Gastrulation is the process that transforms the cellular monolayer of primitive ectoderm to a multilayered embryo, establishing the basic foetal body plan. All the tissues of the embryo proper are derived from the primitive ectoderm (Lawson *et al.* 1991, Tam and Beddington 1992) through a series of complex events involving cellular migration, proliferation, differentiation, pattern formation, apoptosis and organogenesis. The gastrulation stages of murine postimplantation development are schematically represented in figure 1.4.

At 6.5 d.p.c. the primitive streak forms at the posterior region of the embryo. Primitive ectoderm within this region invaginates through the primitive streak, losing epithelial phenotype and emerging as mesoderm, between the primitive ectoderm and visceral endoderm layers. As gastrulation continues, the primitive streak elongates to the embryonic distal tip and mesoderm migrates axially and laterally, reaching the anterior pole by 8.0 d.p.c. Mesodermal derivatives exhibit regional morphological diversity and condense into approximately 65 somite pairs and somitomeres along the anterior-posterior axis. Mesodermal cell types form the precursors of vertebrae, ribs, dermis, kidney and muscle tissues in the developing foetus.

Embryonic endoderm is also induced during gastrulation, emerging from anterior regions of the streak and migrating anteriorly, displacing visceral endoderm. Embryonic endoderm is the developmental primordium of gut, lung and liver tissues. Ectodermal cell types are generated directly from embryonic ectodermal precursors, for example foetal skin and nervous system.

1.4.3 Partitioning of the Primordial Germ Cell Lineage

The process of gastrulation depletes the embryo of all pluripotent cell types, except a population of primordial germ cells (PGCs) which are partitioned from the posterior primitive ectoderm at 7.0 d.p.c. First detected in the extraembryonic mesoderm of the amniotic fold

Figure 1.4 The Gastrulation Stage of Murine Embryogenesis.

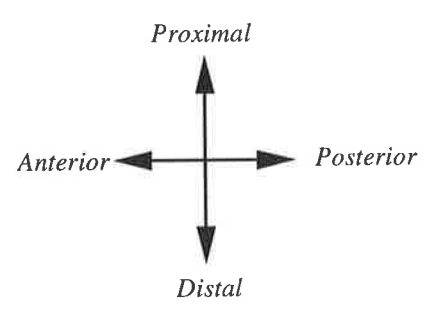
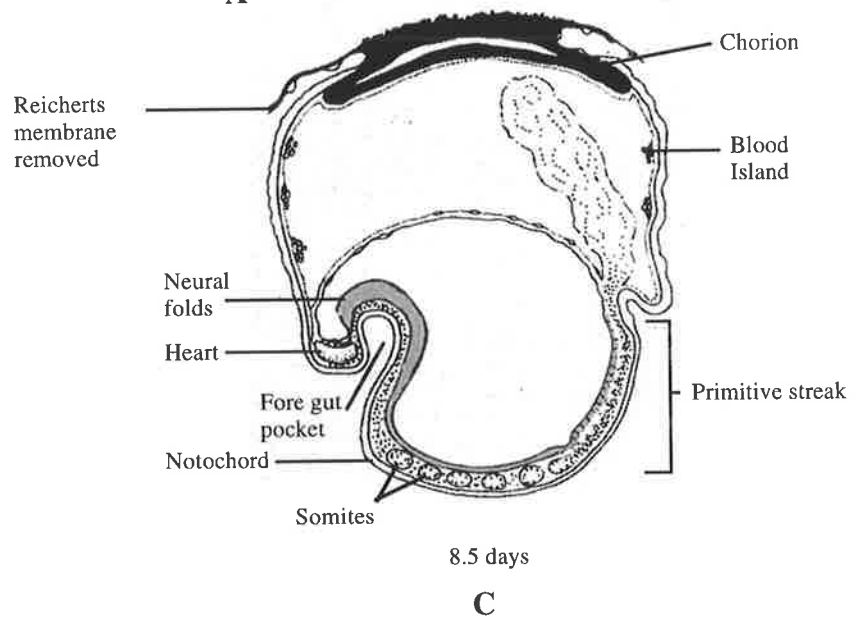
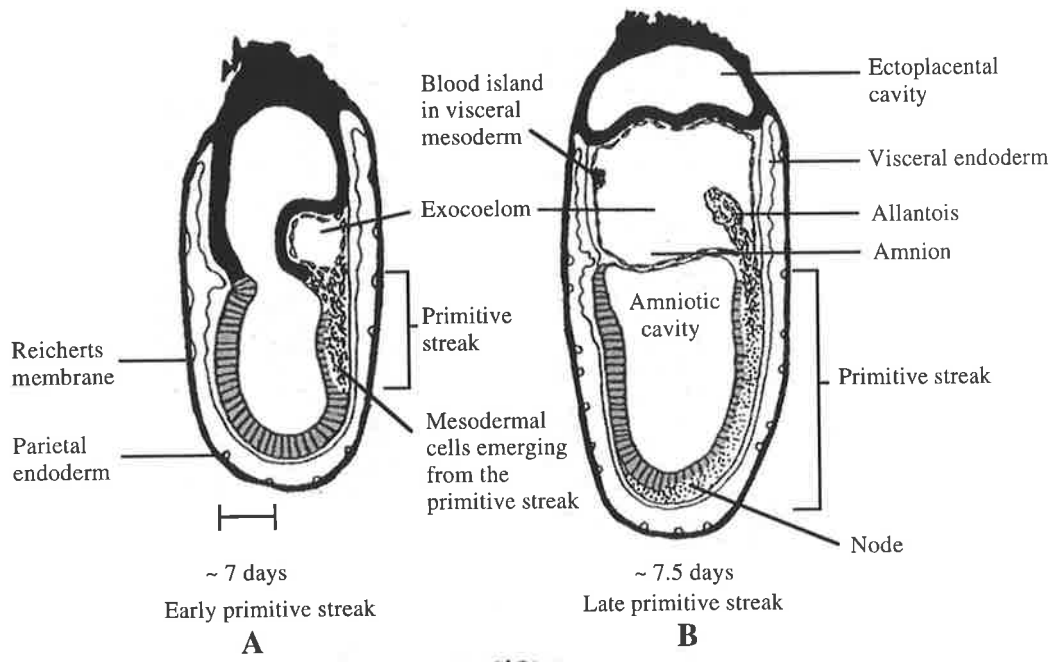
Schematic representation of murine development from 7.0 d.p.c. to 8.5 d.p.c.

(A) At 6.5-7.0 d.p.c. the onset of gastrulation is marked by the formation of the primitive streak, at the posterior end of the primitive ectoderm.

(B) The newly formed mesoderm moves proximally, anteriorly and laterally to generate a distinct layer of embryonic mesoderm.

(C) By 8.5 d.p.c. a variety of differentiated tissues have formed along the embryonic anterior-posterior axis, including somites, notochord, neural tissue, gut and heart precursors.

Adapted from Hogan *et al.* (1986).



(Ginsberg *et al.* 1990), a cluster of approximately 75 PGCs becomes incorporated into the base of the allantois by 8.0 d.p.c. Subsequently, PGCs leave the allantois and migrate through the gut endoderm, reaching the genital ridges by 11.0-11.5 d.p.c. Cell division occurs every 16 hours during migration, resulting in approximately 25,000 PGCs colonising each primordial gonad by 13 d.p.c. (Tam and Snow 1981). The proliferative activity of PGCs is governed by the somatic microenvironment once the genital ridge has been colonised. Female PGCs arrest in prophase of the first meiotic division once entering the presumptive ovary, whereas male germ cells continue to divide in the developing testes until mitotic arrest at 14 d.p.c.

1.5 Pluripotent Cell Populations in Early Mouse Development

One of the characteristics of early mammalian development is the coordinated maintenance, proliferation and differentiation of a population of undifferentiated pluripotent stem cells, which forms the precursors of all embryonic tissues (Fig. 1.5).

The pluripotent cell pool is initially generated by cleavage divisions of a single precursor cell, the fertilised egg. In the mouse, the first differentiation event occurs at the 16- to 32- cell stage, where the undifferentiated pluripotent cells of the ICM, the founder cells of the embryo proper, are partitioned from the trophectoderm. ICM cells that line the blastocoel subsequently differentiate to primitive endoderm. The ICM of the fully expanded blastocyst contracts and is termed the epiblast, and has a restricted capacity to differentiate to primitive endoderm. Morphological and proliferative changes to the pluripotent epiblast accompany implantation. The epiblast proliferates into a solid bud in the former blastocoel, and the programmed cell death of central epiblast cells contributes to the formation of the proamniotic cavity. These processes reorganise the epiblast into a cup shaped epithelial monolayer, the pluripotent primitive ectoderm. The primitive ectoderm is the substrate for the inductive signals of gastrulation, which begins with the formation of the primitive streak. The ingress of primitive ectoderm cells through the primitive streak and the associated morphogenic, migratory and developmental changes, generates the three primary germ layers of the embryo. The pluripotent primordial germ cells are set aside during gastrulation and they migrate to the presumptive gonad. Therefore, the coordinated regulation of pluripotent cell

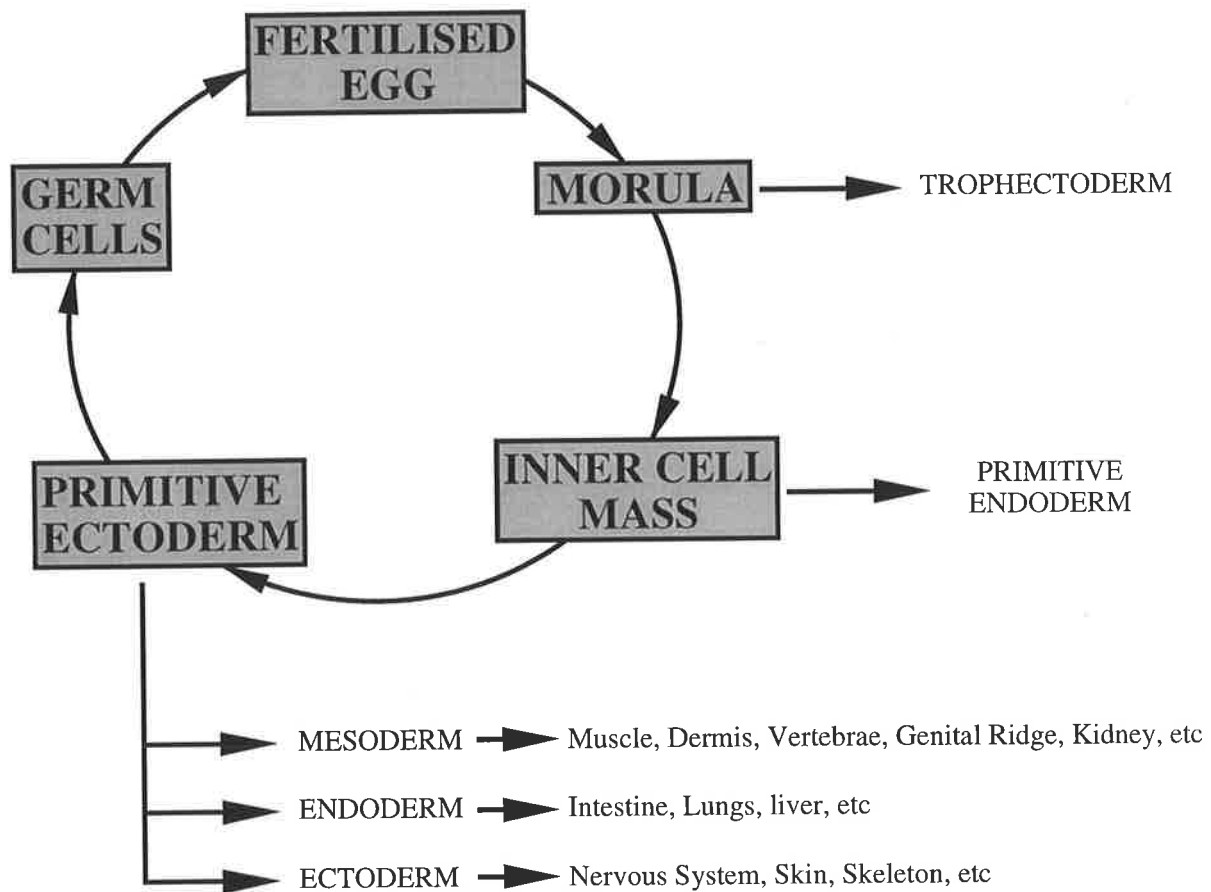


Figure 1.5 Totipotent and Pluripotent Cells in Murine Development.

Murine development begins with a single totipotent cell, the fertilised egg. Subsequent cleavage generates the equivalent totipotent blastomeres of the morula. The pluripotent cell lineage is established by the formation of the inner cell mass and is expanded and maintained during primitive ectoderm formation. The demarcation of the germ cells maintains the pluripotent cell lineage for the next generation. Pluripotent and totipotent cell types express the markers *Oct-4* and alkaline phosphatase and are recognised by the monoclonal antibody SSEA-1. Some of the differentiated tissues arising from each pluripotent cell stage are indicated.

maintenance, proliferation and differentiation is critical for the establishment of extraembryonic tissues, the differentiated lineages of the embryo and the basic body plan. There are therefore likely to be tight controls on the decisions and signals regulating the proliferation and differentiation of the pluripotent stem cell pool (reviewed in Smith 1992).

Experimental manipulation of pre-organogenic murine embryos has demonstrated remarkable developmental plasticity within the pluripotent cell pool. For example, a normal foetus can still be formed even if up to 80% of the primitive ectoderm is ablated in the gastrulating embryo (Snow and Tam 1979). Additionally, the aggregation of multiple pre-implantation embryos results in a single viable mouse of normal size (Petters and Markert 1980). It would seem therefore, that the pluripotent cell pool of the murine embryo is regulated in a flexible manner, capable of adapting to insult by reprogramming developmental fate. The proposal that different pluripotent cell types can interconvert through a common intermediate (Rossant 1993) could highlight the mechanism of pluripotent cell reprogramming.

Deciphering the signals and decisions that coordinate the maintenance, proliferation, differentiation, or reprogramming of the pluripotent cell pool will therefore be essential to reach an understanding of murine embryogenesis.

1.5.1 The Heterogeneous Nature of Murine Pluripotent Cell Populations

The classification of murine pluripotent stem cell types, on the basis of developmental behaviour and morphology, does not appear to delineate the full repertoire of pluripotent cell subtypes. Accumulating experimental evidence indicates the presence of both temporal and spatial heterogeneity within stem cell populations, and suggests that multiple distinct pluripotent cell “states” exist during embryogenesis. Pluripotent cell “states” or subpopulations could reflect the presence of distinct pluripotent cell types, such as the ICM and primitive ectoderm. Conversely, different pluripotent cell “states” could be generated within a homogeneous pluripotent cell population, if heterogeneity results from different cells within the population responding to distinct developmental signals.

Temporal heterogeneity within the stem cell pool is observed during blastocyst formation and development. Firstly, the capacity to differentiate to trophectoderm is not lost

with the allocation of the ICM lineage, but is gradually lost as the ICM develops. (Gardner 1983, Nichols and Gardner 1984, Chisholm *et al.* 1985). This suggests that multiple developmentally distinct ICM cells could exist during blastulation. Secondly, pluripotent embryonic stem (ES) cells can be isolated from the ICM/epiblast cells of the blastocyst just prior to implantation, but have only been isolated from earlier embryos after a period of development in culture (reviewed in Smith 1992). These studies suggest a developmental “maturation” of ICM cells from 3.0-4.5 d.p.c. and indicates the potential for multiple ICM cell types during this period.

Other examples of temporal heterogeneity within stem cells are found during peri-implantation development. Multiple cellular events are associated with blastocyst implantation, leading to massive stem cell proliferation, accompanied by selective apoptosis or survival of pluripotent cell types to form the primitive ectoderm (Coucouvanis and Martin 1995). It is likely that firstly, these pluripotent cells are distinct from preimplantation epiblast cells and secondly, given the complex development of pluripotent cells during this period, several distinct stem cell “states” could exist between 4.5 d.p.c. and 5.5 d.p.c. The implied molecular differences within these developing stem cell types has not been resolved, and the current term “epiblast” encompasses all pluripotent cells during this transition.

The most striking demonstration of heterogeneity within pluripotent cell populations is in the primitive ectoderm. Extensive fate mapping and transplantation experiments have demonstrated regionalised capacity to form certain post-gastrulation tissues within the primitive ectoderm (Tam and Beddington 1987, Tam 1989, Lawson *et al.* 1991, Tam and Beddington 1992, Quinlan *et al.* 1995). These differences are first detected prior to primitive streak formation and become more defined as gastrulation proceeds (Fig. 1.6, reviewed in Hogan *et al.* 1994). Proximal epiblast regions contain precursors of primordial germ cells, extraembryonic ectoderm and amniotic ectoderm, while lateral and posterior primitive ectoderm is populated by cells of predominantly mesodermal fate. Up to 45% of the epiblast consists of ectodermal precursors residing in the anterior and distal regions (Snow 1977). Considerable developmental overlap is found between neighbouring fate map regions. However, individual epiblast cells are not developmentally committed (Lawson *et al.* 1991), with heterotopically grafted cells remaining competent to form derivatives in all three germ

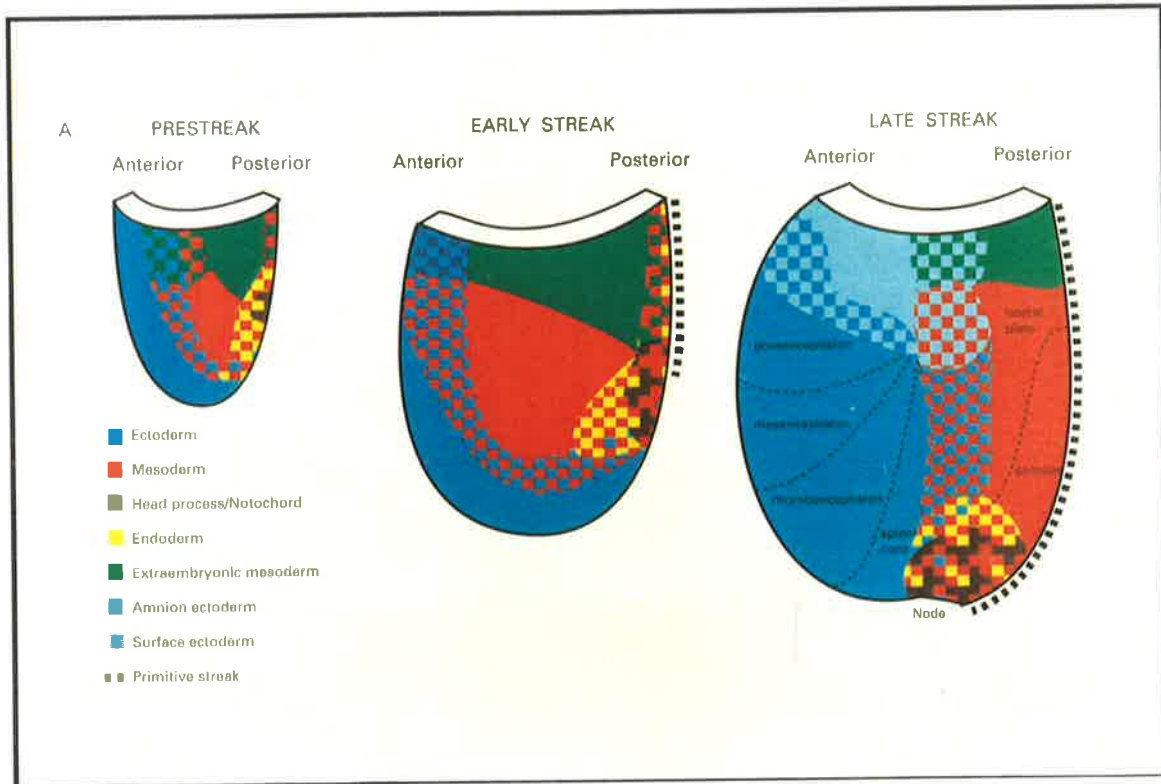


Figure 1.6 Fate Maps of Primitive Ectoderm and Pluripotent Cell Heterogeneity

Fate maps of primitive ectoderm before and during gastrulation. The left half of the embryonic epiblast is shown for prestreak, early streak and late streak development. The origin, boundaries and overlap of prospective tissues are plotted as indicated. Prestreak: 6.0 d.p.c., early streak: 7.0 d.p.c., late streak: 7.5 d.p.c.

Reproduced from Hogan *et al.* (1994).

layers by adapting to the fate of their new location (Beddington 1982). Therefore, while fate mapping indicates that a mosaic of committed epiblast cell types does not exist, it demonstrates the presence of regionalised heterogeneity within the primitive ectoderm and confirms the developmental lability of the pluripotent stem cell pool. The pluripotent cell “states” detected in the primitive ectoderm could represent a homogeneous cell population where different subpopulations respond to distinct developmental signals and therefore exhibit regionalised gene expression.

The molecular basis of primitive ectoderm heterogeneity remains unclear. Recent analysis however has demonstrated that expression of the homeobox gene *Hesx1* (*Rpx*) is restricted to anterior primitive endoderm and primitive ectoderm regions during gastrulation (P. Thomas unpublished data, Hermes *et al.* 1996), possibly identifying and specifying anterior neuroectodermal precursors within the epiblast. This is the first definitive evidence of regionalised gene expression within the early primitive ectoderm and supports the prediction that multiple pluripotent cell subtypes exist and are tightly regulated within this tissue. Whether it may be possible to distinguish additional primitive ectodermal subpopulations on the basis of gene expression correlated to developmental fate has yet to be determined.

These experiments demonstrate the existence of numerous, subtly different, pluripotent cell “states” during early embryogenesis. The identification of pluripotent cell subtypes and analysis of their developmental behaviour will be fundamental to the understanding of the processes that establish and regulate the precursors for gastrulation in the mammalian embryo. It is expected that individual pluripotent cell subtypes will express a repertoire of functionally important, subpopulation-specific genes. Identification of these genes is likely to enable the definition of pluripotent cell subpopulations and investigation of the molecular basis and function of heterogeneity within the developing pluripotent cell pool.

1.5.2 Gene Expression in Murine Pluripotent Cells

The advent of molecular biological techniques has enabled investigation and manipulation of the molecular basis of developmental programs. However, analysis of the pluripotent cells of early murine embryogenesis has been hampered by the small size and inaccessible nature of some stages of development *in utero*. Additionally, genetic approaches

have not identified many factors critical to pluripotent cell development, with embryonic lethality the most likely consequence of pluripotent cell defects, leading to abortion or conceptus resorption. Finally, while the analysis of gastrulation and later embryonic development has benefited greatly from the study of lower organisms, the lack of a pluripotent cell pool in these organisms means that extrapolation to the pluripotent cells of mammalian development is not informative. Despite these difficulties a number of genes and markers have been identified that are restricted to pluripotent stem cell populations or are transiently expressed within the developing stem cell pool.

The POU domain homeobox gene *Oct-4* is the best characterised pluripotent cell specific transcript and is expressed by the fertilised egg, morula, ICM, primitive ectoderm, PGC and germ lines in the testis and ovary (Rosner *et al.* 1990, Schöler *et al.* 1990a, Schöler *et al.* 1990b, Yeom *et al.* 1991). Additionally, *Oct-4* expression is downregulated during the differentiation of these cells and thus provides a specific marker for the pluripotent state. Other markers for pluripotent cell types include the surface carbohydrate recognised by monoclonal antibody SSEA-1 (Solter and Knowles 1978) and endogenous alkaline phosphatase expression (Johnson *et al.* 1977).

Activin peptides are expressed in the early ICM at 3.5 d.p.c., but their expression is reduced during “maturation” to the late stage ICM at 4.5 d.p.c. (Albano *et al.* 1993). The zinc finger transcription factor *Rex-1* is expressed in the ICM. *Rex-1* expression is downregulated as the primitive ectoderm forms and is not detected at 6.0 d.p.c. (Rogers *et al.* 1991). Fibroblast growth factor 5 (*FGF-5*) is not expressed in the ICM or epiblast bud, but expression is upregulated in the primitive ectoderm at 5.25-5.5 d.p.c. (Haub and Goldfarb 1991, Hébert *et al.* 1991). *FGF-4* is not expressed in the morula, but is induced in the ICM and expressed by the epiblast bud. *FGF-4* expression appears to be regionalised to the posterior primitive ectoderm at 6.25 d.p.c. (Niswander and Martin 1992). *Hesx1* is expressed in anterior regions of the primitive ectoderm (1.5.1). Migrating and proliferating primordial germ cells are distinguished by their expression of tissue non-specific alkaline phosphatase (MacGregor *et al.* 1995) and the *c-kit* receptor (reviewed in Besmer *et al.* 1993). As germ cells enter the presumptive gonad and undergo meiotic or mitotic arrest, expression of *c-kit* is downregulated. In contrast to this, spermatocytes are identified by the subsequent

expression of the *Rex-1* (Rogers *et al.* 1991) and *cyclin A1* genes (Sweeney *et al.* 1996). Unfortunately, the temporal expression patterns of most of these markers have not been characterised in sufficient detail to map associated developmental transitions within the pluripotent cell pool. This is generally due to the small size and inaccessibility of early murine embryogenesis, which contributes to technical difficulties when mapping gene expression during these stages.

Gene expression analysis has therefore provided a glimpse into the complex nature of stem cell regulation during early development. This complements the prediction of heterogeneity within the pluripotent cell pool, indicating that distinct temporal and spatial pluripotent cell subpopulations may be identified on the basis of specific gene expression.

1.6 Embryonic Stem (ES) Cells as an *in vitro* Model of Early Development

1.6.1 Isolation of ES cells

The first reported isolation of embryonic stem (ES) cell lines from the ICM of the mouse blastocyst was in 1981 (Evans and Kaufman 1981, Martin 1981). Undifferentiated ES cells can be cultured as colonies on a layer of mitotically inactivated fibroblasts, or in the presence of leukemia inhibitory factor (LIF) or other members of the IL-6 family of cytokines. ES cells express markers of pluripotent cell types including *Oct-4*, SSEA-1 antigen and alkaline phosphatase. Importantly, the developmental pluripotency of ES cells is maintained in the undifferentiated state *in vitro*. ES cells are able to contribute to all tissues of the embryo when injected into the blastocyst, including functional derivatives of the germ line (Bradley *et al.* 1984, Robertson 1987). These criteria demonstrate that cultured ES cells retain full biological potential when reintroduced into the developing embryo. ES cells can be induced to differentiate through alternate pathways *in vitro*, through exposure to different chemical inducers or by cytokine withdrawal (Smith 1991). Cultured ES cells are therefore likely to be capable of mimicking the proliferation and differentiation decisions of embryonic pluripotent cells, thus providing a model system for early development *in vitro*.

The demonstration that ES cell lines can contribute to the germ line has enabled precise manipulation of the murine genome and investigation of gene function *in vivo* (reviewed in

Joyner 1991). Loci of interest can be altered by homologous recombination in undifferentiated ES cell cultures and the modified ES cells reintroduced into the blastocyst. The chimeric mice that result can be bred to homozygosity for the altered allele if the modified ES cells transmit their genotype through the germ line. In this way, targeted alterations of functionally important genes have enabled the investigation of many facets of mammalian development *in vivo*, such as immunology, pattern formation and organogenesis (reviewed in Brandon *et al.* 1995a, Brandon *et al.* 1995b, Brandon *et al.* 1995c).

1.6.2 The Relationship of ES Cells to Embryonic Pluripotent Cells: a Model of the Inner Cell Mass

The embryonic cell type that is represented by cultured ES cells has not been precisely determined (reviewed in Smith 1992). ES cells are isolated from the early epiblast, or late stage blastocyst ICM and retain epiblast-like properties. Like epiblast cells, ES cells are limited in their potential to differentiate to trophoblast and parietal endoderm in culture, cell types that are readily formed by isolated ICM cells. Conversely, after reintroduction to the blastocyst, ES cells have the capacity to colonise trophoblast and extraembryonic endoderm, properties akin to ICM cells (Beddington and Robertson 1989). The gene expression repertoire of ES cells is consistent with that of the ICM, indicating that ES cells form a model of the ICM *in vitro*. ES cells express *Rex-1* (Rogers *et al.* 1991), a marker of the ICM (1.5.2), but not *FGF-5* (Bettess 1993), a marker of the primitive ectoderm (1.5.2).

That ES cells give apparently conflicting information regarding their precise relationship to the ICM and epiblast, may indicate that a direct embryonic equivalent does not exist. Rather, ES cells may be an adapted pluripotent cell type, diverted from normal development by culture conditions and undergoing a proliferation based program. Alternatively, ES cells may represent epiblast cells that have undergone *in vitro* “deprogramming” or loss of developmental commitment, allowing responses to signals provided by host blastocysts upon injection. Interestingly, primordial germ cells can generate ES-like cells in the presence of LIF, steel factor and basic fibroblast growth factor (Matsui *et al.* 1992). This suggests that deprogramming of stem cells may be a generalised feature of a flexible pluripotent state, with

different pluripotent cell types able to interconvert through a common intermediate (Rossant 1993).

1.6.3 Maintenance of ES Cells: the Undifferentiated State

Originally ES cell cultures were maintained by culture on a layer of mitotically inactivated mouse fibroblast cells (Martin and Evans 1975b), with extensive differentiation occurring in the absence of feeders. This implied that cellular differentiation was the default state of an ES cell, with active promotion required for maintenance of the undifferentiated state. This property allowed the discovery of soluble factors that inhibit ES cell differentiation, including members of the IL-6 family of cytokines such as Leukemia inhibitory factor (LIF or DIA), Ciliary neurotrophic factor, Oncostatin M, Cardiotrophin-1 and the Interleukin-6/soluble Interleukin-6 receptor complex (Smith and Hooper 1987, Piquet-Pellorce *et al.* 1994, Pennica *et al.* 1995). The cell surface receptor subunit gp130 appears to be the common signalling component of multisubunit receptors through which these factors act (reviewed in Kishimoto *et al.* 1994). Therefore the identification of these factors, in particular the purification of LIF, has allowed the monoculture of undifferentiated ES cells over extended periods.

1.6.4 Differentiation of ES Cells and *in vitro* Models of Development

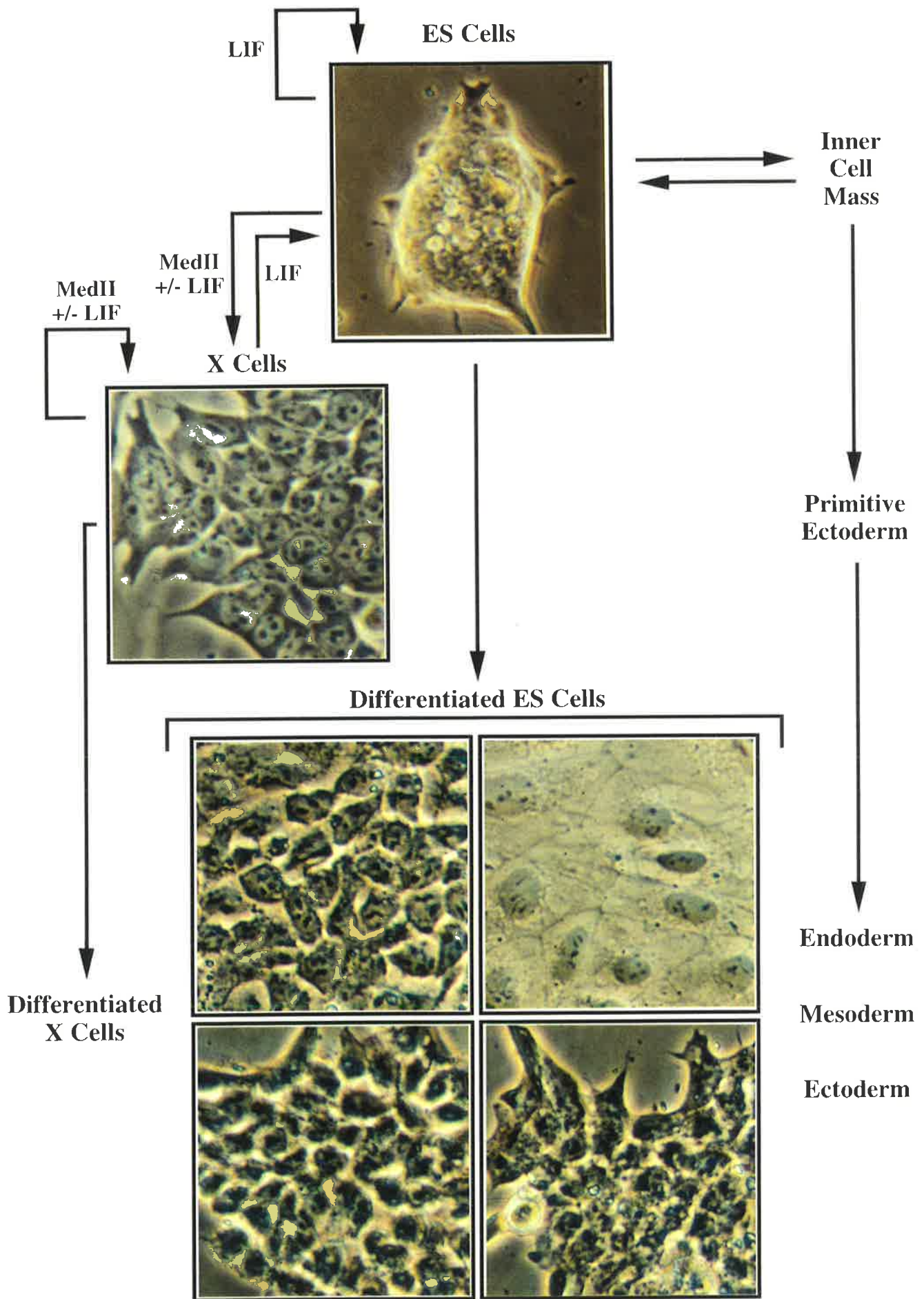
The ability to maintain ES cells with LIF in the absence of feeders and to distinguish undifferentiated cells from differentiated cells on the basis of morphology and gene expression, has allowed the investigation of ES cell differentiation potential and the establishment of *in vitro* models of early murine development. ES cell differentiation regimes include spontaneous differentiation by cytokine withdrawal, chemical induction, or more complex procedures involving ES cell aggregation to embryoid bodies, or conditioned medium-induced differentiation. The ES cell model of early development is comprised of these cell types as summarised in figure 1.7.

Figure 1.7 The Embryonic Stem Cell Model of Early Development.

Schematic representation of the ES cell derived cell types, cell morphology and differentiation pathways of some of the cell types that are a model for early murine embryogenesis *in vitro*. ES cells are isolated from the blastocyst inner cell mass and are cultured in the presence of LIF as compact dome shaped colonies. ES cells are induced to differentiate by the withdrawal of LIF (spontaneous) or by chemical induction, or aggregation and suspension culture. The morphology of some of the differentiated cell types generated by spontaneous (top left), retinoic acid (top right), 3-methoxybenzamide (bottom left), or dimethylsulfoxide (bottom right) induction are indicated.

X cells are generated and maintained by the culture of ES cells in the presence of MedII conditioned medium, with or without LIF. X cells are a morphologically homogeneous monolayer of flattened cells with characteristic nuclei and nucleolus. X cells revert to ES-like cells by MedII withdrawal in the presence of LIF. X cells differentiate with withdrawal of MedII and LIF, or by aggregation and suspension culture, with or without chemical induction. X cells differentiate to alternate mesodermal cell types in response to mesoderm inducing factors of the FGF and TGF β families. The likely embryonic cell types represented by cultures *in vitro* are defined by differentiation potential and gene expression.

ES cells and X cells were cultured for 3 days and photographed under phase contrast optics at 100 X magnification. Differentiated cells were cultured for 5 days and photographed under phase contrast optics at 100 X magnification.



1.6.5 Spontaneous and Chemical Induction of ES Cell Differentiation

A variety of cell types that possess characteristics of differentiated embryonic tissues, can be generated by the spontaneous or chemical induction of ES cell monocultures *in vitro* (Smith 1991, Thomas 1995). The culture of ES cells in the absence of LIF causes spontaneous and terminal differentiation to a range of cell types with ill-defined phenotypes. Residual populations of pluripotent cells often persist in these cultures, as nests of undifferentiated cells surrounded by terminally differentiated cells. Pluripotent cell nests are thought to be caused by the upregulation of LIF expression in differentiated cells, with the resulting feedback loop inhibiting further stem cell differentiation (Rathjen *et al.* 1990a, Mountford *et al.* 1994). Chemical inducers that are used to differentiate ES cells include retinoic acid (RA), 3-methoxybenzamide (MBA) or dimethylsulfoxide (DMSO). RA induction is efficient in this laboratory and generates two major cell types, large cells of fibroblastic phenotype and parietal yolk sac-like cells, while MBA differentiation produces a relatively uniform culture of epithelial-like cells (Smith 1991). DMSO induces ES cell differentiation relatively ineffectively in this laboratory, and results in a variety of poorly defined cell types (Thomas 1995). Residual pluripotent cell types can also persist in chemically induced ES cell cultures. Analysis of homeobox gene expression has indicated that distinct pluripotent cell subpopulations may exist as "differentiation intermediates" during these ES cell differentiation regimes (Thomas 1995).

1.6.6 Differentiation of ES Cells as Embryoid Bodies

Embryoid bodies, formed by the aggregation and culture of ES cells in solution, are globular structures that possess similar features to the early mouse embryo. Simple embryoid bodies are formed after 2-4 days of culture in suspension and contain a central mass of undifferentiated epiblast-like cells surrounded by an outer primitive endoderm-like layer. Prolonged suspension culture results in the generation of an additional ectodermal monolayer, which resembles the primitive ectoderm, surrounding an internal cavity (Doetschman *et al.* 1985, Shen and Leder 1992). A variety of structures that resemble post-gastrulation differentiated tissues can be formed by the differentiation of these cystic embryoid bodies, including beating muscle, neurons and blood islets (Evans and Kaufman 1981, Doetschman

et al. 1985, Robertson 1987). The differentiation of ES cells through embryoid body culture therefore provides an *in vitro* system to study pluripotent cell differentiation in an environment reminiscent of the complex embryo. The application of embryoid body systems to the study of early developmental events is validated by recent analysis of cardiac development (Lyons *et al.* 1995), proamniotic cavitation (Coucouvani and Martin 1995) and neuronal differentiation (Strübing *et al.* 1995).

1.6.7 Expansion of the Pluripotent Cell Model of Early Development: Lineage-specific Differentiation of ES Cells to X Cells in Response to Soluble Biological Factors

The demonstration that cultured ES cells retain the ability to colonise all embryonic tissues upon reintroduction into the blastocyst, including the germ line, indicates that ES cells are capable of generating and responding to the signals that control the proliferation and differentiation decisions of pluripotent cells during early embryogenesis. These attributes enable the manipulation of ES cell culture conditions to identify and isolate factors that commit pluripotent cells to specific developmental lineages, and establish *in vitro* developmental models for inaccessible embryonic events. These approaches have particular relevance to the investigation of early embryonic pluripotent cell development. As indicated above, the pluripotent cell pool is likely to proceed through multiple distinct stages from the ICM to mature germ cells, with developmental and molecular heterogeneity demarcating pluripotent cell "states". Conventional investigation of this system has been limited by the inaccessible nature of early murine development. The controlled differentiation of ES cells *in vitro* could potentially provide stem cell types representing those downstream of the ICM, and convenient systems to analyse pluripotent cell development and heterogeneity *in vitro*.

A screen of multiple conditioned media was carried out in this laboratory (Lake 1996). A morphologically distinctive cell type, termed "X" cells, was generated by the culture of ES cells in media conditioned by a human hepatocarcinoma cell line (MedII cells), in the presence or absence of LIF (Fig 1.7). The differentiation event caused by this conditioned medium (MedII) was relatively rapid, with the transition of the domed-colony ES cell phenotype to a uniform monolayer of X cells occurring within 24 hours. X cells shared distinctive

morphological features with the P19 embryonal carcinoma cell line (Rudnicki and McBurney 1987), such as monolayer growth and distinguishable individual cells with characteristic nuclei. As P19 embryonal carcinoma cells have properties related to the primitive ectoderm, extensive analysis has been undertaken to establish the relationship of X cells with early embryonic pluripotent cell populations.

1.6.8 X Cells as a Model for Primitive Ectoderm

Investigation of gene expression and differentiation potential indicated that X cells represented the pluripotent stem cells of the primitive ectoderm. X cells expressed markers of pluripotent cell types, *Oct-4*, alkaline phosphatase and SSEA-1 antigen (Lake 1996), but not markers of extraembryonic cell types, *α -fetoprotein* and *H19*, or markers of early mesoderm, *Evx-1* and *brachyury* (Bettess 1993). This indicated that X cells were a candidate pluripotent stem cell population. *Rex-1*, a marker of the ICM (1.5.2), was expressed by ES cells and was downregulated in X cells; whereas expression of *FGF-5*, a marker of primitive ectoderm (1.5.2), was induced during X cell formation (Bettess 1993, Lake 1996). This demonstrated that X cell gene expression was consistent with the primitive ectoderm, as distinct from the relationship between ES cells and the late stage ICM. Importantly, while the appearance of the X cell phenotype was rapid (1.6.7), changes in *Rex-1* and *FGF-5* expression only occurred after approximately four to six days, or 2 to 3 passages, in culture (Lake 1996).

X cells were shown to differentiate *in vitro* in response to withdrawal of LIF and MedII, or in response to chemical inducers, or following aggregation as embryoid bodies. These results confirmed that X cells are pluripotent (Lake 1996).

X cells exhibited a differentiation potential consistent with a cell type representative of the primitive ectoderm (see Fig. 1.5). For example, during embryogenesis ICM cells have the capacity to differentiate to primitive endoderm (1.2.3), which is the developmental precursor of visceral endoderm (1.3). Primitive ectoderm cells do not differentiate to primitive endoderm, but acquire the capacity to differentiate to mesodermal cell types in response to inducing factors during gastrulation (1.4.2). Aggregated X cells were restricted in their capacity to differentiate to visceral endoderm compared to ES cells, but demonstrated a greater capacity to generate mesodermal cell types (Lake 1996). Critically, X cells acquired the capacity to

respond to mesoderm inducing factors of the FGF and TGF β families, differentiating to alternate mesodermal cell types (Lake 1996), as defined by expression of the early mesodermal marker *brachyury* (Herrmann *et al.* 1990). These results demonstrated that the differentiation potential of X cells was consistent with the primitive ectoderm. Alteration in the developmental potential of X cells was detected after only 24 hours exposure to MedII, significantly earlier than differences in *Rex-1* and *FGF-5* expression.

The differentiation potential of X cells was analysed *in vivo*, by blastocyst injection. X cells were not capable of contributing to chimera formation, indicating that they were developmentally distinct from ES cells (Lake 1996). However, when X cells are cultured in LIF in the absence of MedII they revert to an ES cell-like phenotype. Reverted X cells regained similar morphology, gene expression and *in vitro* differentiation potential characteristics to ES cells (Lake 1996). Importantly, reverted X cells contributed to chimera formation following blastocyst injection. This demonstrated that pluripotency was maintained in X cells and that reversion generated an ES cell "state" (Lake 1996). This correlates with the predicted capacity of primitive ectoderm to convert to an ES cell-like state (Rossant 1993). Germ line transmission of reverted X cells has not been investigated. These results demonstrated that the differentiation potential of X cells *in vitro* and *in vivo* was consistent with a cell type representative of the primitive ectoderm.

Gene expression and differentiation potential therefore demonstrated that X cells are a distinct pluripotent cell type from ES/ICM cells and can be regarded as an *in vitro* model of the primitive ectoderm.

1.6.9 The Significance of X Cell Formation for *in vitro* Models of Pluripotent Cell Development

The development of detailed *in vitro* models for pluripotent cell renewal, progression or differentiation will simplify the experimental dissection of early embryogenesis. Several properties of X cell formation suggested the basis of an *in vitro* model for the development of the ICM to primitive ectoderm (Fig. 1.7). Firstly, X cell gene expression and differentiation potential were consistent with the primitive ectoderm. Importantly, this complemented the ES cell model of the ICM and expanded the range of pluripotent cell types represented *in vitro*.

Secondly, the appearance of the X cell morphology and alteration in differentiation potential *in vitro* and *in vivo* occurred within 24 hours of exposure to MedII. This indicated that significant developmental changes occurred early during the formation of X cells, generating a distinct pluripotent cell type from ES cells. Thirdly, the alteration in expression of *Rex-1* and *FGF-5* occurred after four to six days, indicating the progression to a primitive ectoderm-like gene expression. This suggested that X cell formation could be a multi-step transition, or a developmental progression of pluripotent cell types. Importantly, several distinct pluripotent cell types may exist during the ES to X cell transition, representing pluripotent cell subpopulations located between the ICM and primitive ectoderm.

The analysis of pluripotent cell subpopulations during the ES to X cell transition has several advantages over investigation of stem cell development during ES cell induction or in embryoid bodies. Firstly, ES cell and X cell cultures are “pure” populations of pluripotent cells, except for variable levels of background differentiation. This contrasts to multiple pluripotent and differentiated cell types that are present during ES cell induction or embryoid body culture. The analysis of pluripotent cell development is therefore likely to be more accessible in the ES to X cell transition than during ES cell differentiation systems. Secondly, the ES to X cell transition is likely to proceed in a more controlled progression than stem cell development during ES cell differentiation. This could enable the generation and investigation of specific stages of pluripotent cell development.

1.7 Generation and Application of Pluripotent Cell Specific Markers

A central requirement for successful dissection of the processes involved in the generation, maintenance and differentiation of the pluripotent cells of early embryogenesis is the isolation of cell type specific markers. This has parallels to the analysis of haematopoiesis, where a series of specific molecular and cell surface markers has led to the functional dissection of haematopoietic stem cell maintenance, proliferation and differentiation during development. Examples of the application of specific markers to characterise differentiation lineages include the tracing of haematopoietic progenitors during T- and myeloid-cell differentiation (reviewed in Barcena *et al.* 1995), and the investigation of complex differentiation hierarchies using lineage specific markers (Brady *et al.* 1995).

Molecular markers of specific pluripotent cell populations would be expected to have several major applications as listed below:

- i) Markers of pluripotent cell types are likely to refine our knowledge of embryogenesis. This could include the identification and analysis of specific pluripotent cell “states” during early embryogenesis. Pluripotent cell “states” could represent either distinct cell types, or within a homogeneous population, stem cell subtypes that respond to distinct developmental signals (1.5.1). This could describe the complexity of pluripotent cell subpopulations in the early mouse embryo and discernment of their spatial and temporal boundaries.
- ii) These markers could be used to follow developmental events that are currently inaccessible. For example, markers for pluripotent cell “states” could be used to identify the source and nature of specific inductive signals during embryogenesis, such as the regulation of primitive endoderm differentiation from the ICM, or signals that arise from the visceral endoderm that specify developmental fate in the primitive ectoderm.
- iii) Markers that define molecular heterogeneity of pluripotent cells would enable the functional analysis of the developing pluripotent cell pool, or of distinct stem cell subpopulations. These approaches would include the analysis of disrupted pluripotent cell development *in vitro*, or in induced or targeted mutations *in vivo*. For example, a lack of β -catenin leads to defects in primitive ectoderm at 7.5 d.p.c. (Haegel *et al.* 1995). Specific markers of primitive ectoderm, or ectodermal subpopulations, may enable the characterisation of aberrant pluripotent cell development in these embryos.
- iv) The gene products of pluripotent cell subtype-specific markers could be functionally involved in the development of these cell populations. Characterisation of these gene products would give insight into the molecular nature of pluripotent cell biology. Importantly, the function of these gene products could be assayed through mutational approaches *in vitro* and *in vivo*, and enable analysis of the development of distinct pluripotent cell subpopulations.
- v) The identification of markers for distinct pluripotent cell subpopulations or “states” could enable the isolation of specific promoter or enhancer elements for these genes. Analysis of these elements could define transcription factors that control marker expression and enable the characterisation of molecular hierarchies that regulate the development of pluripotent cells.

These promoters would also provide tools for the functional analysis of pluripotent cell development *in vitro* and during early embryogenesis.

vi) The application of ES cell technology and precise genome manipulation to non-murine mammalian species has been severely limited by an inability to isolate embryonic stem cell lines and demonstrate germ line transmission of cultured cells. Specific markers might enable precise definition of a transient murine stem cell population represented by ES cells and potential cross species equivalents, leading to enhanced stem cell isolation and characterisation procedures.

Pluripotent cell subpopulation specific markers would therefore define heterogeneity within the stem cell pool and provide powerful tools for the functional analysis and experimental perturbation of pluripotent cell development, both *in vitro* and during early embryogenesis. While specific markers could be derived from diverse cellular molecules such as protein or surface carbohydrates, advances in molecular biological techniques have resulted in mRNA transcript analysis becoming the most efficient and accessible route for marker generation.

1.7.1 Mutational Approaches to Define Markers for Pluripotent Cell Types

Analysis of spontaneous and induced mutations using traditional murine genetics has not enabled broad investigation of stem cell development, as early embryonic defects are often undetectable, leading to abortion or conceptus resorption. Primordial germ cells and their derivatives may be the only pluripotent cell type amenable to mutational analysis, with genetic defects indicated by the ensuing sterility. In this way, the *white spotting* (W) and *steel* (S) loci were identified as encoding factors critical for gametogenesis, the *c-kit* tyrosine kinase receptor and steel factor (*kit*-ligand) respectively (reviewed in Besmer *et al.* 1993). The expression of *c-kit* in primordial germ cells, spermatogonia and growing oocytes subsequently established it as a specific marker for these pluripotent cell types.

1.7.2 Generating Pluripotent Cell Markers by Homology Based Approaches

The identification of markers for mammalian pluripotent cells by the analysis of gene expression in lower organisms such as *Drosophila melanogaster* and *Xenopus laevis* is

difficult, as these organisms do not have a pluripotent cell pool. However, analysis of lower organisms and other mammalian systems has defined protein motifs and domains that are involved with the regulation of developmental processes. The isolation of transcripts with related domains from murine pluripotent cells could identify markers for pluripotent cell subpopulations, due to the tight expression patterns predicted for these control genes. These approaches suffer the limitation of only describing new members of known gene classes, but could define factors involved with the control of pluripotent cell development. For example, *Hesx1 (Rpx)* (1.5.1) was isolated from ES cell cDNA (Thomas and Rathjen 1992, Thomas *et al.* 1995) and a 7.5 d.p.c. embryonic library (Hermesz *et al.* 1996) by amplification with PCR primers designed against homeodomain consensus sequences.

1.7.3 Cloning of Pluripotent Cell Markers from Embryonic Tissue

The isolation of embryonic tissues representing individual stages of pluripotent cell development would be the most direct approach to generate a series of pluripotent cell subtype-specific markers. Unfortunately, the small nature of the developing pluripotent cell pool *in utero*, the existence of complex transient pluripotent cell types and the extraembryonic tissue environment, all contribute to technical difficulties associated with isolation of embryonic material. Limitations with this approach are therefore largely due to the requirement for more material than can be accessed easily. However, dissection based expression analysis is applicable to preimplantation and midgastrulation pluripotent cell stages, where some of these problems can be circumvented.

Sufficient preimplantation embryos for simple or subtracted library construction can be isolated through oviduct or uterine flushing (Rothstein *et al.* 1992, Temeles *et al.* 1994). Although misleading expression could be derived from trophectoderm and primitive endoderm lineages in blastocyst libraries, these libraries could enable gene isolation and definition of pluripotent cell markers for preimplantation embryonic stages. Sufficient primitive ectoderm can be isolated from 7.5 d.p.c. embryos to construct a subtracted library (Harrison *et al.* 1995), from which transient pluripotent cell markers can be isolated. The construction of libraries from specific pluripotent cell types is therefore likely to be a powerful method by which pluripotent cell markers will be isolated. However, embryonic dissection

and library construction are relatively intricate and time consuming processes, and the analysis of multiple different time points during stem cell development, to identify markers of temporal pluripotent cell heterogeneity, has not been undertaken.

1.7.4 Cloning of Pluripotent Cell Markers from Cell Lines *in vitro*

An alternate method to generate pluripotent cell specific markers on the basis of expression pattern is to utilise the proliferation and differentiation capabilities of embryonal carcinoma (EC) and ES cell lines *in vitro*. Transcripts that are restricted to undifferentiated cells, or are expressed early and transiently during stem cell differentiation, could have characteristics expected of pluripotent cell markers representing cell types from the ICM to the early primitive ectoderm. Advantages of *in vitro* pluripotent cell systems are that multiple cell types can be generated (1.6.4), and differences in gene expression between pluripotent and differentiated cell types can be identified. Furthermore, the expression patterns of candidate markers can be validated *in vitro*. This is a significant advantage over the isolation of markers directly from embryonic tissue (1.7.3), as these candidate markers generally have to be validated by more complex analysis *in vivo*. Therefore the EC/ES system is readily amenable to analysis of pluripotent cell gene expression, including the generation of subtracted libraries. An example of a pluripotent cell transcript isolated in this manner *in vitro* is *Rex-1*, a marker of the ICM (1.5.2) identified by the downregulation of expression in F9 EC cells induced to differentiate by retinoic acid (Hosler *et al.* 1989).

1.7.5 Identifying Pluripotent Cell Markers by Gene Traps

Gene trap vectors can be exploited to identify markers that are expressed by and restricted to undifferentiated ES cells. Important features of these vectors, such as an intron splice acceptor site, a promoterless selectable marker and *lacZ* gene, result in transcriptional marking of random loci by insertion when introduced to ES cells (reviewed in Gossler *et al.* 1989, Joyner *et al.* 1992). Insertion events are selected if they reside within transcriptionally active loci and regulation of the identified gene can be examined by *lacZ* staining, either *in vitro* or after reintroduction to the blastocyst. Although gene trap methods would seem ideal

to analyse pluripotent cell development and generate specific markers for the ICM and related stem cell types, interest in gene traps has focused on gastrulation and later stage development.

1.7.6 Cloning of Pluripotent Cell Markers by Differential Display Polymerase Chain Reaction

Given the inherent difficulties with traditional genetic approaches and only moderate success with transcript cloning methods, improved techniques to efficiently identify and isolate stem cell markers are required to address heterogeneity in pluripotent cell populations. Ideally, new approaches should be designed to avoid the problems associated with isolation of embryonic material and be capable of the simultaneous analysis of multiple transcripts from different pluripotent cell types, with markers of interest cloned on the basis of differential expression profiles between pluripotent cell populations. Additionally, such approaches should be combined with systems to validate the expression patterns of candidate markers *in vitro*.

Differential display polymerase chain reaction (DDPCR) was developed by Liang and Pardee (1992) and fulfils these requirements when combined with the advantages of pluripotent ES cell culture systems. DDPCR is schematically represented in figure 1.8.

In DDPCR procedures, first strand cDNA synthesis is primed with a poly d(T) primer from the start of the poly(A) tail. A subset of the total transcript pool is reverse transcribed. PCR amplification is carried out with the same 3' primer in combination with a short 5' primer of arbitrary sequence. A product of random length that terminates within the poly(A) tail is generated for an individual transcript. Annealing temperatures during amplification are kept low, to facilitate promiscuous binding of the short 5' primer during initial cycles. Reaction products are resolved by conventional gel electrophoresis. Selection of appropriate primer combinations and annealing conditions leads to the generation of 50-100 products in a single reaction. Multiple independent reactions can be compared to identify differential gene expression between mRNA samples from different cell populations. A large scale screen for differential gene expression is performed by varying the primer combinations, thus generating and comparing multiple alternate banding profiles. Markers exhibiting differential expression can be isolated by excision of a particular band, reamplification and cloning. A common

Figure 1.8 Differential Display Polymerase Chain Reaction.

Schematic representation of the identification and isolation of cDNA markers by differential display polymerase chain reaction (DDPCR).

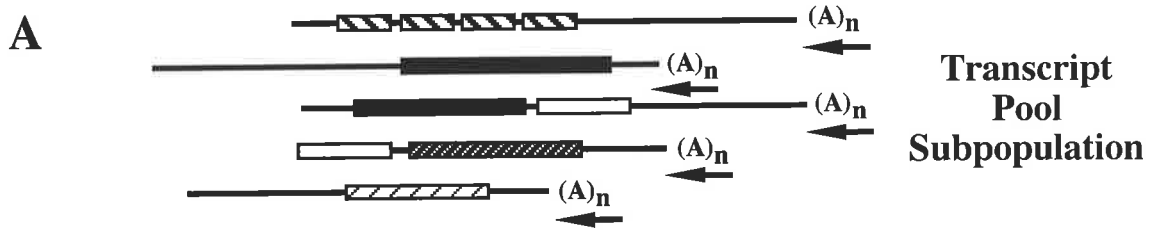
(A) mRNA representing a subpopulation of the total transcript pool is reverse transcribed with a poly(T) primer (arrow) anchored to the 5' end of the poly(A) tail. Boxed regions represent encoded protein domains.

(B) First strand cDNA template is amplified by PCR with the same 3' primer used in reverse transcription, in combination with an arbitrary 5' oligonucleotide (arrowhead) that binds at a randomly distributed position at low annealing temperatures. The reaction incorporates radioactive label to accurately visualise the generated fragments ("Hot PCR").

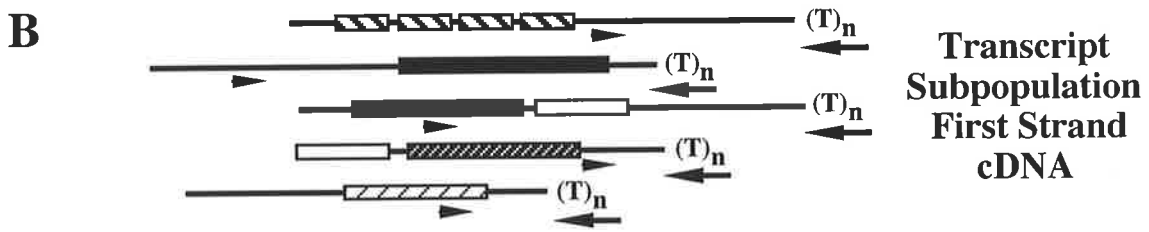
(C) The reaction products are separated on a denaturing acrylamide gel. Multiple RNA samples are compared simultaneously and differential expression between samples detected. Candidate sample-specific markers are excised from the dried down gel, eluted and reamplified without label, using the same primers.

(D) The candidate marker fragments represent regions at the 3' termini of transcripts. Reamplified products are cloned via ligation to an appropriate vector.

(E) The ligation reaction of a candidate marker represents a library of different fragments, generated by specific marker reamplification and inherent variable levels of background amplification products. Identification of individual clones that verify the original expression profile is performed via validity procedures such as northern blot or *in situ* hybridisation.



Reverse Transcription



Hot PCR

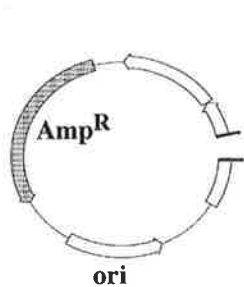
C

Resolve Products and Identify Candidate Markers

Fragment Excision, Elution and Reamplification

D

Candidate Marker Fragment



Restriction Digestion and Ligation

E

Library of Cloned Fragments

Screening and Verification

MARKER

problem with DDPCR is the contamination of markers with background amplification products, which can be reamplified and cloned in place of the required transcript. This necessitates the development of screening procedures for identification of the clones which represent the desired transcript (Callard *et al.* 1994).

The effectiveness of DDPCR in detecting differential gene expression between cell types and tissues has been established in a wide range of systems. Examples of these include analysis of haematopoietic development in the yolk sac and embryoid bodies (Guimarães *et al.* 1995), androgen dependence in prostate carcinoma (Blok *et al.* 1995), cocaine and amphetamine regulation in the rat brain (Douglass *et al.* 1995), immune response in *Drosophila* (Asling *et al.* 1995), oestrogen regulation in rat hepatoma cells (Diel *et al.* 1995), TGF β regulation in human osteoblasts (Subramaniam *et al.* 1995) and human breast cancer (Liang *et al.* 1992). DDPCR has previously been used to analyse murine preimplantation embryogenesis (Zimmermann and Schultz 1994, Davis *et al.* 1996) and is likely to enhance the analysis of gene expression during these stages of embryogenesis. However, the inability to distinguish the gene expression in pluripotent cells from that of differentiated cell types, the trophoblast and primitive endoderm, could complicate the identification of pluripotent cell specific markers by this approach.

The use of DDPCR is most effective when several closely related cell types are available, as is the case with the ES cell model of development, because most of the displayed transcripts are common and subtle molecular differences between cell types can be examined. These factors therefore indicate that DDPCR is likely to be a powerful technique for identifying restricted gene expression in the pluripotent or differentiated cell types generated from ES cells *in vitro*. Key advantages of this approach are the ability to screen for transient expression in pluripotent cell subtypes present during the ES to X cell conversion and to exclude transcripts that are expressed by terminally differentiated cells. A panel of markers that identifies multiple transient pluripotent cell types *in vitro* could enable the definition and functional investigation of pluripotent cell heterogeneity (1.7) during the ICM to primitive ectoderm stages of embryogenesis.

1.8 Aims and Approaches

Pluripotent cell populations perform critical functions during the development of the early mammalian embryo, generating precursors for formation of the three germ cell layers, establishment of tissue pattern and axis formation during gastrulation, and ensuring the continuation of the species through the germ line. Experimental manipulation of murine embryos has demonstrated that the developing pluripotent cell pool is tightly but flexibly regulated and can respond to embryonic signals, environmental cues or injury by altering cellular fate. Correspondingly, mounting developmental and molecular evidence points to the existence of multiple pluripotent cell “states” or subtypes, with the stem cell pool of early development featuring both temporal and spatial heterogeneity. The elucidation of individual pluripotent cell subtypes is an essential precursor to the understanding of the regulation of the stem cell pool and the decisions that govern pluripotent cell maintenance, progression or differentiation within the early embryo.

Embryonic stem (ES) cells are derived from the pluripotent cells of the early mouse embryo and represent the pluripotent cells of the ICM/early epiblast. ES cells retain developmental pluripotency, as demonstrated by competency to colonise all murine tissues including the germ line, following reintroduction to the blastocyst. Therefore ES cells remain sensitive to all the developmental signals that are present during early embryogenesis, properties that may be exploited to examine pluripotent cell biology *in vitro*. ES cells can be maintained in the undifferentiated state by LIF and related cytokines, or differentiated through cytokine withdrawal, chemical inducers or embryoid body aggregation, thus mimicking embryonic pluripotent cell proliferation and differentiation decisions. Potential exists for the generation of subtly different pluripotent cell types from ES cells *in vitro*, either transiently during differentiation procedures, or through the manipulation of culture conditions, as demonstrated by the generation of X cells, a model for the pluripotent primitive ectoderm. Therefore the ES cell model of early development is likely to provide multiple pluripotent cell types for the investigation of restricted gene expression and generation of stem cell subtype-specific markers. Such analysis could define markers that address the molecular nature of temporal and spatial heterogeneity within pluripotent cells during the development of the ICM to primitive ectoderm.

The general aim of this work was to develop methods for the identification of pluripotent cell markers for transient stem cell populations in the developing mouse embryo and hence provide molecular insight to the question of pluripotent cell heterogeneity.

The approach taken was to identify differentially expressed genes during the ES cell to X cell transition, to define markers for pluripotent cell subpopulations *in vivo*. This approach has significant advantages over alternate strategies for the identification of markers for pluripotent cell types. Firstly, distinct pluripotent cell subpopulations are predicted to exist during the ES to X cell transition. The analysis of gene expression at timed points during the ES to X cell transition could therefore address molecular heterogeneity during the inner cell mass to primitive ectoderm transition *in vivo*. Secondly, the pluripotent cell types present during the ES to X cell transition are closely related and are therefore amenable to DDPCR analysis. This would enable multiple transcripts to be screened simultaneously and transcripts cloned on the basis of differential expression profiles between ES cells and X cells. Thirdly, candidate markers of pluripotent cell subpopulations could be verified by gene expression analysis *in vitro*, prior to expression analysis during early embryogenesis.

CHAPTER 2

Materials and Methods

CHAPTER 2

MATERIALS AND METHODS

2.1 Abbreviations

| | |
|----------------|--|
| APS | ammonium persulphate |
| ATP | adenosine triphosphate |
| bp | base pair |
| BCIG | 5-bromo-4-chloro-3-indolyl-beta-D-galactosidase |
| BCIP | 5-bromo-4-chloro-3-indolyl-phosphate |
| bisacrylamide | N'N'-methylene-bisacrylamide |
| BSA | bovine serum albumin |
| CHAPS | 3-[(Cholamidopropyl)dimethylammonio]-1-propane-sulfonate |
| CIP | calf intestinal phosphatase |
| c.p.m. | counts per minute |
| d.p.c. | days post-coitum |
| DIG | digoxigenin |
| DMF | dimethylformamide |
| DMSO | dimethylsulfoxide |
| DTT | dithiothreitol |
| <i>E. coli</i> | Escherichia coli |
| EtBr | ethidium bromide |
| EDTA | ethylene diamine tetra acetic acid |
| FCS | foetal calf serum |
| HEPES | N-2-hydroxyethyl piperazine-N-ethane sulphonic acid |
| IPTG | isopropyl- β -D-thiogalactopyranoside |
| kb | kilobase pairs |
| MBA | 3-methoxybenzamide |
| min. | minute(s) |
| MOPS | 3-[N-morpholino]propane sulphonic acid |
| NBT | nitro blue tetrazolium chloride |
| NP-40 | Nonidet P40 |
| O/N | overnight |
| PBS | phosphate buffered saline |
| PEG | polyethylene glycol |
| PFA | paraformaldehyde |
| RA | <i>all-trans</i> -retinoic acid |
| RNase | ribonuclease |
| rNTPs | ribonucleotide phosphates |

| | |
|----------|--|
| RT | room temperature |
| SDS | sodium dodecyl sulphate |
| sec. | second |
| TEMED | N, N, N', N'-teramethyl-ethenediamine |
| Tween 20 | polyoxyethylenesorbitan monolaurate |
| U | unit(s) |
| UV | ultra violet |
| X-gal | 5-bromo-4-chloro-3-indol- β -D-galactopyranoside |

2.2 Materials

2.2.1 Chemicals and Reagents

Sigma Chemical Co. supplied the following chemicals: Acrylamide, bisacrylamide, agarose (Type 1), ampicillin, BSA, EtBr, EDTA, Heparin, MOPS, RA, rNTPs, salmon sperm DNA, SDS, TEMED, Tris base, tRNA (from brewer's yeast).

Sources for other important reagents were as follows: IPTG, BCIG; Progen. BCIP, Anti-DIG Fab antibody-alkaline phosphatase conjugate, NBT and glycogen; Boehringer Mannheim. APS, phenol, NP-40 and PEG 6000; BDH Chemicals. Paraformaldehyde; Merck. Poly dT₁₂₋₁₈ primer and Sepharose CL-6B; Pharmacia. Gluteraldehyde; Probing and Structure. Dithioerythritol and DTT; Diagnostic Chemicals Ltd. Tween 20; ICN. Sequagel 6; National Diagnostics.

All chemicals and reagents were of analytical grade.

2.2.2 Radiochemicals

[α -³²P] dATP (3000 Ci/mmol), [α -³³P] dATP (1500 Ci/mmol), [α -³⁵S] dATP (1000-1500 Ci/mmol) and [α -³²P] UTP (3000 Ci/mmol) were supplied by Bresatec Ltd.

2.2.3 Kits

| | |
|--------------------------------|-----------------------------|
| T ⁷ Sequenase kit | Pharmacia |
| Gigaprime kit | Bresatec Ltd. |
| Bresaclean kit | Bresatec Ltd. |
| RPA II kit | Bresatec Ltd. |
| ampliFINDER RACE kit | Clontech |
| 7-deaza-dGTP decompression kit | USB, Amersham Life Sciences |

2.2.4 Enzymes

Restriction endonucleases were supplied by Pharmacia and New England Biolabs. Other enzymes were obtained from the following sources: CIP and Proteinase K; Boehringer Mannheim. *E. coli* DNA polymerase I (Klenow fragment), Taq Polymerase, RNasin, DNase 1, T4 DNA ligase, SP6 RNA polymerase and T7 RNA polymerase; Bresatec. RNase A; Sigma. RNase T1; Ambion. T7 DNA polymerase and T3 RNA polymerase; Pharmacia. Avian Myeloblastosis Virus (AMV) reverse transcriptase; Molecular Genetic Resources. T4 RNA ligase; New England Biolabs. Pfu DNA Polymerase; Stratagene. Superscript II reverse transcriptase; Gibco BRL.

2.2.5 Buffers and Solutions

AP buffer: 100 mM NaCl, 50 mM MgCl₂, 100 mM Tris-HCl pH 9.5, 0.1% Tween 20

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

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

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

GEB: 10 mM Tris-HCl, 1 mM EDTA, 0.1% SDS, pH 7.5

GLB: 50% glycerol, 0.1% SDS, 0.05% bromophenol blue, 0.05% xylene cyanol

GTE: 50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, pH 7.6

β-mercaptoethanol/PBS: 100 mM β-mercaptoethanol (Sigma) in 14 ml PBS. Fresh solution made every two weeks.

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

PBS: 136 mM NaCl, 2.6 mM KCl, 1.5 mM KH₂PO₄, 8 mM Na₂HPO₄, pH 7.4.
Sterilised by autoclaving (20 psi for 25 min. at 140°C).

RIPA: 150 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl pH 8.0, 0.5% sodium deoxycholate, 0.1% SDS

1 x SD buffer: 33 mM Tris-HAc pH 7.8, 62.5 mM KAc, 10 mM MgAc, 4 mM spermidine, 0.5 mM dithioerythritol

SSC: 150 mM NaCl, 15 mM sodium citrate

STET: 50 mM Tris-HCl pH 8.0, 50 mM EDTA, 8% sucrose, 5% Triton X-100

TAE: 40 mM Tris-acetate, 20 mM NaAc, 1 mM Na₂EDTA, pH 8.2

1 x Taq buffer: 67 mM Tris-HCl pH 8.8, 16.6 mM (NH₄)₂SO₄, 0.2 mg/ml gelatin, 0.45% Triton X-100

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

TBST: 136 mM NaCl, 2.7 mM KCl, 25 mM Tris-HCl pH 7.5, 0.1% Tween 20

Tfb 1: 30 mM KAc, 100 mM RbCl, 10 mM CaCl₂, 15% (v/v) Glycerol, pH 5.8
(adjusted with 0.2 M Acetic Acid)

Tfb 2: 10 mM MOPS, 75 mM CaCl₂, 10 mM RbCl, 15% Glycerol, pH 6.5 (adjusted with 1 M KOH)

- TNM: 30 mM Tris-HCl pH 7.6, 150 mM NaCl, 15 mM MgCl₂, 0.4% NP40
- Trypsin: 0.1% trypsin (Difco) and 1 x EDTA Versene buffer solution (CSL). Sterilised by filtration through a 0.2 µM filter (Whatmann).
- TUNES: 10 mM Tris-HCl pH 8.0, 7 M urea, 0.35 M NaCl, 1 mM EDTA, 2% SDS

2.2.6 Plasmid Vectors

Cloning vectors

i) pBluescript KS and SK were kind gifts from Dr. Blair Hopwood.

Cloned DNA sequences

- ii) The *GAPDH* (*mGAP*) cDNA clone in pGEM3Z was a kind gift from Prof. Peter Rathjen and contained a 300 bp *HindIII/PstI* fragment from the 5' end of the mouse *GAPDH* gene (Rathjen *et al.* 1990b).
- iii) The *Oct-4* cDNA clone in pBluescript was a kind gift from Dr. Hans Schöler and contained a 462 bp *StuI* cDNA fragment of positions 491 to 953 of the *Oct-4* cDNA sequence (Schöler *et al.* 1990b). Double strand *Oct-4* probes were isolated by *EcoRI/XhoI* digestion.

2.2.7 Riboprobe Templates

Riboprobe templates were prepared as described in section 2.3.6, DIG labelled riboprobes were generated as described in section 2.3.29.

i) *Icm1* (A03/360)

Riboprobes for *in situ* hybridisation were generated from the A03/360 5' plasmid containing a 256 bp *Icm1* insert (3.4). The antisense transcript was generated by *HindIII* digestion and transcription with T7 RNA polymerase. The sense transcript was generated by *BamHI* digestion and transcription with T3 RNA polymerase.

ii) *Psc1* (B04/400)

Riboprobes for *in situ* hybridisation, northern blot and ribonuclease protection were generated from the B04/400 plasmid containing a 458 bp insert (3.4), corresponding to positions 1661 to 2041 of the *Psc1* cDNA. The antisense transcript was generated by *BamHI* digestion and transcription with T3 RNA polymerase. The sense transcript was generated by *HindIII* digestion and transcription with T7 RNA polymerase.

iii) *Oct-4*

Riboprobes for *in situ* hybridisation were generated from the 462 bp *Oct-4* cDNA clone (2.2.6iii). The antisense transcript was generated by *EcoRI* digestion and transcription with T3 RNA polymerase. The sense transcript was generated by *XhoI* digestion and transcription with T7 RNA polymerase.

iv) *GAPDH*

Riboprobes for ribonuclease protection were generated from the 300 bp *GAPDH* cDNA clone (2.2.6ii). The antisense transcript was generated by *BamHI* digestion and transcription with SP6 RNA polymerase.

2.2.8 Oligonucleotides

DNA primers were synthesized by Bresatec Ltd. 5' arbitrary DDPCR primers, OPA and OPB 10mers kits, were purchased from Operon technologies. Primer sequences are shown 5'-3', except for "Anchor" which is 3'-5'. *EcoRI* restriction sites are in bold. Generation of the 3' homeodomain primer from the homeodomain consensus is indicated.

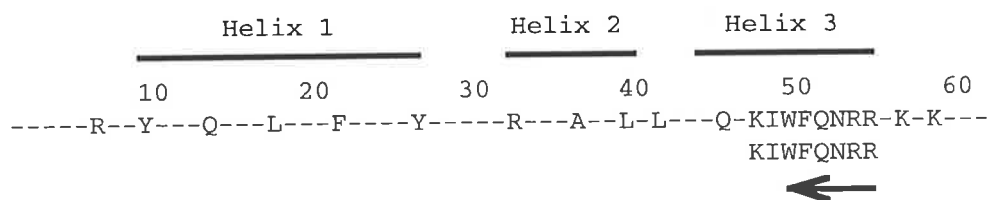
General sequencing primers

T7: TAATACGACTCACTATAGGGAGA
 T3: ATTAACCCTCACTAAAGGGA

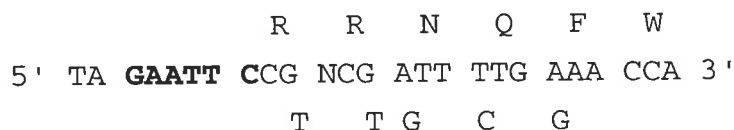
DDPCR primers

3' homeodomain primer, 3'-hp:

The homeodomain consensus (Scott *et al.* 1989).



Dashes indicate variable residues, conserved residues at invariant positions are indicated. Numbers indicate size of the domain, dark lines indicate helices within the homeodomain structure. KIWFQNRR indicates the highly conserved region in the third α -helix of all homeodomain proteins. 3'-hp (arrow) is a degenerate complimentary oligonucleotide to the sequence WFQNRR:



3'-hp: TAGAATTCCG/TNCG/TA/GTTC/TTGA/GAACCA

(N=A, C, G, or T; / indicates position of degeneracy)

Arbitrary 5' primers:

| | | | |
|---------|------------|---------|------------|
| OPA-01: | CAGGCCCTTC | OPB-01 | GTTTCGCTCC |
| OPA-02: | TGCCGAGCTG | OPB-02 | TGATCCCTGG |
| OPA-03: | AGTCAGCCAC | OPB-03 | CATCCCCCTG |
| OPA-04: | AATCGGGCTG | OPB-04: | GGACTGGAGT |
| OPA-05: | AGGGGTCTTG | OPB-05: | TGCGCCCTTC |
| OPA-06: | GGTCCCTGAC | OPB-06: | TGCTCTGCCC |
| OPA-07: | GAAACGGGTG | OPB-07: | GGTGACGCAG |

| | | | |
|---------|------------|---------|------------|
| OPA-08: | GTGACGTAGG | OPB-08: | GTCCACACGG |
| OPA-09: | GGGTAACGCC | OPB-09: | TGGGGGACTC |
| OPA-10: | GTGATCGCAG | OPB-10: | CTGCTGGGAC |
| OPA-11: | CAATCGCCGT | OPB-11: | GTAGACCCGT |
| OPA-12: | TCGGCGATAG | OPB-12: | CCTTGACGCA |
| OPA-13: | CAGCACCCAC | OPB-13: | TTCCCCCGCT |
| OPA-14: | TCTGTGCTGG | OPB-14: | TCCGCTCTGG |
| OPA-15: | TTCCGAACCC | OPB-15: | GGAGGGTGT |
| OPA-16: | AGCCAGCGAA | OPB-16: | TTTGCCCGGA |
| OPA-17: | GACCGCTTGT | OPB-17: | AGGGAACGAG |
| OPA-18: | AGGTGACCGT | OPB-18: | CCACAGCAGT |
| OPA-19: | CAAACGTCGG | OPB-19: | ACCCCCGAAG |
| OPA-20: | GTTGCGATCC | OPB-20: | GGACCCTTAC |

PscI RACE and RT/PCR primers

| | |
|----------------|--|
| Anchor: | 3'-NH ₃ -GGAGACTTCCAAGGTCTTAGCTATCACTTAAGCAC-P-5' |
| Anchor primer: | CTGGTTCGGCCACCTCTGAAGGTCCAGAATCGATAG |
| 582: | TAGAATTCGTACCGCTCATAGTCTCTCCAC |
| 1064: | TAGAATTCGGCAGAGCAACTTCATCAACAATA |
| 1275: | TGGGAAGCACAACAGAAGGT |
| 36: | TTTACTTTGATTGTTGTTCC |
| 5' primer: | ATGAATTCCCAGGGAGTAGGTGAAGT |

2.2.9 Bacterial Strains

All recombinant plasmids were maintained in *E. coli* DH5 α (a kind gift from Dr. Blair Hopwood). Library screening was carried out in host *E. coli* BB4 (purchased from Clontech). O/N cultures diluted with an equal volume of 80% glycerol were used for long term storage of stock strains and transformed bacteria at -80°C.

| | |
|----------------|--|
| DH5 α : | <i>supE44</i> <i>deltalac</i> U169 (<i>phi80 lacZdeltaM15</i>) <i>hsdR17</i> <i>recA1</i> <i>endA1</i> <i>gyrA96</i> <i>thi-1</i> <i>relA1</i> |
| BB4: | <i>supF58</i> <i>supE44</i> <i>hsdR514</i> <i>galK2</i> <i>galT22</i> <i>trpR55</i> <i>metB1</i> <i>tonA</i> <i>deltalac</i> U169 F'[<i>proAB</i> ⁺ <i>lacI</i> ^q <i>lacZdeltaM15</i> Tn10(<i>tet</i> ^r)] |

2.2.10 Bacterial Growth Media

Growth media were 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. Luria (L) broth: 1% (w/v) Bacto-tryptone (Difco), 0.5% (w/v) yeast extract (Difco), 1% (w/v) NaCl, pH 7.0 (adjusted with NaOH).

LMM broth: 1% (w/v) Bacto-tryptone (Difco), 0.5% (w/v) NaCl, 0.4% (w/v) maltose, 0.2% MgSO₄.

Psi broth: 2% (w/v) Bacto-tryptone (Difco), 0.5% (w/v) yeast extract (Difco), 0.5% (w/v) MgSO₄, pH 7.6 (adjusted with 1 M KOH).

Solid Media: Agar plates were prepared by supplementing the above media with 1.5% Bacto-agar (Difco). LMM agarose for library screening was prepared by dissolving 0.7 g agarose/100 ml LMM broth.

Ampicillin (100 µg/ml) was added where appropriate for growth of transformed bacteria to maintain selective pressure for recombinant plasmids.

2.2.11 Tissue Culture Cell Lines and Media

ES Cell Lines

The ES cell lines used throughout the course of this work were MBL-5 and D3 (Lindsay Williams, Ludwig Institute, Melbourne, Australia).

Other Cell Lines

MedII cells (unpublished data).

Media

ES DMEM medium: 67.4 g Dulbecco's Modified Eagle Medium (Gibco BRL), 18.5 g NaHCO₃ and 6.25 ml (40 mg/ml) Gentamicin (Delta West) dissolved in 5 litres sterile water.

Incomplete ES cell medium: 90% ES DMEM medium, 10% FCS (Commonwealth Serum Laboratories), 1% glutamine, 0.1% β-mercaptoethanol/PBS.

Complete ES cell medium: Incomplete ES cell medium with 1% COS cell conditioned medium containing LIF (Smith, 1991).

X cell medium: (1:1) MedII conditioned medium:ES cell medium (complete or incomplete).

MedII conditioned medium: medium was isolated from MedII cells cultured in incomplete ES cell medium for 5 days.

All tissue culture media were filter sterilised.

2.2.12 DNA Markers

*Hpa*II digested pUC19 markers were purchased from Bresatec Ltd. Band sizes (bp): 501, 489, 404, 331, 242, 190, 147, 111, 110, 67, 34, 26.

*Eco*R1 digested SPP-1 phage markers were purchased from Bresatec 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.

2.2.13 Miscellaneous Materials

| | |
|---------------------------------------|------------------------|
| 3MM chromatography paper | Whatmann Ltd. |
| X-Omat AR and Bio-max diagnostic film | Kodak |
| Nytran nylon membrane (0.45 μ M) | Schleicher and Schuell |
| Petri-dishes, 6-, 24-, 96-well trays | Corning, Falcon |
| Terisaki plates | Bresatec Ltd. |
| Genescreen filters | Dupont |

2.3 Molecular Methods

2.3.1 Restriction Endonuclease Digestion of DNA

Plasmid DNA was digested in SD buffer with 1-2 U of enzyme/ μ g DNA and incubation at the appropriate temperature for 30 min. to 2 hours. Complete digestion of DNA was assayed by agarose gel electrophoresis.

2.3.2 Agarose Gel Electrophoresis

Agarose gel electrophoresis (0.8% to 3% w/v agarose in TAE) was carried out using horizontal mini-gels prepared by pouring 10 ml of gel solution onto a 7.5 cm x 5.0 cm glass microscope slide. Agarose mini-gels were submerged in 1 x TAE and samples containing 1 x GLB were electrophoresed at 100 mA. Nucleic acid was visualised by staining gels with EtBr (4 mg/ml in water) and exposure to medium wavelength UV light. Appropriate bands were removed from preparative gels using a sterile scalpel blade.

2.3.3 Polyacrylamide Gel Electrophoresis

i) Non-denaturing gels

A 30% acrylamide stock solution was prepared by dissolving 29 g acrylamide and 1 g bis-acrylamide in a final volume of 100 ml MQ water. Polyacrylamide gels were prepared by pouring a 50 ml gel solution containing 5%-10% acrylamide in 1 x TBE between 15 cm x 16 cm glass plates separated by 1 mm spacers. Prior to pouring the gel, 500 μ l of 10% APS and 60 μ l TEMED were added to the gel solution to initiate polymerisation. The gel was allowed to set for 30-60 min. and pre-electrophoresed at 30 mA in 1 x TBE for 10 min. DNA samples containing 1 x GLB were electrophoresed at 30 mA. After electrophoresis, the glass plates

were separated and the gel stained with EtBr, destained in water for 10 min. and DNA bands excised under medium wavelength UV light using a sterile scalpel blade.

ii) Denaturing gels

6% polyacrylamide sequencing gels (20 cm x 40 cm x 0.4 mm, or 35 cm x 40 cm x 0.4 mm) were prepared from Sequagel 6 solutions according to the manufacturer's instructions. Once the gel had set (approximately 20 min.), the comb was removed and wells flushed with water. Denaturing gels were pre-electrophoresed for 30 min. at 50 W (small gel) or 75 W (large gel), the wells were flushed with 1 x TBE before loading samples and electrophoresis at the same setting (about 50°C). After the gel had run, the glass plates were prised apart and the gel transferred to dry 3MM Whatmann paper. The gel was dried down at 65°C under vacuum. Radioactivity was detected by exposure to X-ray film at RT or -80°C with intensifying screens, or by exposure to phosphorimager screens.

2.3.4 Purification of Linear DNA Fragments

Linear DNA fragments greater than 200 bp in length were purified using the Bresaclean kit according to the manufacturer's instructions. Smaller DNA fragments were separated by non-denaturing gel electrophoresis (2.3.3i), excised and eluted O/N in 400 µl GEB. The following day, the eluate was transferred to a fresh eppendorf tube and the DNA was precipitated by addition of 40 µl 3 M NaAc, 1 µl glycogen and 1 ml ethanol and chilled at -80°C for 30 min. After centrifugation at 14 K for 15 min. at RT the DNA was resuspended in MQ H₂O.

2.3.5 Agarose Minigels for DNA Southern Blot

Agarose minigels for southern blot analysis were prepared as described in section 2.3.2. DNA was transferred to Nytran Nylon membrane as described in section 2.3.24, except that 0.4 M NaOH was used for the capillary transfer and weights were 0.1-0.5 kg.

2.3.6 Preparation of Vector DNA

i) Standard Vectors

5-10 µg midi-prep DNA (2.3.11) was digested with the appropriate enzyme(s) in 1 x SD buffer as described in section 2.3.1. 1 U of CIP, 10 µl 10 x CIP buffer (Boehringer Mannheim Biochemicals) and MQ H₂O to a final volume of 100 µl were added to restriction endonuclease digestions and incubated at 37°C for 1 hour. Vectors were purified by agarose gel electrophoresis (2.3.2) and Bresaclean (2.3.4).

ii) DDPCR Cloning Vectors

The cloning vector for DDPCR product ligations (pBluescript KS *EcoRI*/T) was prepared in the following manner. 10-20 µg pBluescript KS was digested with *EcoRV* (2.3.1), extracted with an equal volume of 1:1 phenol:chloroform and the aqueous layer precipitated with 1/10th volume NaAc pH 5.2, 2 µl glycogen and 2.5 volumes ethanol (-80°C

for 30 min.). Vector DNA was resuspended and T-tailed (described in Austbel *et al.* 1994) in a 50 μ l reaction containing 1.5 mM MgCl₂, 1 x Taq buffer, 2 mM dTTP and 1 U Taq polymerase, at 72°C for 2 hours under mineral oil. The pBluescript KS T vector was extracted and precipitated (as above) and resuspended in 20 μ l MQ H₂O. pBluescript KS T was digested with *Eco*RI (2.3.1) to generate the pBluescript KS *Eco*RI/T cloning vector. The DDP-PCR cloning vectors were purified by agarose gel electrophoresis (2.3.2) and Bresaclean (2.3.4).

2.3.7 DNA Ligation Reactions

Ligation reaction contained 25-40 ng purified vector (2.3.6), 50-100 ng DNA insert (2.3.4), ligation buffer (50 mM Tris-HCl pH 7.4, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP) and 1 U T4 DNA ligase, and were incubated at RT for 1-5 hours.

2.3.8 Preparation of Competent Cells

A single *E. coli* colony was used to inoculate 5 ml Psi broth and grown in an orbital shaker O/N at 37°C. A 1:30 subculture was made in 15 ml Psi broth and grown for 90 min. at 37°C, or until O.D. 0.6 was obtained. A 1:20 subculture was made in 100 ml Psi broth and grown at 37°C until the O.D. was 0.5-0.6. The bacterial cells were chilled on ice for five min. and harvested by centrifugation at 6 K, for 5 min., at 4°C (SS34 rotor in Sorvall RC-5 centrifuge). The cells were resuspended in 0.4 volume of Tfb 1 and chilled on ice for a further 5 min. The cells were re-harvested by centrifugation (as above) and resuspended in a 0.04 volume of Tfb 2. After incubation on ice for 15 min., 50 μ l 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.

2.3.9 Transformation of Competent Bacterial Cells

Approximately 10 ng of ligated DNA (2.3.7) was added to 50 μ l of thawed competent cells and placed on ice for 30 min. The cells were heat shocked at 42°C for 2 min. and then incubated in 1 ml of LB at 37°C 30 min. The cells were pelleted by centrifugation and the supernatant decanted. The cells were resuspended in the residual supernatant, spread onto L broth ampicillin (100 μ g/ml) plates and transformant colonies grown O/N at 37°C. Where blue/white colour selection was required, plates were prepared by spreading with 50 μ l IPTG (50 μ g/ml) and 20 μ l X-gal (50 μ g/ml in DMF).

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

1.0 ml L broth ampicillin (100 μ g/ml) was inoculated with a single transformant 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 aerated and resuspended in 200 μ l STET buffer by vortexing. The bacteria were lysed at 100°C for 45 sec. Samples were centrifuged for 10

min. and the cell debris and chromosomal DNA removed with a sterile toothpick. Plasmid DNA was precipitated by addition of 200 μ l isopropanol and incubation at -20°C for 5 min. Plasmid DNA was pelleted by centrifugation for 10 min. and resuspended in 20 μ l MQ H_2O .

2.3.11 Rapid Large Scale Preparation of DNA (50 ml Midi-prep)

50 ml L broth ampicillin (100 $\mu\text{g}/\text{ml}$) was inoculated with a single transformant colony and incubated O/N at 37°C with shaking. Cell cultures were transferred to oakridge tubes, pelleted by centrifugation at 5 K, for 5 min., at 4°C (SS34 rotor in Sorvall RC-5 centrifuge) and resuspended in 3 ml GTE. 6 ml of fresh lysis solution (0.2 M NaOH / 1% SDS) was added, mixed gently by inversion and placed on ice for 5 min. 4.5 ml of acetate solution (3M KAc / 2M HOAc pH 5.8) was added, mixed gently by inversion and placed on ice for 5 min., then mixed vigorously and placed on ice for 15 min. The cell debris was pelleted by centrifugation (12 K, 10 min., 4°C ; SS34 rotor in a Sorvall RC5 centrifuge) and the supernatant was transferred to a new tube. Nucleic acid was precipitated at RT for 5 min. by the addition of 8 ml iso-propanol, followed by centrifugation (12 K, 10 min, 4°C ; SS34 rotor in a Sorvall RC5 centrifuge). The pellet was drained, resuspended in 400 μ l H_2O and transferred to a eppendorf tube. RNA was removed by the addition of 2 μ l RNaseA (20 mg/ml) and incubation at 37°C for 30 min. 8 μ l 10% SDS and 2.5 μ l Proteinase K (20 mg/ml) was added and incubated at 37°C for 15 min. Plasmid DNA was extracted once with an equal volume of 1:1 phenol:chloroform, once with chloroform and was precipitated with 100 μ l 7M NH_4Ac and 1 ml ethanol (-20°C , 30 min.). Plasmid DNA was pelleted, dried *in vacuo* and resuspended in 200 μ l H_2O .

2.3.12 Generation of Sequencing Subclones

Sequencing of large cDNA fragments was carried out by generating a series of overlapping 200-600 bp subclones. 1 μg of isolated cDNA insert (2.3.4) was digested with either *AluI*, *HaeIII*, *RsaI*, *HpaII* or *Sau3A I* (2.3.1) in 20 μ l. The reactions were terminated by incubating at 65°C for 5 min. and digested fragments were endfilled by adding 2 μ l 10 x SD buffer, 4 μ l 5 mM dNTPs (dATP, dCTP, dGTP, dTTP), 14 μ l MQ H_2O and 2 U Klenow DNA polymerase and incubation at 37°C , 15 min. Digestion products from blunt cutting enzymes (*AluI*, *HaeIII*, *RsaI*) were endfilled due to 5' overhangs generated during isolation of the original cDNA insert. Fragments were purified on 10% acrylamide TBE gels (2.3.3i) and ligated (2.3.7) into a blunt cut vector.

2.3.13 Double Stranded Sequencing of Plasmid DNA

i) Denaturation of plasmid DNA

15 µl mini-prep (2.3.10) or midi-prep DNA (2.3.11) was incubated at 37°C for 15 min. with 1.5 µl RNase A (20 mg/ml). Plasmid DNA was denatured by the addition of 3.5 µl 1 M NaOH / 1 mM EDTA and incubation at 37°C for 15 min. Denatured plasmid was purified by centrifugation at 1.8 K for 3 min. on a Sepharose CL-6B spin column.

ii) Sequencing of plasmid DNA

Dideoxy sequencing reactions were carried out using the Pharmacia T7 polymerase sequencing kit according to manufacturer's instructions. Reaction products were separated on a 6% polyacrylamide denaturing gel (2.3.3ii).

iii) Decompression sequencing

Decompression reactions were carried out using the USB 7-deaza-dGTP sequencing kit according to manufacturer's instructions. Reaction products were separated on a 6% polyacrylamide denaturing gel (2.3.3ii).

2.3.14 Library Screening

i) Preparation of plating bacteria

10 ml LMM broth was inoculated with a single bacterial colony and grown O/N at 37°C in an orbital shaker.

ii) Library plating

5×10^4 recombinant phage were added to 200 µl plating bacteria and incubated at 37°C for 15 min. 9 ml LMM agarose at 50°C was transferred to the phage / bacteria mixture and overlaid onto 15 cm L agar plates. Bacteriophage plaques were formed by incubating at 37°C for 5-7 hours. When the plaques had grown to a sufficient size, the plates were cooled at 4°C for at least 1 hour before plaque lifts.

iii) Plaque lifts

Genescreen filters were placed onto library plates for 60 sec. (1st lift) or 3 min. (2nd lift) and transferred to 3MM Whatmann to air dry. Filters were then autoclaved for 2 min. and nucleic acid was bound to the filter by UV crosslinking. Second and third round screens were carried out on 10 cm plates and nitrocellulose filters were used. After lifts, nitrocellulose filters were transferred to 3MM Whatmann paper soaked in denaturing solution (0.5 M NaOH, 1.5 M NaCl) for 5 min. Filters were then transferred to 3MM Whatmann paper soaked in neutralising solution (0.5 M Tris-HCl pH 8.0, 1.5 M NaCl) for 5 min. and rinsed in 20 x SSC for 5 min. After air drying filters for 30 min. nucleic acid was bound to the filter by UV crosslinking.

2.3.15 Lambda Zapping

cDNA clones were isolated from third round positive plaques using the zapping protocol provided by Clontech.

2.3.16 Reverse Transcription

i) With AMV reverse transcriptase

2 (or 10) μg total RNA in a volume of 4 μl was denatured at 65°C for 10 min. before snap cooling on ice. The denatured RNA was added to a 16 (or 36) μl reverse transcription reaction mix containing 50 mM Tris-HCl pH 8.5, 6 mM MgCl_2 , 40 mM KCl, 1 mM DTT, 1.5 mM each dNTP, 40 U RNasin, 250 ng oligo d(T) primer (or 500 ng gene specific primer) and 15 U AMV reverse transcriptase and incubated for 1 hour at 41°C.

ii) With Superscript II reverse transcriptase

10 μg total RNA was reverse transcribed according to the manufacturer's instructions and diluted to 100 μl with MQ H_2O .

2.3.17 Reverse Transcription PCR (RT/PCR)

PCR mixes contained 5 μl cDNA template (2.3.16ii), 25 mM Tris-HCl pH 9.1, 16 mM $(\text{NH}_4)_2\text{SO}_4$, 250 μM each dNTP, 200 ng each primer, 2 U Taq polymerase and 0.025 U Pfu polymerase in a final volume of 43 μl . The reactions were overlaid with mineral oil and were "hot started" at 94°C for 1 min. prior to addition of 7 μl 25 mM MgCl_2 (3.5 mM final). Typical thermal cycling parameters were: (1) 94°C, 1 min. (hot start); (2) 94°C, 15 sec.; (3) 50°C, 15 sec.; (4) 72°C, 3 min.; (5) cycle (2-4) 40 times; 72°C, 5 min. PCR reactions were amplified in a MJ research PTC-100 thermal cycler.

2.3.18 Differential Display PCR

The reverse transcription reactions used for DDPCR were a modification of AMV reverse transcriptase conditions described in section 2.3.16. Reactions were primed with 250 ng 3'-hp (2.2.8) and the reaction mixes were preheated to 50°C before denatured RNA template was added. The 20 μl reverse transcription reactions were incubated for 1 hour at 50°C under mineral oil, to prevent evaporation. The reactions were then incubated at 94°C for 5 min. to inactivate the reverse transcriptase. cDNA solution was not diluted prior to PCR.

1 μl of the reverse transcription mix was used as template in 25 μl PCR reactions containing 1.5 mM MgCl_2 , 1 x Taq buffer, 20 μM each dNTP, 500 ng 3'-hp (2.2.8), 37.5 ng arbitrary 10mer (2.2.8), 5 μCi [α - ^{33}P] dATP (or 10 μCi [α - ^{35}S] dATP) and 1 U Taq polymerase. The DDPCR reactions were amplified in a MJ Research PTC-100 thermal cycler for 40 cycles of: (1) 94°C, 30 sec.; (2) 42°C, 30 sec.; (3) 72°C, 30 sec. followed by a final extension at 72°C for 5 min. 5 μl samples were removed after amplification, mixed with 3 μl 90% formamide / dye solution, boiled and run on 8M Urea / 6% polyacrylamide sequencing gels (2.3.3ii). After electrophoresis, gels were transferred to Whatmann 3MM paper, dried and taped down in an autoradiographic cassette. X-ray film was carefully aligned to the top left corner of the cassette and the gel was exposed to film for 1 to 3 days. Sequencing reactions or dye migration were used as markers of product sizes.

To reamplify DDPCR products of interest, 26G needle holes were punched in the X-ray film surrounding the individual band. The film was carefully realigned to the top left corner of the cassette and the position of the DDPCR product was marked through the holes with a fine pencil. A gel / Whatmann slice was excised from the marked region of the dried gel and the DDPCR product was eluted in 100-200 μ l of 0.5 M NH_4Ac , 10 mM MgAc_2 , 1 mM EDTA pH 8.0, 0.1% SDS at 50°C for 5-20 hours. 1 μ l of eluted DDPCR product was then used directly as a template in 100 μ l PCR reamplification reactions under the same conditions as the original DDPCR, omitting radioactive label. A 5 μ l sample of reamplified product was checked on an agarose gel (2.3.2). The remainder of the reamplified product was precipitated with 1/10th volume NaAc pH 5.2, 2 μ l glycogen and 2.5 volumes ethanol (-80°C for 30 min.). The DNA pellet was resuspended in 20 μ l and digested with *EcoRI* (2.3.1). Digested product was purified by non-denaturing acrylamide gel electrophoresis (2.3.3i), eluted (2.3.4) and ligated (2.3.7) to pBluescript KS *EcoRI*/T (2.3.6ii).

2.3.19 Rapid Amplification of cDNA Ends (RACE/PCR)

RACE PCR was carried out using the ampliFINDER kit (Clontech) according to manufacturer's instructions, with the following modifications: Reverse transcription conditions were the same as for AMV reverse transcriptase (2.3.16), except 1 μ g of $(\text{A})_n^+$ RNA was used in a reaction primed by 250 ng of a gene specific 3' primer. RNA was hydrolysed after 1 hour with 6 μ l 2 M NaOH at 65°C, 30 min., then neutralised with 2 μ l 6 M Acetic acid. 1st strand cDNA was purified by addition of 80 μ l NaI and Bresaclean procedure (2.3.4) and eluted in 50 μ l. cDNA was precipitated with 5 μ l 3 M NaAc pH 5.2, 2 μ l glycogen and 150 μ l ethanol at -80°C for 30 min., pelleted, dried *in vacuo* and resuspended in 10 μ l MQ H_2O . 4 pM Anchor oligonucleotide (2.2.8) was ligated to 2 μ l purified cDNA in a 10 μ l reaction containing T4 RNA ligase buffer (50 mM Tris-HCl pH 7.8, 10 mM MgCl_2 , 1 mM β -mercaptoethanol, 1 mM ATP) and 10 U T4 RNA ligase at 37°C for 6 hours under mineral oil. Ligation reactions were terminated by heating at 94°C for 2 min. and 1 μ l ligated cDNA template was used undiluted in RACE/PCR analysis using the 5' anchor primer (2.2.8) and a nested gene specific 3' primer. Amplification conditions were essentially identical to those described for RT/PCR in section 2.3.17.

2.3.20 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 stored at -80°C until use. Cell pellets were thoroughly resuspended in 2 ml ice cold TNM and lysed by vigorous pipetting 10 times. After incubation on ice for 5 min., the nuclei were pelleted by centrifugation (3 K for 5 min.) and the supernatant decanted and mixed thoroughly with 2 ml TUNES. 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 . After spinning at 10 K for 30 min. at 4°C , the RNA pellet was resuspended in 450 μl of sterile water, and transferred to an eppendorf tube for re-precipitation as above. The RNA was pelleted for 15 min., resuspended in an appropriate volume of water, and the concentration was determined by spectrophotometry. RNA samples were stored at -20°C .

2.3.21 Isolation of RNA from Tissue Samples

RNA from tissue samples was isolated using the acid guanidium thiocyanate method (Chomczynski and Sacchi 1987).

2.3.22 Isolation of $(A)_n^+$ RNA

Oligo-d(T) cellulose beads were prepared by soaking in 0.1 M NaOH and rinsed thoroughly in MQ H_2O before resuspending in pre-wash solution (0.5 M NaCl, 0.4 M Tris-HCl pH 7.5, 10 mM EDTA, 0.5% SDS). 500 μg cytoplasmic RNA was precipitated and resuspended in 100 μl MQ H_2O before denaturation at 80°C for 2 min. and snap cooling on ice. 100 μl resuspension solution (1 M NaCl, 0.4 M Tris-HCl pH 7.5, 20 mM EDTA, 1% SDS) was added to the RNA before transferring to 100 μl oligo-dT cellulose beads. After incubation at room temperature for 5 min. the beads were washed twice with 0.5 M NaCl / 10 mM EDTA and the $(A)_n^+$ RNA eluted twice in 150 μl MQ H_2O at 65°C for 5 min. RNA was precipitated by addition of 650 μl ethanol and 8 μl 4M NaCl, resuspended in MQ H_2O , and stored at -20°C .

2.3.23 RNA Gels for Northern Blot Analysis

1.3% agarose gels for northern blot analysis were prepared by dissolving 3.9 g agarose in 225 ml MQ water. Once the gel solution had cooled to 60°C , 25 ml 10 x MOPS and 15 ml 20% formaldehyde (freshly prepared by dissolving 4 g paraformaldehyde in 20 ml MQ H_2O) were added before pouring the gel.

$(A)_n^+$ RNA samples (approximately 10 μg) were prepared for electrophoresis as follows:

| | |
|---------------------|-------------------------------|
| 11.25 μl | RNA + MQ H_2O |
| 5 μl | 10 x 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. and snap cooled on ice. GLB was added before loading into wells. 5 μg *EcoRI* digested SPP-1 phage markers were also loaded. Northern gels were run at 6 V / cm gel length in 1 x MOPS until the bromophenol blue dye had reached the bottom of the gel.

2.3.24 Northern Blot Transfer

The lane containing the DNA markers 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 Nytran nylon membrane using capillary transfer. Two pieces of Whatmann 3MM paper were pre-wetted in 20 x SSC and placed over a platform so that the edges of the paper were submerged in 20 x SSC. The gel was placed wells facing down on the damp Whatmann paper (avoiding bubbles) and parafilm placed around the gel to avoid short circuiting. The nylon membrane was pre-wetted in 20 x SSC and then carefully placed over the gel such that no air bubbles were trapped between the filter and the gel. Two pieces of Whatmann paper, pre-wetted in 20 x SSC, were then placed on top of the membrane, followed by a 1 cm stack of dry Whatmann paper. A 5 cm stack of paper towels was then added and a glass plate placed on top of the paper towels. A 0.8 kg weight was placed on the glass plate and the gel was transferred for at least 24 hours. Following transfer, the RNA was cross-linked to the filter by exposure to 120 mJ of UV radiation in a Stratagene UV Stratalinker™ before pre-hybridisation.

2.3.25 Radioactive DNA Probes

DNA probes were prepared using the Gigaprime labelling kit with 50 μCi [α - ^{32}P] dATP in 20 μl reactions. The reactions were terminated by addition of 20 μl stop buffer (50% glycerol, 0.1 M EDTA pH 7.0, 0.02% bromophenol blue, 0.02% SDS) and 60 μl MQ H₂O and incubation at 65°C for 10 min. Unincorporated label was removed from the probe by centrifugation on a Sepharose CL-6B column (1.8 K, 3 min.).

2.3.26 Radioactive RNA Probes

Riboprobes were synthesised as described by Krieg and Melton (1987), using 60 μCi [α - ^{32}P] UTP. Unincorporated label was removed using a Sephadex G-50 column spun at 3.2 K for 5 min.

2.3.27 Hybridisation of Radioactive Probes to Nylon and Nitrocellulose Filters

i) DNA probes

Prehybridisation solution contained 1 M NaCl, 50% deionized formamide, 1% SDS, 10% PEG, 50 mM Tris-HCl pH 7.4, 100 $\mu\text{g}/\text{ml}$ sonicated salmon sperm DNA, 100 $\mu\text{g}/\text{ml}$ denatured tRNA and 5 x Denhardt's. Filters were prehybridised for a minimum of 4 hours at 42°C in 10 ml prehybridisation solution per filter. DNA probes (2.3.25) were boiled for 5 min., snap-cooled on ice before adding to filters. Filters were probed O/N at 42°C in a Hybaid hybridisation oven. Plaque lift filters (2.3.14iii) were hybridised in petri dishes (up to 20 per dish), with enough prehybridisation solution to cover all filters.

ii) RNA probes

Northern filters were prehybridised for a minimum of 4 hours at 65°C in 10 ml hybridisation solution containing 5 x SSC, 60% formamide, 5 x Denhardt's, 20 mM sodium phosphate pH 6.8, 1% SDS, 100 µg/ml sonicated salmon sperm DNA and 100 µg/ml denatured tRNA. RNA probes (2.3.26) were denatured by heating at 85°C for 2 min., snap-cooled on ice and added to Hybaid cylinders. Hybridisation reactions were carried out O/N at 65°C.

2.3.28 Ribonuclease Protection Analysis

RNase protection analysis was carried out using the RPAII kit according to the manufacturer's instructions. 120,000 c.p.m. of single stranded antisense *Psc1* riboprobe (2.2.7ii) and 12,000 c.p.m. of single stranded antisense *GAPDH* riboprobe (2.2.7iv) were added to each RNA sample. RNase digestion products were separated on a 7 M urea / 6% TBE polyacrylamide gel and visualised using autoradiography and phosphorimager analysis.

2.3.29 Digoxigenin (DIG) Labelled RNA Probes

DIG labelled riboprobes were prepared by transcription reactions containing 1-2 µg linearized plasmid, 40 mM Tris-HCl pH 7.5, 6 mM MgCl₂, 2 mM spermidine-HCl, 10 mM DTT, 2 U RNasin, 1 x DIG labelling mix and 40 U RNA polymerase in a total volume of 40 µl. Transcription reactions were incubated at 37°C for 2-3 hours before the template was removed by addition of 40 U RNase free DNase and incubation at 37°C for 15 min. Reactions were precipitated at -20°C (1 hour-O/N) after addition of 60 µl MQ H₂O, 8 µl 250 mM EDTA, 11 µl 4 M LiCl and 330 µl 100% ethanol. After centrifugation for 15 min., the probe was resuspended in 200 µl of RNase free MQ H₂O. A 5 µl sample of resuspended probe was used to assess riboprobe yield and quality by agarose gel electrophoresis (2.3.2).

2.3.30 *in situ* Hybridisation of Cell Monolayers

Cell monolayers were rinsed twice with PBS and fixed O/N in 4% PFA / PBS solution at 4°C. The following day, the cells were rinsed twice on ice with PBT (0.1% Tween 20 / PBS solution) and dehydrated on ice by the addition of 25% methanol / PBS for 5 min., 50% methanol / PBS for 5 min., 75% methanol / PBS for 5 min. and stored in 100% methanol at -20°C. Cells were rehydrated to PBS using the methanol on ice series, and rinsed three times at room temperature in PBT. Cells were then washed for 3 x 20 min. with RIPA buffer and postfixed in 4% PFA / 0.2% glutaraldehyde in PBT for 20 min. Plates were rinsed 3 x 5 min in PBT and washed for 5 min. in 1:1 hybridisation buffer:PBT (hybridisation buffer: 50% deionized formamide, 5 x SSC, 0.1% Tween-20, 50 µg/ml heparin). Cells were washed for a further 5 min. in hybridisation buffer, and prehybridised in hybridisation buffer containing 100 µg/ml denatured salmon sperm DNA and 100 µg/ml denatured yeast tRNA for 1-6 hours at 68°C. Prehybridisation and hybridisation steps were carried out in a sealed humidified box

containing paper towels soaked in 50% formamide. DIG-labelled riboprobes were denatured at 80°C for 10 min., snap cooled on ice, added to fresh hybridisation buffer containing denatured salmon sperm DNA and yeast tRNA, and hybridised to cells O/N at 68°C. Probes were diluted 1:200 for the hybridisation step. The following day, cell monolayers were washed in 50% formamide / 2 x SSC / 0.1% Tween 20 for 5 min. and 3 x 30 min. at 68°C. The cells were then washed three times at room temperature in TBST followed by incubation in 10% FCS / TBST for one hour at room temperature. Anti-digoxigenin Fab fragments-Alkaline Phosphatase conjugate (Boehringer Mannheim) was diluted 1:2000 in 1% FCS / TBST, preblocked for 1 hour at 4°C and incubated with cells O/N at 4°C. Next day, cells were washed for 3 x 5 min. and 6 x 30 min. in TBST and 3 x 10 min. in AP buffer. The cells were developed in AP buffer containing 0.45 mg/ml NBT and 0.18 mg/ml BCIP in the dark until purple staining appeared (1-8 hours). The staining reaction was terminated by rinsing several times with PBT / 1 mM EDTA. Cells were viewed and photographed under phase contrast and bright field optics using a Nikon Diaphot inverted microscope with Ectachrome 100 slide film.

2.3.31 Embryonic *in situ* Hybridisation

In situ hybridisation to preimplantation and postimplantation murine embryos was carried out with a modification of the whole mount digoxigenin method of Rosen and Beddington (1993). Embryos were treated with 2 ml solution volumes in 24 well trays using processing baskets (2.5.2).

Embryos were dissected and dehydrated as described in section 2.5.3 and 2.5.4. Dehydrated embryos in 100% methanol were rehydrated on ice with 75% methanol / PBT (0.1% Tween 20 / PBS solution) for 5 min., 50% methanol / PBT for 5 min., 25% methanol / PBT for 5 min. and rinsed three times with PBT for 5 min. at RT. Embryos were bleached with 6% H₂O₂ / PBT for 1 hour and rinsed three times with PBT for 5 min. at RT. Embryos were then washed with RIPA buffer for 3 x 20 min. at RT, rinsed twice in PBT and postfixed in 4% PFA / 0.2% glutaraldehyde in PBT for 20 min. at RT. Embryos were washed 3 x 5 min in PBT and washed for 5 min. in 1:1 hybridisation buffer:PBT (hybridisation buffer: 50% deionized formamide, 5 x SSC, 0.1% Tween-20, 50 µg/ml heparin). Embryos were washed for a further 5 min. in hybridisation buffer and prehybridised in hybridisation buffer containing 100 µg/ml denatured salmon sperm DNA and 100 µg/ml denatured yeast tRNA for 1-6 hours at 68°C. Prehybridisation and hybridisation steps were carried out in a sealed humidified box containing paper towels soaked in 50% formamide. DIG-labelled riboprobes were denatured at 80°C for 10 min., snap cooled on ice, added to fresh hybridisation buffer containing denatured salmon sperm DNA and yeast tRNA, and hybridised to embryos O/N at 68°C. Probes were diluted 1:200 for the hybridisation step. The following day, embryos were washed in solution I (50% formamide, 5 x SSC pH 4.5, 0.1% Tween 20, 0.5% CHAPS) for 5 min. at 68°C, 70% solution I / 30% 2 x SSC for 5 min. at 68°C, and 30%

solution I / 70% 2 x SSC for 5 min. at 68°C. Embryos were then washed twice in 2 x SSC, 0.1% CHAPS for 30 min. at 68°C and twice in 0.2 x SSC, 0.1% CHAPS for 30 min. at 68°C. Embryos were then rinsed twice in PBT at RT and washed three times in TBST for 5 min., followed by incubation in 10% FCS / TBST for one hour at room temperature. Preblocked anti-digoxigenin Fab fragments-Alkaline Phosphatase conjugate (Boehringer Mannheim) was prepared as described in section 2.3.33, and incubated with the embryos for 4-6 hours at 4°C with gentle rocking. Embryos were washed for 3 x 10 min. then O/N in TBST. The following day, embryos were washed 2 x 10 min in TBST and 3 x 10 min. in AP buffer. 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 for staining. Stainings were developed in AP buffer containing 0.45 mg/ml NBT and 0.18 mg/ml BCIP 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% PFA / 0.2% gluteraldehyde in PBT for 20 min. and embryos were stored in 0.4% PFA / PBT. Embryos were viewed and photographed with Ectachrome 160T slide film under phase contrast and bright field optics using a Nikon inverted microscope, or under bright field using a Leica MZ-8 dissecting microscope.

2.3.32 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. Homogenate was spun at 10 K, for 10 min. at 4°C (SS34 rotor in Sorvall RC-5 centrifuge) and the supernatant was discarded. The pellet was washed in ice-cold acetone and centrifuged as above. The pellet was ground to a powder on Whatmann 3MM paper, air dried and stored in an airtight tube at 4°C.

2.3.33 Antibody Preblocking

50-100 µl embryo powder (2.3.32) 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% FCS / TBST and 1 µl anti-digoxigenin Fab fragments-Alkaline Phosphatase conjugate (Boehringer Mannheim) was added *per in situ*. The antibody was blocked for 1-5 hours at 4°C with gentle rocking and centrifuged for 5 min. at 4°C. The supernatant was diluted with 1% FCS / TBST to give 2 ml per well of *in situ*.

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.

2.3.35 Phosphorimager Analysis, Autoradiograph Scanning and Image Manipulation

Gels and filters were exposed to Storage Phosphor Screens (Molecular Dynamics) and processed using a Molecular Dynamics PhosphorImager running the ImageQuant software package. Quantitation by volume intergration was carried out using ImageQuant. PhosphorImager files were manipulated using the Adobe Photoshop™ and MacDraw Pro programs, and printed using a Hewlett Packard LaserJet 5MP printer.

Photographic slides of embryonic *in situ* hybridisation analysis were scanned. Scanned images were manipulated using the Adobe Photoshop™ and PowerPoint programs, and converted to photographic images.

2.3.36 Sequencing Software and Database Searches

Sequencing reactions were read using a digitiser and MacDNASIS software (Hitachi). Contiguous sequence alignments were performed using MacDNASIS.

Sequence 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.

2.4 Tissue Culture Methods

2.4.1 Maintenance of ES Cells

ES cells were routinely maintained as described by Smith (1991), in ES medium at 37°C in 10% CO₂. Cells were processed for *in situ* hybridisation (2.3.30) or harvested as follows: cells were washed twice in PBS, incubated with trypsin (1 ml) at 37°C and transferred into 4 ml complete ES cell medium. The cells were spun at 1.2 K for 5 min., gently resuspended in 10 ml complete ES medium (2.2.11) and re-seeded at clonal density (20-40 fold dilution). ES cells were re-seeded every 3-4 days.

2.4.2 Differentiation of ES Cells

i) Generation of X cells

ES cells were seeded at clonal density in X cell medium (with or without LIF). X cell colonies were detected within 24 hours of induction. X cells were maintained in X cell medium (2.2.11), and were passaged and processed (2.3.30) for *in situ* hybridisation, or harvested (2.4.1) as described for ES cells. The nomenclature used for the ES to X cell conversion was based on the passage number of the culture (0-3) and the presence (+) or absence (-) of supplementing LIF within the X cell medium. For example, the induction of X

cells in the presence of LIF and subsequent culture for two passages results in a culture termed X+2 cells.

Reversion of X cells to ES cell like colonies: X cells were seeded in complete ES cell medium (2.2.11) and passaged every 3-4 days.

ii) Spontaneous Induction

ES cells were seeded at 1000 cells/cm² in incomplete ES medium (2.2.11). The cells were grown for 8-9 days, with daily changes of medium after the initial 72 hours, and were processed (2.3.30) for *in situ* hybridisation, or harvested (2.4.1) as described for ES cells.

iii) Retinoic Acid Induction

ES cells were seeded at 1 x 10⁴ cells/cm² in incomplete ES medium (2.2.11) containing 10U/ml recombinant LIF (Amrad) and allowed to adhere O/N. The medium was replaced the following day with identical medium containing 2 µM retinoic acid, and the cells cultured for a further 72 hours with daily changes of (RA containing) medium. The cells were cultured for a further 2-4 days in incomplete ES cell medium (without LIF) before being processed (2.3.30) for *in situ* hybridisation, or harvested (2.4.1) as described for ES cells.

iv) 3-Methoxybenzamide Induction

ES cells were seeded at 1000 cells/cm² in incomplete ES medium (2.2.11) containing 100 U/ml recombinant LIF (Amrad) and allowed to adhere O/N. The following day, the medium was replaced with identical medium containing 5 mM 3-methoxybenzamide (MBA), and the cells induced for 48 hours. The cells were then processed (2.3.30) for *in situ* hybridisation, or harvested (2.4.1) as described for ES cells.

v) Dimethylsulfoxide Induction

ES cells were seeded at 1000 cells/cm² in incomplete ES medium (2.2.11) containing 100 U/ml recombinant LIF (Amrad) and allowed to adhere O/N. The culture medium was replaced the following day with identical medium containing 1.5% DMSO. Induction was maintained for 72 hours with daily changes of DMSO containing medium. Following induction, cells were cultured in medium without LIF for a further 3-4 days before being processed (2.3.30) for *in situ* hybridisation, or harvested (2.4.1) as described for ES cells.

2.5 Embryological Methods

2.5.1 Mouse Strains and Matings

Embryos to be used for *in situ* hybridisation were generally produced from Balb/C x C57Bl/6J matings, or occasionally were F1's or F2's from CBA x C57Bl/6J matings. Preimplantation embryos were generated by natural matings or standard superovulation conditions (Hogan *et al.* 1994), postimplantation embryos were derived from natural matings. Mice were maintained on a 14 hour light / 10 hour dark cycle and supplied with food and water *ad libitum*.

2.5.2 Embryo Processing Baskets

For complex *in situ* hybridisation procedures, preimplantation and postimplantation embryos were processed through solutions in siliconised plastic processing baskets. Baskets were made (in a fume hood) by firstly removing the bottom of a screw capped eppendorf tube with a heated scalpel blade. The top of the tube (without lid) 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. Processing baskets were suitable for both 4 well and 24 well tissue culture dishes.

2.5.3 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. As slight variation in developmental progression occurred within litters and between litters, individual embryos were reclassified according to idealised developmental time points such as in Kaufman 1992, Hogan *et al.* 1994. Isolation and dissection of preimplantation and postimplantation embryos was carried out according to Hogan *et al.* 1994 with the following modifications:

i) Preimplantation embryos

Uteri were removed and embryos were flushed from oviduct or uteri with dissection medium (20% FCS / ES DMEM, 10 mM HEPES pH 7.4). The zona pellucida was not removed.

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

Uteri were removed and individual pregnancies isolated in dissection medium (20% FCS / ES DMEM, 10 mM HEPES pH 7.4). The mesometrial side of the pregnancy and early decidua was prised open and embryos were "shelled" out of the decidua with fine forceps.

2.5.4 Embryo Fixation and Dehydration

Dissected preimplantation and postimplantation embryos were handled using siliconised glassware. Embryos were rinsed once in PBS and once in 4% PFA / PBS. Embryos were transferred to siliconised processing baskets and fixed in 4% PFA / PBS for 4-16 hours at 4°C. Embryos were rinsed twice on ice with PBT (0.1% Tween 20 / PBS solution) and dehydrated on ice by 25% methanol / PBS for 5 min., 50% methanol / PBS for 5 min., 75% methanol / PBS for 5 min. and 100% methanol. Embryos were stored at -20°C in processing baskets, in 2 ml 100% methanol in sealed 10 ml polypropylene tubes.

2.5.5 Animal Manipulations

All procedures involving animals were carried out with the approval of the University of Adelaide Animal Ethics Committee.

CHAPTER 3

*A Method for the Identification of Differentially
Expressed Genes During Pluripotent Cell Progression
in vitro*

CHAPTER 3

A METHOD FOR THE IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES DURING PLURIPOTENT CELL PROGRESSION *IN VITRO*

3.1 Introduction

The developmental progression of pluripotent cells during early embryogenesis is fundamental for the processes that establish the embryonic body plan and tissue primordia. While the developmental potential of these cell populations has been studied by transplantation and fate mapping experiments (1.5, 1.5.1), a paucity of specific markers has severely limited molecular dissection of their development (1.5.2). Detailed molecular analysis of these developmental stages requires markers that are specifically expressed in subpopulations of the pluripotent cell pool. The existence of stem cell subpopulations and corresponding specific gene expression is indicated by multiple changes to pluripotent cell morphology and the specific timing of embryonic events (1.5.1). The limited array of markers can be attributed to the small size and inaccessibility of various stages of early murine embryogenesis, coupled with the impracticality of mouse genetic screens for early developmental defects. The generation of pluripotent cell markers could therefore require the development of alternate isolation strategies.

3.1.1 *In vitro* Models of the Inner Cell Mass and Primitive Ectoderm

ES cells are induced to form an homogeneous monolayer of flattened cells, termed X cells (1.6.7), when cultured in the presence of medium conditioned by a transformed human hepatic cell line (MedII cells; Lake 1996). Morphologically indistinguishable X cells are generated in the presence or absence (+ or -) of supplementary LIF. In this analysis, X cells were grown for up to three passages (X+0, 1, 2, or 3; X-0, 1, 2, or 3).

Gene expression and differentiation potential analysis indicates that ES and X cells are distinct pluripotent cell types (1.6.8). In particular, X+3 and X-3 cells, collectively termed Xp3 (X cells, passage 3), have gene expression characteristics akin to the primitive ectoderm.

The conversion of ES cells to Xp3 cells therefore represents a model of the ICM to primitive ectoderm transition *in vitro* (1.6.9, Lake 1996). Additionally, several distinct timed subpopulations of pluripotent cells are predicted to exist during the ES to X cell transition (1.6.9), which could represent specific “windows” of pluripotent cell development during the ICM to primitive ectoderm transition. Therefore, the identification of differentially expressed genes during the ES to X cell transition has the potential to generate a series of molecular markers for distinct pluripotent cell subpopulations during early embryogenesis.

3.1.2 Differential Display Polymerase Chain Reaction and Gene Class Enrichment

Differential display PCR (DDPCR) is potentially a powerful method to identify differentially expressed transcripts in the ES to X cell transition, as candidate markers for pluripotent cell subpopulations *in vivo*. DDPCR can be used to exploit inherent strengths of the ES model of early development, by comparing simultaneously the expression of numerous genes in multiple pluripotent and terminally differentiated cell types (1.7.6).

Modification to DDPCR to bias the amplification of particular gene classes could have significant benefits. The identification of certain gene types, for example transcription factors, cell signalling molecules or growth factors, is often advantageous when investigating developmental processes, as these types of factors are commonly key regulatory signals for developmental events. Correspondingly, the expression domains of these regulatory gene classes are often spatially or temporally restricted, thus signifying their capacity to serve as cell type-specific markers. Homeodomain DNA binding proteins represent an example of such a gene family.

3.1.3 Homeobox Genes and Expression in Pluripotent Cell Subpopulations

The homeodomain is a 60 amino acid DNA binding motif that has been identified in over 65 mouse genes (reviewed in Burglin 1994). Expression and functional analysis during murine embryogenesis has demonstrated that homeobox genes are often critical for pattern formation and specification of regional identity. However, besides *Oct-4* (1.5.2) and *Oct-6*

(Suzuki *et al.* 1990), relatively little is known about homeobox gene expression or control functions in the pluripotent cells of early embryogenesis.

Murine homeobox genes are grouped into approximately twenty classes (reviewed in Burglin 1994). The homeodomain is defined by the conservation of residues at approximately twenty one invariant positions. In particular, a short region that is critical to DNA binding is highly conserved within the third α -helix, near the carboxyl termini of all homeodomains (2.2.8). In addition to these invariant positions, sequence conservation throughout the homeodomain defines an alternate consensus for each distinct homeobox class. Therefore, a degenerate 3' PCR primer that recognises the majority of homeobox gene transcripts can be designed to the common region near the carboxyl terminal of the homeodomain (Thomas 1995), termed 3'-hp (described in 2.2.8). Degenerate 5' PCR primers can be designed to regions of class specific consensus and are capable of recognising the majority of members within a homeodomain class (Thomas 1995).

A RT/PCR screen for homeobox genes of the *Antennapedia* and *Engrailed* classes, using 3'-hp in combination with class specific 5' primers, identified thirteen different transcripts expressed in ES cell cultures, including three novel homeobox sequences (Thomas 1995). Expression analysis during ES cell differentiation indicated that some of these genes, for example *Gbx-2*, could provide molecular markers specific for pluripotent cell subpopulations (Chapman 1994, Thomas 1995). However, this approach was limited by the necessity to screen individually different homeobox classes, and the analysis of homeobox genes from additional classes was not undertaken. Furthermore, extensive expression analysis of isolated homeobox genes was required to identify those restricted to pluripotent cell subpopulations.

The development of a modified DDPCR regime that is capable of enriching for homeobox gene transcripts could streamline the identification of homeobox genes expressed in pluripotent cell subpopulations. Modification of the differential display oligonucleotides could enable enrichment for homeobox gene transcript amplification. An example of this may be to replace the poly(A) anchored 3' primer with 3'-hp (2.2.8), a degenerate 3' primer that should recognise the majority of homeobox sequences. DDPCR with 3'-hp and arbitrary 5' primers could enable the amplification of homeobox gene sequences, by specific annealing of

3'-hp to the homeodomain consensus and random priming of 5' oligonucleotides. Advantages of this approach are firstly, the potential to amplify homeobox transcripts irrespective of their particular class, and secondly, the opportunity to restrict analysis to homeobox gene transcripts demonstrated to be expressed in specific pluripotent cell subpopulations.

3.1.4 Chapter Summary

This chapter describes the development of methods for the identification and isolation of markers for pluripotent cell populations present during the ES to X cell transition. Modifications to DDPCR conditions were made in an attempt to develop a method capable of enriching for homeobox gene transcripts. Using this technique, a screen for pluripotent cell markers that were restricted to ES cell populations was carried out. Two novel, non-homeobox, markers that were expressed by ES cells and early X cell types were identified and isolated, A03/360 and B04/400. The analysis of A03/360 and B04/400 expression by DDPCR and northern blot indicated that at least three distinct pluripotent cell subpopulations could be identified during the ES to X cell transition. The results detailed in this chapter therefore describe the combination of two experimental systems into a powerful method for the identification of markers for pluripotent cell subpopulations during early embryogenesis. Lineage specific differentiation of ES cells *in vitro* was used to provide a source of stem cell subpopulations and DDPCR was applied to identify and isolate transcripts within these cultures that defined specific pluripotent cell subtypes.

3.2 Modification and Trial of Differential Display PCR Conditions

3.2.1 Differential Display PCR Modifications

All modifications to differential display PCR conditions using the homeodomain 3' primer, 3'-hp (2.2.8), were carried out on ES cell RNA, or involved the comparison of ES cell gene expression with that of differentiated ES cell types. The primary objectives of these experiments were to establish differential display profiles with the 3'-hp primer, to trial reamplification and cloning procedures, and to assess the potential for enrichment of

homeobox gene transcripts. Total cytoplasmic RNA preparations were used for all DDPCR analysis (2.3.20).

3.2.1i Amplification Profiles With 3'-hp

The capacity of 3'-hp to generate differing amplification profiles in combination with varying 5' arbitrary 10 mers (2.2.8), was assessed by DDPCR amplification (2.3.18) of ES cell cDNA (Fig. 3.1). MBL-5 ES cell RNA was reverse transcribed with oligo d(T) primer at 42°C (2.3.16i) and first strand cDNA was amplified with 3'-hp in combination with the arbitrary 5' 10 mers OPA-01, OPA-02, OPA-03, OPA-04 or OPA-05.

3-10 major fragments resulted from DDPCR with each arbitrary 10 mer, with discernible products ranging in size from approximately 20 bp to 250 bp. The amplification profile was distinct for each 10 mer, although two common bands (arrowheads) were present in all reactions and were likely to be generated by PCR amplification with 3'-hp alone. All reactions exhibited a high background level of amplification, with minor bands at every nucleotide position on the gel and a smear obscuring fragments larger than approximately 250 bp. The background amplification was thought likely to complicate specific reamplification and cloning procedures required for successful DDPCR. This was also indicated by the co-amplification of fragments derived from 18S ribosomal RNA during cloning and sequencing assessment of OPA-05/3'-hp reactions (data not shown). This indicated that the differential display reverse transcription or PCR conditions required modification to reduce the level of background amplification.

3.2.1ii Assessment of Template Preparation Methods

In an attempt to reduce the high level of background amplification in 3'-hp DDPCR, alternative reverse transcription and PCR template preparation strategies were assessed (Fig. 3.2). Reverse transcription of MBL-5 ES cell cytoplasmic RNA was primed with either oligo d(T), or with 3'-hp at 42°C for 1 hour (2.3.16i). The cDNA was then used directly in DDPCR with OPA-03/3'-hp (2.3.18), or RNA was removed from reverse transcription reactions by NaOH hydrolysis and the cDNA purified by Bresaclean (2.3.4) prior to DDPCR analysis.

Figure 3.1 Analysis of 3'-hp Differential Display PCR Primer Combinations and Profiles.

10 μ g MBL-5 ES cell cytoplasmic RNA was reverse transcribed (2.3.16i) with oligo d(T) primer (2.2.1) for 1 hour at 42°C. 1/40th of undiluted first strand cDNA was used as the template for DDPCR (2.3.18) using 3'-hp in combination with either OPA-01, OPA-02, OPA-03, OPA-04, or OPA-05 (2.2.8). [α -³⁵S] dATP was incorporated. Products common to all reactions were likely to be generated by amplification with 3'-hp alone (arrowheads). Products larger than approximately 250 bp were obscured by a high background (not shown). Approximate sizes are indicated.

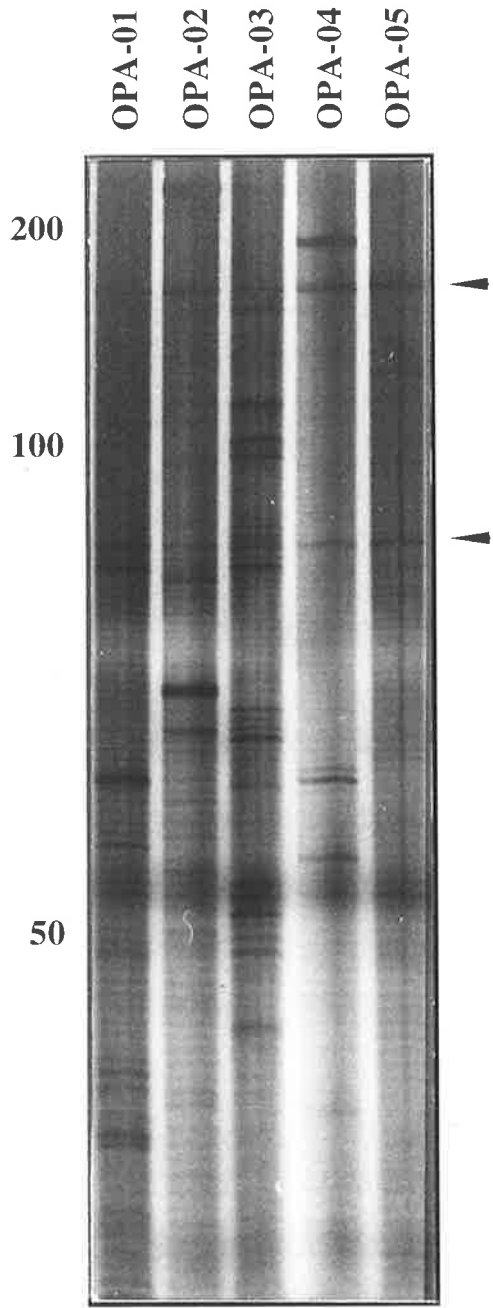


Figure 3.2 Reduction in 3'-hp DDPCR Background Amplification.

3'-hp DDPCR template preparation methods were altered to investigate and reduce the source of background amplification products. DDPCR (2.3.18) with OPA-03/3'-hp (2.2.8), incorporating [α - ^{33}P] dATP, was performed on the following cDNA templates. 3'-hp DDPCR reactions were run on the same gel, intervening lanes are not shown. Approximate sizes are indicated.

(1) Unpurified oligo d(T) primed template:

10 μg MBL-5 ES cell cytoplasmic RNA was reverse transcribed (2.3.16i) with oligo d(T) primer (2.2.1) for 1 hour at 42°C in a 40 μl reaction volume. 1/40th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR.

(2) Purified oligo d(T) primed template:

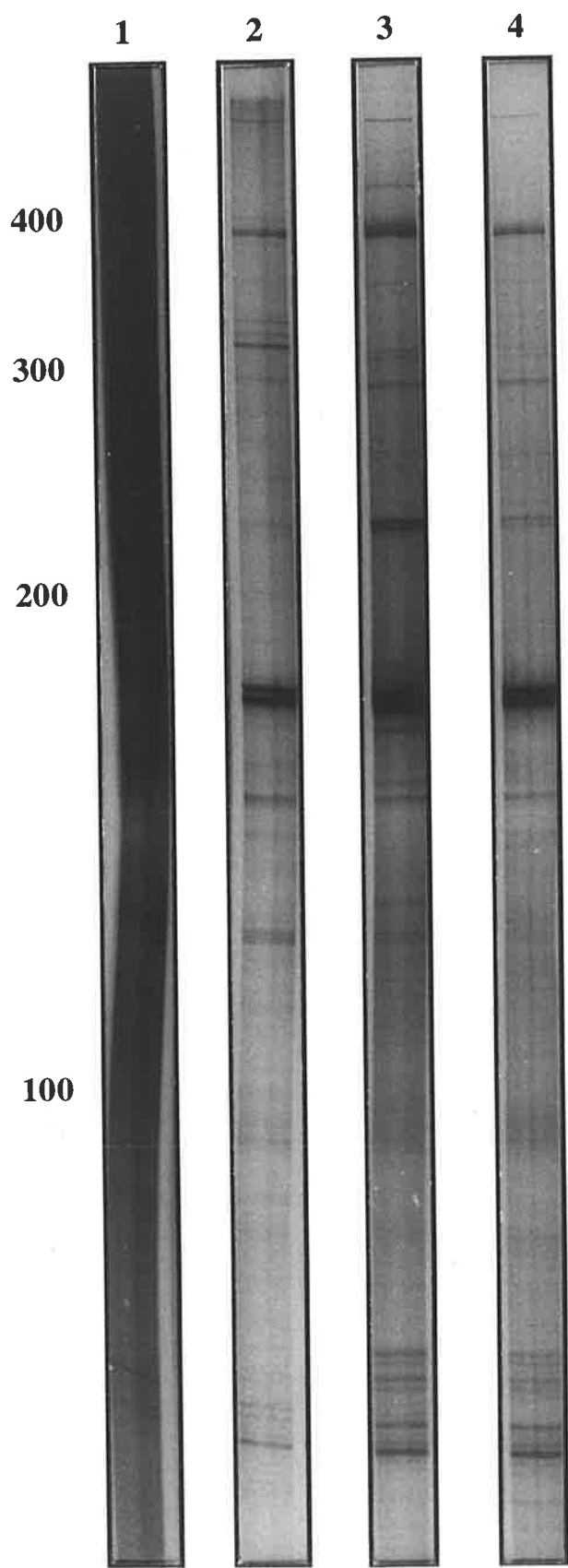
30 μl of unpurified oligo d(T) primed template (1) was hydrolysed with 2 μl 6M NaOH (65°C, 30 min.), neutralised with 2 μl 6M Acetic acid, purified by Bresaclean (2.3.4) and eluted in 30 μl . 1/30th purified template was used in 3'-hp DDPCR.

(3) Unpurified 3'-hp primed template:

10 μg MBL-5 ES cell cytoplasmic RNA was reverse transcribed (2.3.16i) with 3'-hp (2.2.8) for 1 hour at 42°C in a 40 μl reaction volume. 1/40th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR.

(4) Purified 3'-hp primed template:

30 μl of unpurified 3'-hp primed template (3) was hydrolysed with 2 μl 6M NaOH (65°C, 30 min.), neutralised with 2 μl 6M Acetic acid, purified by Bresaclean (2.3.4) and eluted in 30 μl . 1/30th purified template was used in 3'-hp DDPCR.



RNA hydrolysis and cDNA purification of oligo d(T) primed reverse transcription (lane 2) markedly reduced background amplification products compared to unpurified oligo d(T) primed template (lane 1). The visualisation of displayed products was enhanced, particularly for fragments larger than approximately 200 bp (compare lanes 1 and 2). Reverse transcription with 3'-hp led to a similar reduction in background, with (lane 4) or without (lane 3) template purification prior to DDPCR. The three modifications that reduced amplification background (lanes 2, 3, 4) generated DDPCR profiles that were similar but not identical, with differences relating to the presence or absence of particular bands, or variation in banding intensity. These results indicated that high levels of background amplification in 3'-hp DDPCR could be reduced by either removing RNA and reverse transcription reaction components from oligo d(T) primed cDNA template, or by using unpurified or purified 3'-hp primed cDNA as a template.

3.2.1iii 3'-hp Display and Detection of Differential Gene Expression

The application of 3'-hp DDPCR to the detection of differential transcript expression between multiple cell types was demonstrated by comparison of DDPCR reactions (2.3.18) carried out on mRNA from MBL-5 ES cells and terminally differentiated MBL-5 ES cell types. ES cells (2.4.1) were differentiated by the withdrawal of LIF (SP, 2.4.2ii), or by chemical induction with retinoic acid (RA, 2.4.2iii), 3-methoxybenzamide (MBA, 2.4.2iv) or dimethylsulfoxide (DMSO, 2.4.2v). Cytoplasmic RNA from these cell types was reverse transcribed with 3'-hp at 52°C for 1 hour. Total RNA from a 10.5 d.p.c. murine embryo (2.3.21) was included as a “non-pluripotent” control sample. Unpurified cDNAs were amplified with OPA-01/3'-hp or OPA-02/3'-hp (Fig. 3.3A, B respectively). Amplified products were labelled according to the arbitrary 5' primer used (OPERON primer kit “A” and oligo number, 2.2.8) and approximate fragment size. Due to the low background in these reactions, reverse transcription with 3'-hp at 52°C was used for all subsequent DDPCR analysis.

The OPA-01/3'-hp banding profile (Fig. 3.3A) was evenly distributed and generated approximately 20 distinct major amplification products and 50-80 minor products, as defined by low level amplification or blurred banding. Most products appeared as single bands,

although some products were doublets, a common feature of subsequent gels. This was potentially a consequence of slightly different migration rates of individual DNA strands in denaturing gels (Sambrook *et al.* 1989) and could be compounded by variable promiscuous adenosine addition by Taq DNA polymerase terminal transferase activity (Clark 1988). The banding profile of individual transcripts could be grouped into distinct classes of expression patterns. The majority of transcripts belonged to class I, and were expressed by all the ES cell derived cell types and the 10.5 d.p.c. embryonic control. These transcripts were therefore likely to represent housekeeping genes. Control 10.5 d.p.c. embryonic RNA exhibited many distinct additional bands, or class II transcripts, for example A01/360. This was likely to highlight specific gene expression in embryonic cell lineages not present during the ES cell differentiation regimes. Class III transcripts were specifically upregulated in individual ES differentiation regimes. For example, A01/400 was specifically upregulated during SP differentiation, and A01/600 was upregulated during RA induced differentiation. These transcripts may potentially identify differentiated cell types that are restricted to specific ES cell differentiation regimes.

The OPA-02/3'-hp expression profile (Fig. 3.3B) demonstrated similar features of major and minor banding and product distribution to the OPA-01/3'-hp reactions (Fig. 3.3A). Class I transcripts were also detected, for example A02/400. Approximately five class II transcripts were identified, for example A02/250 and A02/335. No class III transcripts were displayed with the OPA-02/3'-hp primer combination, but two additional transcript classes were identified. Class IV transcripts were expressed in ES cells and differentiated ES cell types, but not in 10.5 d.p.c. embryonic cell types, including A02/240, A02/340, A02/350 and A02/420. This signifies that cell types generated by these ES cell differentiation regimes may not be represented at 10.5 d.p.c. of development. A Class V transcript was expressed in multiple differentiated ES cell types. Two closely migrating doublets, collectively termed A02/290, were not detected in undifferentiated ES cells, but were upregulated during cellular differentiation and in the 10.5 d.p.c. control. More specifically, A02/290 was expressed at a higher level in SP and RA cells than in MBA and DMSO cell types. This verifies that SP and RA induced ES cells may be more closely related to each other than to MBA and DMSO

Figure 3.3 3'-hp DDPCR of Embryonic Stem Cells and Differentiated Cell Types.

2 µg cytoplasmic RNA was reverse transcribed (2.3.18) with 3'-hp for 1 hour at 52°C in a 20 µl volume. 1/20th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR (2.3.18) with OPA-01 (2.2.8). [α -³³P] dATP was incorporated. DDPCR products were labelled by the arbitrary primer and approximate size. ES, undifferentiated MBL-5 ES cells (2.4.1); SP, spontaneous (-LIF) induction (2.4.2ii); RA, retinoic acid induction (2.4.2iii); MBA, 3-methoxybenzamide induction (2.4.2iv); DMSO, dimethylsulfoxide induction (2.4.2v); 10.5, total RNA from a 10.5 d.p.c. embryo (2.3.21). 3'-hp DDPCR reactions were run on the same gel, intervening lanes between ES and SP reactions are not shown. Approximate sizes are indicated.

(A) OPA-01/3'-hp DDPCR of embryonic stem cells and differentiated cell types.

Examples of differential or restricted transcripts:

| | | |
|------------|---------|--------|
| Class II: | A01/360 | (10.5) |
| Class III: | A01/400 | (SP) |
| | A01/600 | (RA) |

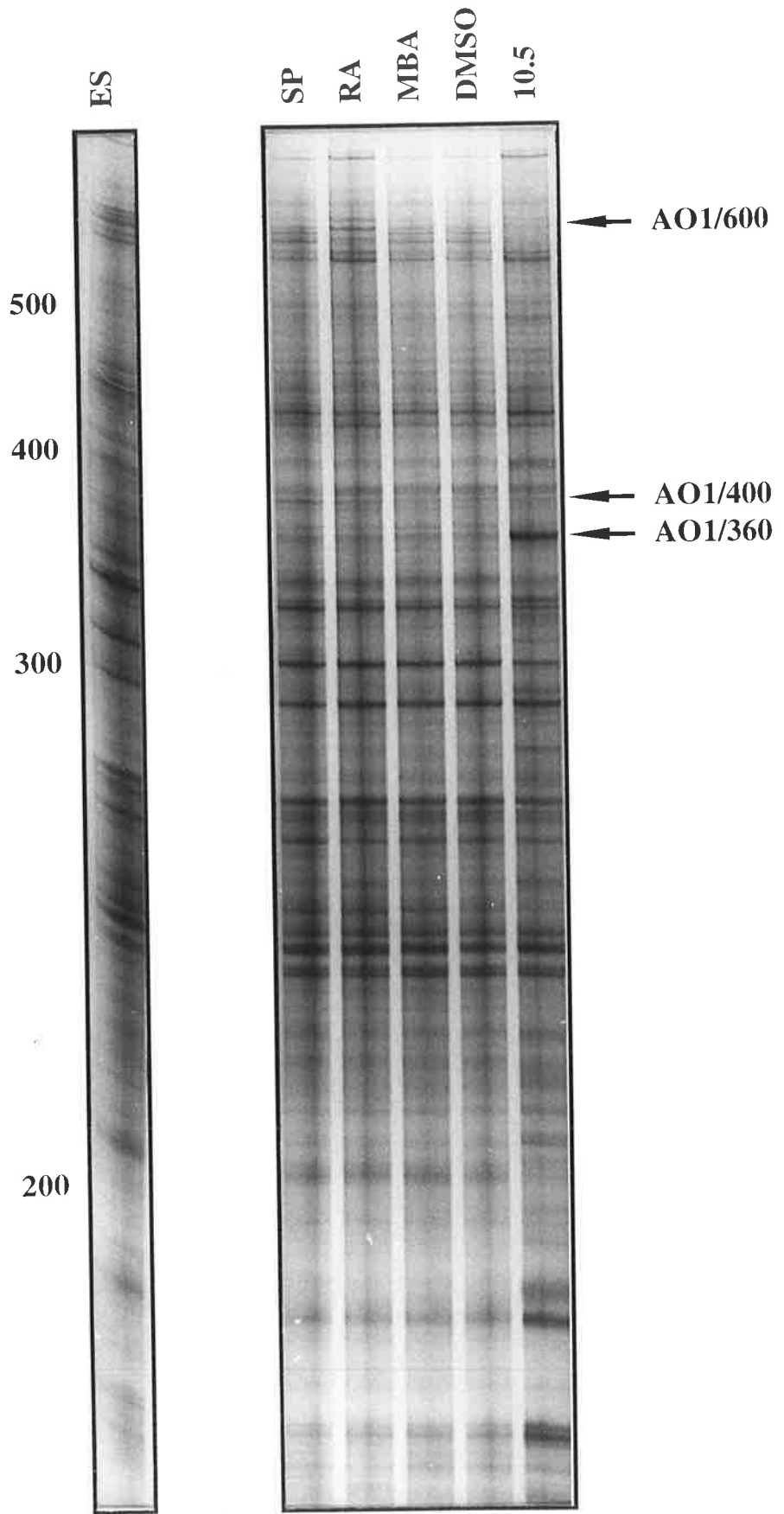


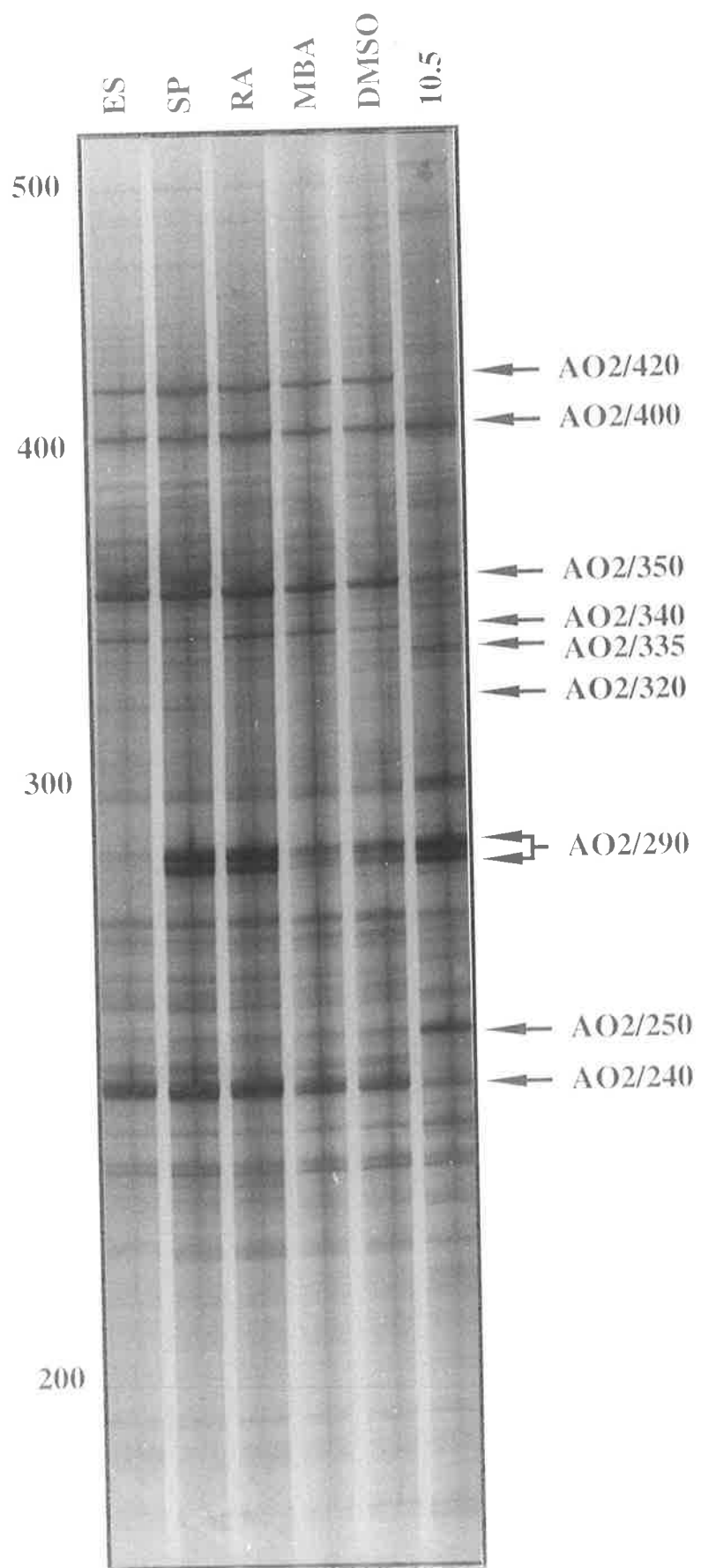
Figure 3.3 3'-hp DDPCR of Embryonic Stem Cells and Differentiated Cell Types.

2 µg cytoplasmic RNA was reverse transcribed (2.3.18) with 3'-hp for 1 hour at 52°C in a 20 µl volume. 1/20th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR (2.3.18) with OPA-02 (2.2.8). [α -³³P] dATP was incorporated. DDPCR products were labelled by the arbitrary primer and approximate size. ES, undifferentiated MBL-5 ES cells (2.4.1); SP, spontaneous (-LIF) induction (2.4.2ii); RA, retinoic acid induction (2.4.2iii); MBA, 3-methoxybenzamide induction (2.4.2iv); DMSO, dimethylsulfoxide induction (2.4.2v); 10.5, total RNA from a 10.5 d.p.c. embryo (2.3.21). Approximate sizes are indicated.

(B) OPA-02/3'-hp DDPCR of embryonic stem cells and differentiated cell types.

Examples of common, differential or restricted transcripts:

| | | |
|-----------|------------------------------------|----------------------------|
| Class I: | A02/400 | (common) |
| Class II: | A02/250, A02/335 | (10.5) |
| Class IV: | A02/240, A02/340, A02/350, A02/420 | (ES, SP, RA, MBA, DMSO) |
| Class V: | A02/290 | (SP, RA, 10.5 > MBA, DMSO) |



induced ES cells, or that they may share a common differentiation pathway as proposed by Thomas (1995).

3'-hp DDPCR enabled the detection of differential gene expression during ES cell differentiation and at 10.5 d.p.c. of embryogenesis, identifying five distinct class of transcript expression profiles. Application of 3'-hp DDPCR to the ES to X cell transition may therefore enable the identification of transcripts with a range of different expression profiles. This approach is likely to define markers of specific pluripotent cell subpopulations.

3.2.2 3'-hp DDPCR Reamplification and Cloning Procedures

Before 3'-hp DDPCR could be applied to the generation of pluripotent cell markers, it was necessary to establish an effective reamplification and cloning strategy to isolate prospective markers from DDPCR gels. Seven bands from the OPA-02/3'-hp expression profile (Fig. 3.3B), A02/240, A02/250, A02/350, A02/400, A02/420 and separate top (A02/290a) and lower (A02/290b) doublets from the A02/290 product, were selected to trial marker isolation procedures. Band positions were precisely determined from autoradiographic exposures and fragments were excised as dried gel slices. Gel slice eluate was reamplified (2.3.18) and the products visualised by agarose gel electrophoresis (2.3.2). Original and reamplified products were expected to contain single adenosine additions at the 3' end of each DNA strand, resulting from the terminal transferase activity of Taq DNA polymerase (Clark 1988). As 3'-hp contained an engineered *Eco*RI restriction site, fragments were to be ligated directionally via *Eco*RI at the 3' end and a single adenosine overhang at the 5' end, to an *Eco*RI/T overhang prepared pBluescript KS vector (2.3.6ii).

Reamplified products were digested with *Eco*RI and purified by polyacrylamide gel electrophoresis (2.3.3i) as shown in figure 3.4. The estimated size of some products (3.2.1iii) were slightly inaccurate, but original names were retained for simplicity. Eluted fragments appeared as a distinct major band after reamplification and digestion, with the exception of A02/400. This demonstrated that fragments of specific size could be purified and reamplified from dried DDPCR gels. Furthermore, internal *Eco*RI sites in the A02/290a, A02/290b and A02/420 fragments generated a single major digestion product for each reamplification, with small digestion products running off the gel or not sharply resolved (not

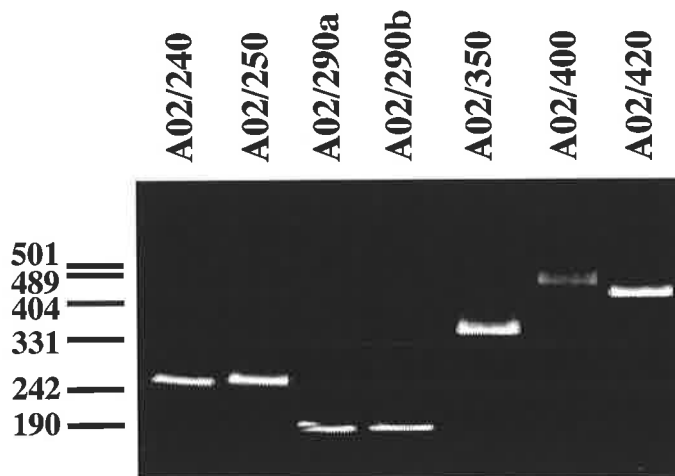


Figure 3.4 Reamplification of DDPCR Products From Dried Acrylamide Gels.

Products from OPA-02/3'-hp DDPCR of embryonic stem cells and differentiated cell types (Fig. 3.3B) were excised, eluted and reamplified (2.3.18). Reamplified A02/240, A02/250, A02/350, A02/400, A02/420, A02/290a and A02/290b products were digested with *EcoRI* and purified by electrophoresis on a non-denaturing 5% polyacrylamide gel (2.3.3i). A02/400 did not reamplify effectively. A02/290a, A02/290b and A02/420 contained an internal *EcoRI* site, with small digestion products migrating at the bottom or off the gel (not shown). The sizes of marker fragments are indicated.

shown). This indicated that for these reactions, the majority of reamplified product was likely to be comprised of a single sequence, as background amplification products would be expected to generate digestion products of a different size.

These results indicated that elution and reamplification procedures had the potential to purify differentially expressed transcripts from 3'-hp DDPCR gels. This may be more effective than reamplification of DDPCR fragments (1.7.6), with a reduction in the number of transcripts displayed per reaction possibly contributing to lower levels of background products during reamplification.

3.2.3 Sequence Analysis of 3'-hp DDPCR Products

To assess the types of transcripts amplified by 3'-hp DDPCR, *EcoRI* digested A02/240, A02/250 and A02/350 fragments were directionally cloned into *EcoRI*/T overhang prepared pBluescript KS (2.3.6ii). Undigested A02/290a, A02/290b and A02/420 fragments were cloned into T overhang prepared pBluescript (2.3.6ii). For rapid characterisation, a single clone for each ligated product was sequenced in one direction. Directionally cloned fragments were sequenced from the 3' end. This would enable the identification of homeobox gene transcripts, by detection of conserved residues at invariant positions of the homeodomain, adjacent to the binding site of 3'-hp (2.2.8). Sequences and conceptual translations in all six reading frames were compared to the GenBank and EMBL nucleotide and translated nucleotide data bases (2.3.36). As these fragments were not to be characterised further and the primary objective was to obtain rapid information on the sequence types present, detailed information regarding novel sequences, sequence comparisons, fragment size or position within original transcript were not obtained and are not included in this analysis.

A02/240 was approximately 90% similar to a coding region segment of the human electron transfer flavoprotein α subunit (Finocchiaro *et al.* 1988; Accession no. HUMETF). A02/350 was a fragment of the previously identified murine preimplantation gene (Temeles *et al.* 1994; Accession no. UO1140). A02/250 and A02/420 were not homeodomain sequences and did not contain significant similarity to any nucleotide or open reading frames within the data bases. No differences between the sequences of A02/290a and A02/290b were detected

by this preliminary analysis, indicating they represented strand migration differences of the same amplification product (3.2.1iii). A02/290 encoded residues 152 to 274 of the homeobox gene *Hox-d8*, a member of the clustered Hox genes involved in embryonic pattern formation (Izpisua-Belmonte *et al.* 1990, Accession no. A43562). The correct orientation of primer sequences in the cloned A02/290 product indicated that the *Hox-d8* transcript was amplified through specific binding of 3'-hp to the homeodomain sequence, in combination with the arbitrary 5' primer OPA-02 (data not shown).

Preliminary cloning and sequence analysis of OPA-02/3'-hp products therefore demonstrated that 3'-hp DDPCR was potentially an effective marker isolation strategy. 3'-hp DDPCR had the capacity to amplify either homeodomain fragments through specific binding of 3'-hp to homeodomain sequences, or to amplify transcripts randomly through the non-specific binding of 3'-hp to random positions within transcripts, as discussed in 3.2.4.

3.2.4 Discussion of 3'-hp DDPCR Modifications

Modifications to the standard DDPCR primer combinations were made in an attempt to enrich for homeobox gene transcripts. The DDPCR 3' primer was replaced with a degenerate 3' oligonucleotide that should recognise the majority of homeobox transcripts, by binding to a conserved region within the homeobox consensus. This primer, 3'-hp, exhibited different amplification profiles with varying 5' primers, thus indicating the capacity to identify independent sets of transcripts with different primer combinations. Refinement of first strand cDNA template preparation methods led to a significant reduction in the level of background amplification products found in initial 3'-hp DDPCR reactions. This was important as background amplification has the potential to significantly limit successful marker isolation, by contamination during reamplification and cloning procedures (Callard *et al.* 1994).

The modified primer composition of 3'-hp DDPCR led to a reduction in banding complexity in this analysis (3.2.1iii) and additional reactions (3.3.2), with approximately 5 to 10 major products displayed for each primer combination, compared to 50 to 100 in the original technique (Liang and Pardee 1992). While this restricted the number of transcripts identified by each profile and was therefore likely to increase the number of primer combinations required to detect a desired marker, it may also have contributed to lowering

background amplification. This was supported by the demonstration that PCR reamplification of selected transcripts resulted in the amplification of a single major product. These features of 3'-hp DDPCR may lessen the requirement for extensive screening procedures to identify the correct clone within a library of background ligation products (Callard *et al.* 1994).

Expression analysis during ES cell differentiation demonstrated that 3'-hp DDPCR retained the capacity to detect transcripts that are differentially expressed between cell populations. This potentially identified several markers that were restricted to individual or subsets of differentiated ES cell types.

Preliminary analysis of the types of sequences displayed with 3'-hp DDPCR indicated that both randomly primed and homeobox gene transcripts could be amplified. However, *Hox-d8* was the only homeodomain transcript isolated by this analysis, sequencing of further fragments (data not shown), or during a subsequent screen for pluripotent cell markers (3.3). In total, the clones of approximately 20 reamplified products generated from 3'-hp DDPCR were analysed, indicating that homeodomain transcripts were enriched under these conditions, but probably not extensively. This contrasts with a differential display method to enrich the amplification of members of the zinc finger family of transcription factors (Stone and Wharton 1994). In this approach, reverse transcription was primed with an arbitrary 3' oligonucleotide and differential display performed with the same 3' primer in combination with 5' zinc finger primers. Five zinc finger clones were identified by the analysis of twelve amplification products. This suggested that the reverse transcription or PCR conditions in 3'-hp DDPCR were not conducive to a high level of enrichment for homeobox gene transcripts. Conversely, the level of enrichment may be greater in cell populations known to express a large repertoire of homeobox genes, for example at 10.5 d.p.c. of embryogenesis (reviewed in Duboule 1994).

For the purposes of marker isolation, 3'-hp display was regarded as a technique that generated DDPCR profiles from random positions within transcripts, with potential for the amplification of homeobox gene transcripts. Importantly, 3'-hp DDPCR from random positions is more likely to amplify fragments within open reading frames than the original DDPCR technique, which primarily amplifies 3' untranslated regions (Liang and Pardee 1992). This was advantageous during sequence analysis of isolated 3'-DDPCR clones

(3.2.3, 3.4), as it enabled the comparison of transcript coding regions and the identification of homologous sequences in other genes or open reading frames.

3.3 DDPCR Screen for ES Cell Restricted Pluripotent Cell Markers

The capacity of 3'-hp DDPCR to detect and isolate cell type-specific markers was demonstrated by the analysis of gene expression during ES cell differentiation (3.2.1iii, 3.2.2). The application of 3'-hp DDPCR to the ES to X cell transition therefore had the potential to define markers for specific pluripotent cell subpopulations. The key strategy in this approach was to screen for transcripts that were specific to particular stem cell subtypes present during the ES to X cell conversion (1.6.9) and to exclude transcripts expressed by terminally differentiated cells. This was potentially a more powerful approach than the analysis of ES cell differentiation alone, as gene expression analysis suggests that a staged series of pluripotent cell subpopulations is generated during the ES to X cell transition (1.6.9). Additionally, the definition of markers for pluripotent cell subpopulations using ES cell differentiation regimes is complicated by the mixed populations of differentiated and undifferentiated cells that are generated (1.6.5). In comparison, morphology and gene expression analysis has demonstrated that X cell cultures are relatively uniform pluripotent cell populations at a given time point (Lake 1996). Application of 3'-hp DDPCR to the ES to X cell transition therefore provides a system for the identification of markers for subpopulations of pluripotent cells.

Markers for pluripotent cell subpopulations could refine our knowledge of embryogenesis, enable analysis of inaccessible developmental events, provide insight into the molecular basis of pluripotent cell heterogeneity and enable functional analysis of pluripotent cell biology (1.7). Specific markers for ES cells would be particularly useful as they could enable the identification of the embryonic equivalent of ES cells, which has not been defined rigorously (1.6.2). ES cells provide a route for precise transgenesis in the mouse but transfer of this technology to non-murine species has been hampered by an inability to isolate ES cell lines from these organisms (1.6.1). The identification of the embryonic origin of mouse ES cells could therefore lead to alternative isolation strategies for ES cell lines from commercially valuable species.

3.3.1 Identification of Early Pluripotent Cell Specific Markers

To screen for pluripotent cell specific transcripts with expression restricted to ES cells, the gene expression of ES cells, X cells and terminally differentiated ES cell types was compared. MBL-5 ES cells were maintained in the presence of LIF or converted to X cells with MedII in the presence or absence (+ or -) of supplementary LIF, for up to three passages (X+0, 1, 2, or 3; X-0, 1, 2, or 3). X cells of the same passage number were collectively named when appropriate, for example X+1 and X-1 cells were referred to as Xp1. Differentiated MBL-5 ES cell types were generated as before (3.2.1iii). Gene expression in selected combinations of these cell types was compared by 3'-hp DDPCR.

RNA samples were reverse transcribed with 3'-hp and first strand cDNA was amplified with 3'-hp in combination with forty different arbitrary 5' primers (OPERON primer kits OPA and OPB, 2.2.8). The majority of primer combinations displayed approximately five to ten major products, with the maximum being twenty five and the minimum zero. This indicated that the expression patterns of up to 400 transcripts may have been assayed by this analysis. The majority of products were class I transcripts (3.2.1iii), common to all the ES cell-derived cell types assayed and were therefore likely to represent housekeeping gene mRNAs. A significant proportion of bands represented Class II transcripts (3.2.1iii), expressed only in the 10.5 d.p.c. embryonic control, or class III or V transcripts that were induced during ES cell differentiation (3.2.1iii). Nine products, from four different primer combinations, exhibited banding profiles that were potentially specific for pluripotent cell populations and consistent with ES cell restricted expression. These products were termed class VI transcripts. The differential display profiles of eight of these class VI transcripts are shown in figures 3.5, 3.6 and 3.7.

3.3.2 Expression Profile of ES and Early Pluripotent Cell Markers

Figure 3.5 shows the 3'-hp DDPCR profile (2.3.18) of ES cell, X cell and differentiated ES cell types using the OPA-02 5' oligonucleotide. The banding profile was generated from independent reverse transcriptions from the previous OPA-02 profile (Fig. 3.3B), demonstrating consistency in amplification and detection of differential gene expression during ES cell differentiation. Products, A02/240, A02/295, A02/320, A02/350,

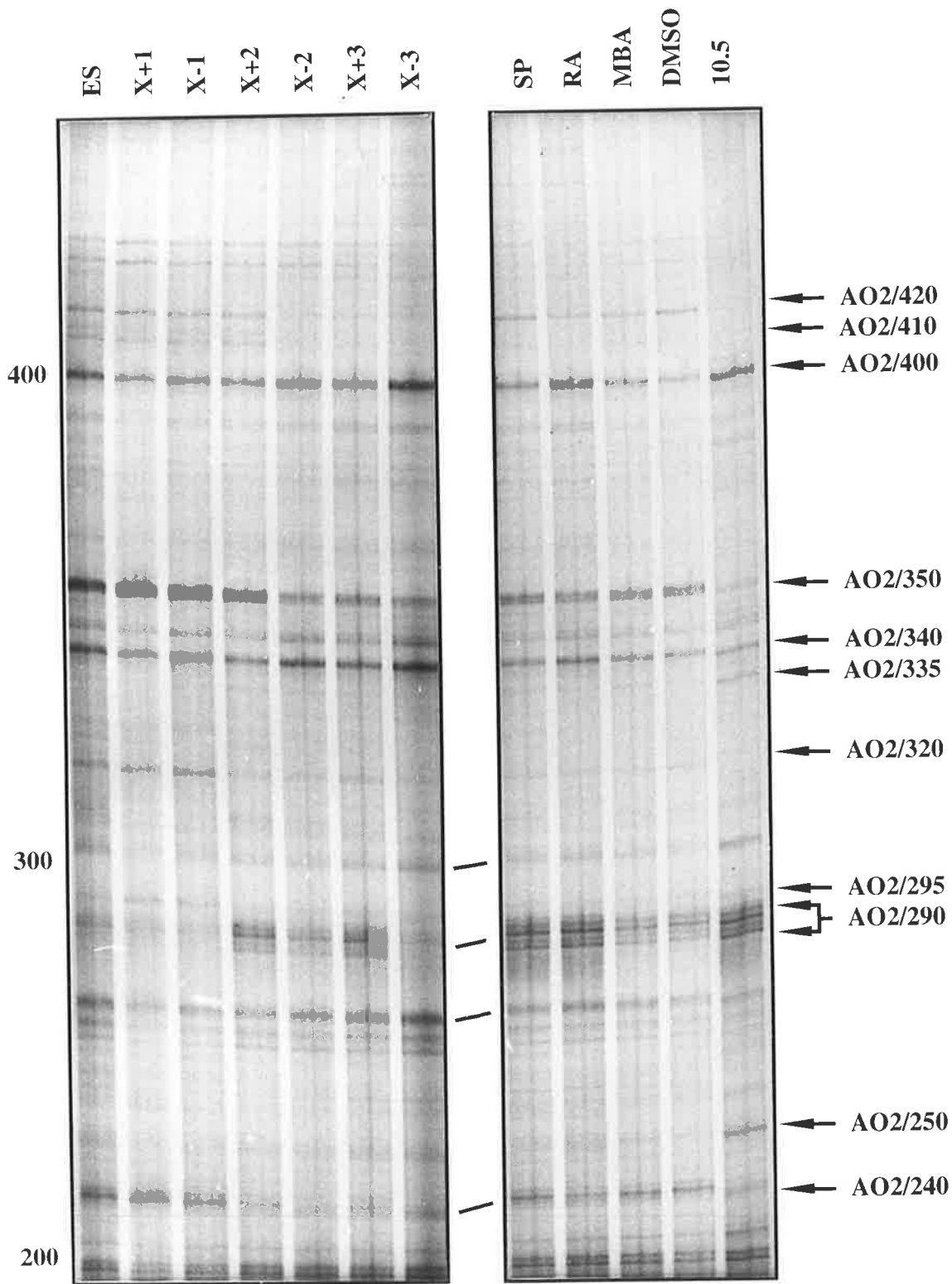
Figure 3.5 Identifying Markers of Embryonic Stem Cells:

A02/295, A02/320 and A02/410.

2 µg cytoplasmic RNA was reverse transcribed (2.3.18) with 3'-hp for 1 hour at 52°C in a 20 µl volume. 1/20th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR (2.3.18) with OPA-02 (2.2.8). [α -³³P] dATP was incorporated. DDPCR products were labelled by the arbitrary primer and approximate size. ES, undifferentiated MBL-5 ES cells (2.4.1); X+/-1, X cells + or - LIF / 1 passage; X+/-2, X cells + or - LIF / 2 passages; X+/-3, X cells + or - LIF / 3 passages (2.4.2i); SP, spontaneous (-LIF) induction (2.4.2ii); RA, retinoic acid induction (2.4.2iii); MBA, 3-methoxybenzamide induction (2.4.2iv); DMSO, dimethylsulfoxide induction (2.4.2v); 10.5, total RNA from a 10.5 d.p.c. embryo (2.3.21). A02/240, A02/350 and A02/420 were not significantly downregulated in differentiated cell types and were disregarded. A02/290 (*Hox-d8*, 3.4.2) expression in X+/-2 and X+3 cells was likely to be an indicator of background differentiated cell types in these cultures. 3'-hp DDPCR reactions were run on the same gel, intervening lanes between X-3 and SP reactions are not shown. Approximate sizes are indicated.

ES and early pluripotent transcripts:

| | |
|-------------------|--------------------|
| Class VI: A02/295 | (ES, X+/-1 > 10.5) |
| A02/320 | (ES, X+/-1) |
| A02/410 | (ES, X+/-1 > X+2) |



A02/410 and A02/420 exhibited downregulation of expression as ES cells converted to X cells. However, as A02/240, A02/350 and A02/420 were not significantly downregulated in differentiated ES cell types, they were disregarded as candidate pluripotent specific transcripts. A02/295 and A02/320 were detected in ES and Xp1 cells, but were downregulated in Xp2, Xp3 and differentiated ES cells. A02/410 shared a similar pattern, but was also detectable in X+2 cells. Therefore A02/295, A02/320 and A02/410 were considered as candidate ES or early pluripotent subpopulation markers, or class VI transcripts (3.3.1). Class VI transcripts were termed early pluripotent cell markers because they were expressed in ES cells and during early X cell passages, and were therefore predicted to be expressed in the ICM and in early pluripotent cell types downstream of the ICM *in vivo*.

Expression of the homeobox gene *Hox-d8* (A02/290, Fig. 3.3B) was detected in Xp2 and X+3 cultures using OPA-02/3'-hp DDPCR (Fig. 3.5). Given the expression of *Hox-d8* during spontaneous and retinoic acid induced differentiation (Figs. 3.3B, 3.5), this was likely to be an indicator of background differentiated cell types present in these cultures.

Figure 3.6 shows the candidate ES cell and Xp1 markers, A03/280 and A03/360, identified in the OPA-03/3'-hp DDPCR profile (2.3.18). A closely migrating doublet and upper band with identical expression were collectively classified A03/280. A03/280 and A03/360 were both expressed in ES cells and Xp1 cells, but were significantly downregulated in Xp3 cells. A03/280 and A03/360 were detected at a reduced level in differentiated cell types and were barely detectable in the 10.5 d.p.c. embryonic control. This indicated that A03/280 and A03/360 were also likely to be downregulated during ES cell differentiation, and were therefore class VI transcripts.

Figure 3.7 demonstrates the OPB-04/3'-hp expression pattern (2.3.18) of three candidate ES and early pluripotent cell markers in ES cell, X+2 and X+3 cultures and in RA and MBA-induced ES cells. B04/340 (arrowhead) was detected at a low level in ES and MBA cells, but not in X+2, X+3 or RA differentiated cells. B04/400 was expressed at the highest level in ES cells, was downregulated in X+2 cells, and was significantly downregulated in X+3 cells. B04/400 was also downregulated in RA and MBA-induced ES cells, possibly indicating that expression was restricted to residual stem cell types in these cultures (1.6.5). B04/410 (open dot) exhibited a similar banding profile to B04/400, but appeared to be

Figure 3.6 Identifying Markers of Embryonic Stem Cells:

A03/280 and A03/360.

2 µg cytoplasmic RNA was reverse transcribed (2.3.18) with 3'-hp for 1 hour at 52°C in a 20 µl volume. 1/20th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR (2.3.18) with OPA-03 (2.2.8). [α -³³P] dATP was incorporated. DDPCR products were labelled by the arbitrary primer and approximate size. ES, undifferentiated MBL-5 ES cells (2.4.1); X+/-1, X cells + or - LIF / 1 passage; X+/-3, X cells + or - LIF / 3 passages (2.4.2i); SP, spontaneous (-LIF) induction (2.4.2ii); RA, retinoic acid induction (2.4.2iii); MBA, 3-methoxybenzamide induction (2.4.2iv); DMSO, dimethylsulfoxide induction (2.4.2v); 10.5, total RNA from a 10.5 d.p.c. embryo (2.3.21). 3'-hp DDPCR reactions were run on the same gel, intervening lanes between X-3 and SP reactions were removed and are not shown. Approximate sizes are indicated.

ES and early pluripotent transcripts:

Class VI: A03/280, A03/360

(ES, X+/-1 > SP, RA, MBA, DMSO)

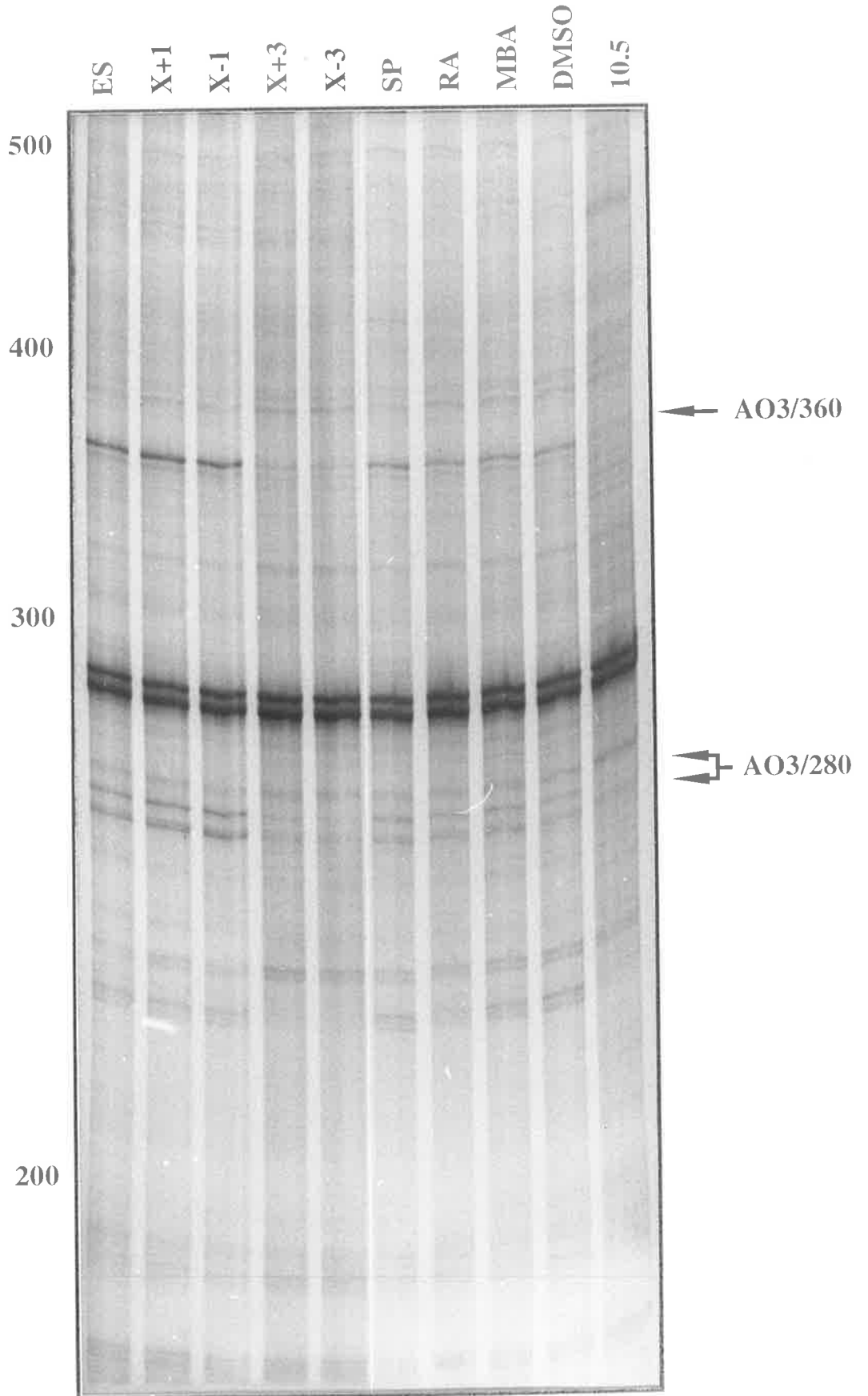


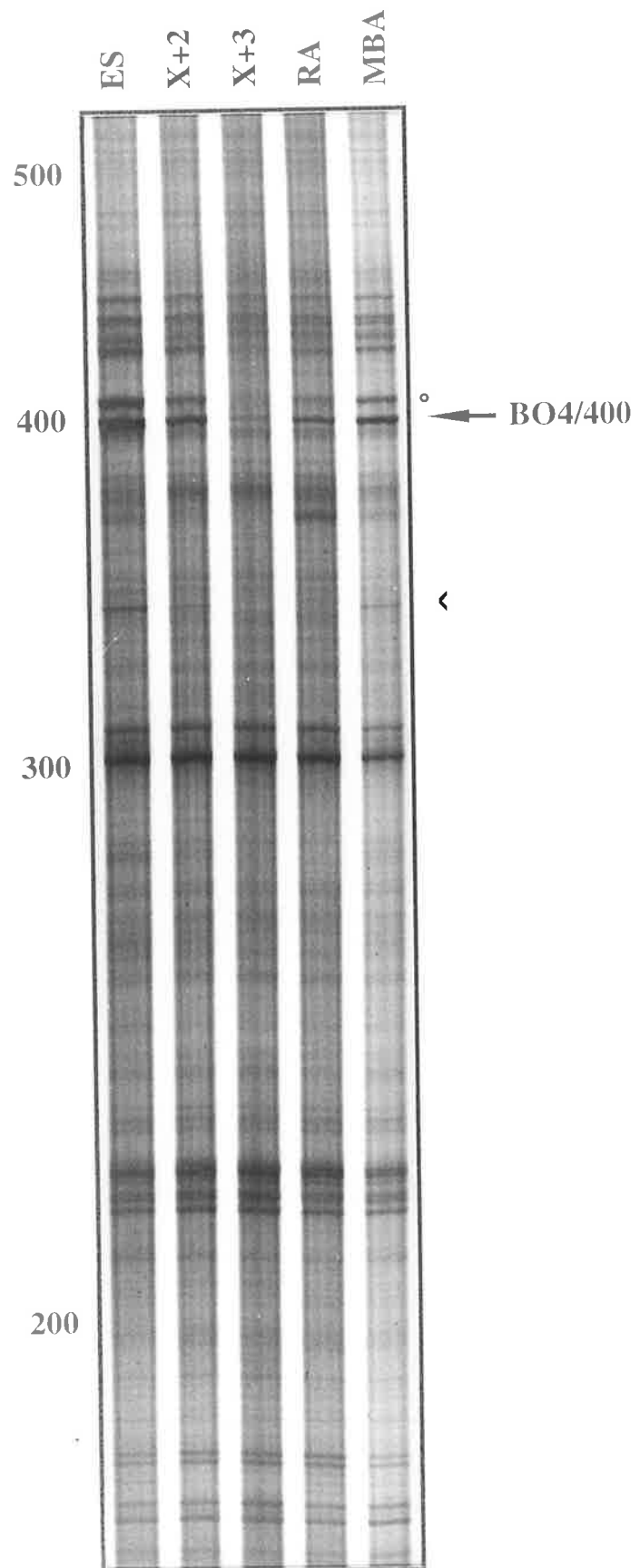
Figure 3.7 Identifying Markers of Embryonic Stem Cells:

B04/340, B04/400 and B04/410.

2 µg cytoplasmic RNA was reverse transcribed (2.3.18) with 3'-hp for 1 hour at 52°C in a 20 µl volume. 1/20th of undiluted first strand cDNA was used as the template for 3'-hp DDPCR (2.3.18) with OPB-04 (2.2.8). [α -³³P] dATP was incorporated. DDPCR products were labelled by the arbitrary primer and approximate size. ES, undifferentiated MBL-5 ES cells (2.4.1); X+/-1, X cells + or - LIF / 1 passage; X+/-2, X cells + or - LIF / 2 passages; X+/-3, X cells + or - LIF / 3 passages (2.4.2i); SP, spontaneous (-LIF) induction (2.4.2ii); RA, retinoic acid induction (2.4.2iii); MBA, 3-methoxybenzamide induction (2.4.2iv); DMSO, dimethylsulfoxide induction (2.4.2v); 10.5, total RNA from a 10.5 d.p.c. embryo (2.3.21). Approximate sizes are indicated.

ES and early pluripotent transcripts:

| | |
|-------------------------------|----------------------|
| Class VI: B04/340 (arrowhead) | (ES > X+2, MBA) |
| B04/400 | (ES > X+2, MBA > RA) |
| B04/410 (open dot) | (ES > X+2, MBA > RA) |



expressed at a lower level, most noticeable in RA and MBA differentiated cells. Both B04/400 and B04/410 (open dot) appeared to be expressed at a higher level in MBA-induced cultures than in RA cells. This may indicate variable levels of residual stem cell populations in these cultures. B04/340, B04/400 and B04/410 were therefore class VI transcripts and candidate markers of ES and early pluripotent cell subpopulations.

3.3.3 Summary of 3'-hp DDPCR Screen for Early Pluripotent Cell Markers

The comparative expression patterns of the identified class VI products, transcripts restricted to ES cell or early pluripotent cell types, are summarised in figure 3.8 and allow some general observations to be made. All these products were expressed at the highest level in ES cells, with significant downregulation occurring during X cell formation and little or no expression in Xp3 cells. The expression patterns of these markers in OPA-02 and OPA-03 reactions did not vary between ES and Xp1 cells, demonstrating that ES cells and early passage X cell types were closely related. This also indicated that these transcripts were markers for early pluripotent cells, rather than embryonic stem cell specific. Therefore, the expression patterns of these class VI transcripts indicate that they subdivide the *Oct-4*⁺ pool of pluripotent cells *in vitro* into distinct cell types, and identify an early pluripotent cell subpopulation during the lineage specific differentiation of ES cells to X cells.

During this analysis, no products were identified with expression patterns restricted to specific X cell populations. Specific markers of X cell types were expected because previous analyses have demonstrated alteration in gene expression during the ES to X cell transition (Bettess 1993, Thomas 1995, Lake 1996), for example induction of *FGF-5* during the development of X cells (1.6.8). Therefore, a potential disadvantage of 3'-hp DDPCR was that the reduction in the number of products displayed (3.2.4, 3.3.1) may have resulted in a restricted sample of transcripts being analysed. Using the original DDPCR technique (1.7.6) numerous X cell specific transcripts have been detected (S. Sharma, unpublished data) and consist of three distinct transcript classes. Class VII transcripts are not expressed by ES cells but are induced in all X cell types. A class VIII transcript has been identified that is rapidly induced in Xp0 cultures but is downregulated by Xp3 cells. A series of products that are

| | A | | | | | | | | | B | |
|----------------|---------------------|-----|-----|-----|----|-----|----|----|------|-----|----|
| | RELATIVE EXPRESSION | | | | | | | | | | R |
| | ES | Xp1 | Xp2 | Xp3 | SP | RA | MB | DM | 10.5 | | |
| A02/295 | + | + | - | - | - | - | - | - | +/- | NA | |
| A02/320 | ++ | ++ | - | - | - | - | - | - | - | NA | |
| A02/410 | + | + | +/- | - | - | - | - | - | - | NA | |
| A03/280 | +++ | +++ | | - | + | + | + | + | - | + | + |
| A03/360 | +++ | +++ | | - | + | + | + | + | - | + | + |
| A16/430 | +++ | | - | - | | + | - | | | +/- | - |
| B04/340 | ++ | | + | - | | - | + | | | + | NA |
| B04/400 | +++ | | ++ | - | | + | ++ | | | + | + |
| B04/410 | +++ | | ++ | - | | +/- | + | | | NA | |

Figure 3.8 Summary of ES and Early Pluripotent Cell Markers.

(A) Expression patterns of ES and early pluripotent cell markers, summarised from figures 3.5, 3.6 and 3.7. Relative expression levels are indicated by + or -. Empty boxes represent cell types that were not assayed. X cells of the same passage, grown in the presence or absence of LIF, were grouped. ES, ES cells; Xp1, X cells + or - LIF / 1 passage; Xp2, X cells + or - LIF / 2 passages; Xp3, X cells + or - LIF / 3 passages; SP, spontaneous (-LIF) induction; RA, retinoic acid induction; MB, 3-methoxybenzamide induction; DM, dimethylsulfoxide induction; 10.5, 10.5 d.p.c. embryo. A02/410 was expressed in X+2 cells but not in X-2 cells. A16/430, B04/340, B04/400 and B04/410 expression was not assayed in X-2 or X-3 cells.

(B) ES and early pluripotent cell marker reamplification and verification procedures, summarised from 3.4, 3.4.1 and 3.4.2. (R) reamplification and cloning of DDPCR fragments; (E) verification of expression pattern by northern blot. Success is indicated by + or -; NA, not attempted. A03/280, A03/360 and B04/400 were reamplified, cloned and verified the expected expression pattern by northern blot. A16/430 did not reamplify effectively and clones that verified the expected expression pattern were not identified. Analysis of B04/340 was discontinued after sequence analysis revealed it was not novel.

induced in Xp2 and Xp3 cultures are termed class IX transcripts. These transcripts subdivide X cell types into multiple pluripotent cell subpopulations.

The downregulation of early pluripotent cell marker expression (Fig. 3.8) appeared to occur either between Xp1 and Xp2 cultures, or between Xp2 and Xp3 cells. This timing was consistent with the downregulation of *Rex-1* expression during X cell formation (Lake 1996). Further analysis may indicate whether these transcripts are simultaneously downregulated during X cell development, and therefore define a major transition point during the development of early passage X cells to a primitive ectoderm-like cell.

All the class VI markers exhibited marked downregulation in at least some of the differentiated cell cultures. This was likely to indicate a restriction of marker expression upon induction of differentiation, with later expression potentially limited to residual stem cell nests within these cultures (1.6.5).

The comparison of gene expression in ES cells and X cells therefore demonstrated that transcripts that were differentially expressed between these cell types could be detected. Isolation of these products is likely to provide markers for early subpopulations of *Oct-4*⁺ pluripotent cells *in vivo* and distinct X cell subpopulations *in vitro*.

3.4 Isolation and Cloning of ES and Early Pluripotent Cell Markers

Five of the putative early pluripotent cell markers identified by the 3'-hp DDPCR screen were selected for reamplification and cloning, as indicated in figure 3.8. A03/360 and the top band of A03/280 (Fig. 3.6), B04/340 and B04/400 (Fig. 3.7) were excised, eluted and reamplified (2.3.18). The A16/430 product from OPA-16/3'-hp DDPCR (not shown), reamplified poorly and screening to isolate the correct clone was not pursued. A03/280, B04/340 and B04/400 fragments were cloned by digestion with *EcoRI* and ligation into an *EcoRI*/T overhang pBluescript KS vector (2.3.6ii). A03/360 contained an internal *EcoRI* site, at nucleotide positions 252 to 257 (Fig. 3.10A). The 256 bp 5' A03/360 fragment was cloned into pBluescript KS *EcoRI*/T and the 108 bp 3' A03/360 fragment was cloned into an *EcoRI* pBluescript KS vector (2.3.6i).

A single clone for each marker fragment was sequenced in one direction to identify homeobox gene transcripts (3.2.3) or previously characterised genes. B04/340 corresponded

to a hydrophobic murine histocompatibility antigen (Loveland *et al.* 1990; Accession no. L07095) and was eliminated from subsequent analysis. A03/280 spanned approximately position 2836 bp to 3070 bp of a *LINE-1* (Loeb *et al.* 1986; Accession no. M13002) repetitive retrotransposon element expressed as a 7 kb mRNA in EC cells (Schichman *et al.* 1992) and during preimplantation development (Packer *et al.* 1993). The expression of A03/280 was found to be greater in ES cells than Xp3 cells by a preliminary northern blot (data not shown) but was not analysed further. A03/360 and B04/400 were novel sequences. A03/360 was not homologous to any described sequence and B04/400 contained a region of homology to a predicted *C. elegans* open reading frame (3.4.2). A03/360 and B04/400 were re-sequenced as described in figure 3.10 and selected for more detailed characterisation. No transcripts derived from homeobox genes were identified.

3.4.1 Validation of A03/360 and B04/400 Expression Patterns

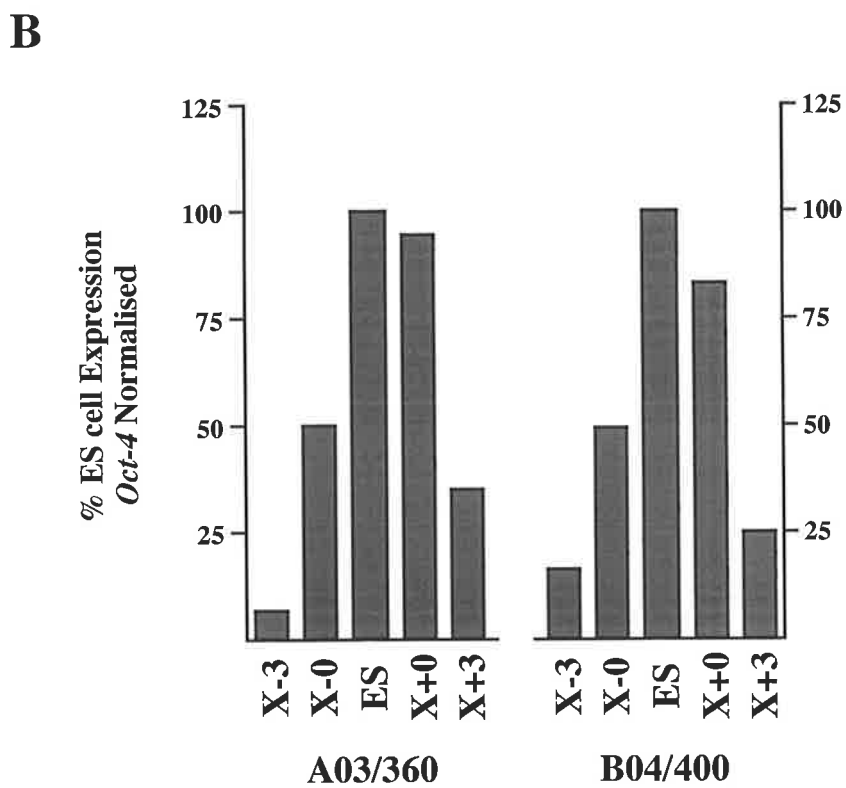
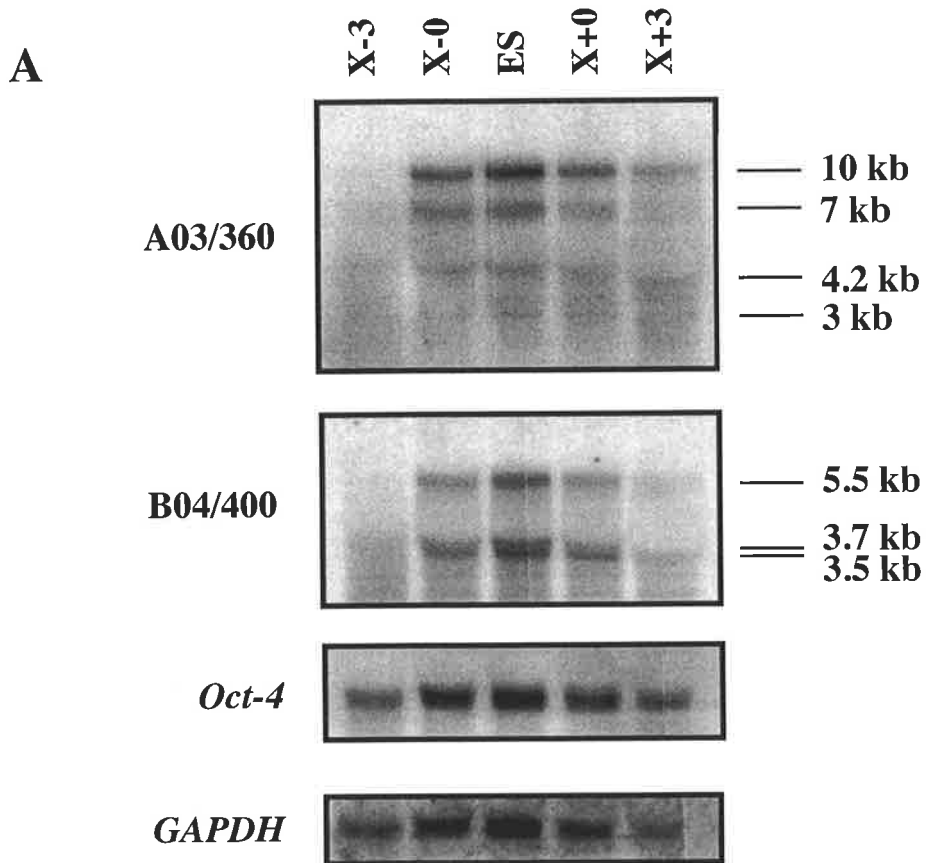
Cloning of transcripts following DDPCR can be complicated by stochastic background amplification products in the primary DDPCR which carry through to reamplification and cloning. This can lead to considerable difficulty when attempting to identify clones that verify the original expression profile (Callard *et al.* 1994). It is therefore necessary to validate the expression patterns of individual cDNA clones, thus demonstrating the isolation of the desired fragment. These validity procedures were required for A03/360 and B04/400, despite the indication that reamplification procedures from 3'-hp DDPCR gels result in the generation of a single major species (3.2.2). Northern blot analysis of A03/360 and B04/400 expression was carried out in conjunction with Mr. M. Bettess (Department of Biochemistry, University of Adelaide).

To demonstrate that the cloned A03/360 and B04/400 cDNAs recognised transcripts that were restricted to ES and early pluripotent cell populations, northern blot analysis of ES and X cell populations was carried out. A northern blot of D3 ES cell, Xp0 and Xp3 cell RNAs was sequentially probed with the 5' (256 bp) A03/360 cDNA fragment (3.4, Fig. 3.10A), the B04/400 cDNA fragment (3.4, Fig. 3.10B), *Oct-4* (2.2.6iii) and *GAPDH* (2.2.6ii) controls and is shown in figure 3.9A. Both A03/360 and B04/400 verified the

Figure 3.9 Northern Blot Verification of A03/360 and B04/400 Expression in ES and X cells.

(A) 10 μg (A)_n⁺ RNA isolated from ES cells and X cells was electrophoresed on a 1.3% agarose gel (2.3.23) and blotted (2.3.24). The northern filter was hybridised with a double stranded A03/360 probe, isolated by *EcoRI/HindIII* digestion of the 256 bp 5' cDNA clone (3.4, Fig. 3.10A). The filter was washed once in 2 X SSC / 0.1% SDS at 42°C for 15 min. and once at 65°C for 30 min. The filter was then washed in 0.1 X SSC / 0.1% SDS at 65°C for 15 min. The northern blot was exposed in a phosphorimager cassette overnight. The filter was stripped and reprobbed with a double stranded B04/400 probe, isolated by *EcoRI/HindIII* digestion of the 458 bp cDNA clone (3.4, Fig. 3.10B) and washed once in 2 X SSC / 0.1% SDS at 42°C for 15 min. and once at 65°C for 30 min. The filter was then washed in 0.2 X SSC / 0.1% SDS at 65°C for 15 min. The northern blot was exposed in a phosphorimager cassette overnight. The filter was stripped, reprobbed with *Oct-4* (2.2.6iii), washed and exposed as above. The filter was stripped, reprobbed with *GAPDH* (2.2.6ii) as a loading control and exposed as above. The estimated sizes of A03/360 and B04/400 transcripts are indicated. ES, undifferentiated D3 ES cells. X cells were generated by the culture of D3 ES cells in MedII (2.4.2i). X-3, X cells -LIF / 3 passages; X-0, X cells -LIF / 0 passages; X+0, X cells +LIF / 0 passages; X+3, X cells +LIF / 3 passages. *Oct-4* transcript: 1.6 kb. *GAPDH* transcript: 1.5 kb.

(B) Graphic representation of A03/360 and B04/400 expression in ES and X cells. The expression level of the 10 kb A03/360 and the 5.5 kb B04/400 transcripts were quantitated by volume integration (2.3.35), normalised against *Oct-4* expression and depicted as percentage of ES cell expression.



expected expression profile of ES cell expression and downregulation of expression during X cell formation. This demonstrated the isolation of the desired DDPCR fragment.

The A03/360 probe identified two major transcripts at 10 kb and 7 kb and minor transcripts at 4.2 kb and 3 kb (Fig. 3.9A). The 10 kb A03/360 transcript exhibited the highest expression level. The B04/400 probe identified three transcripts with similar expression levels, at 5.5kb and a doublet at 3.7kb and 3.5kb (Figs. 3.9A, 3.10C). A03/360 and B04/400 were expressed at the highest level in ES cells and their expression was downregulated in Xp0 cells. Expression of A03/360 and B04/400 was significantly downregulated in Xp3 cells.

Oct-4 is expressed at a relatively consistent level in pluripotent cell types during X cell formation (Lake 1996). The diminished expression of *Oct-4* and *GAPDH* observed in Xp3 cultures (Fig. 9A) was therefore attributed to a reduced proportion of pluripotent cell types in these cultures, or an under-loading of these lanes. To take account of this and clarify the expression of A03/360 and B04/400 during X cell formation, the 10 kb A03/360 and 5.5 kb B04/400 transcripts were quantitated by volume integration (2.3.35) and normalised against *Oct-4*. The *Oct-4* normalised expression levels of A03/360 and B04/400 are presented as a percentage of ES cell expression in figure 3.9B. A03/360 and B04/400 were detected at the highest level in ES cells (100%) and were expressed in X+0 cells at 94% and 83% respectively and their expression was downregulated to 50% and 49% respectively in X-0 cultures. Expression of A03/360 and B04/400 was downregulated in X+3 cultures to 35% and 25% respectively, and to 7% and 16% respectively in X-3 cultures.

This analysis highlighted several important features of the expression of A03/360 and B04/400 during the formation of X cells. Firstly, as expected from DDPCR analysis (3.3.3), these markers were restricted to an early subpopulation of *Oct-4*⁺ pluripotent cells, demonstrated by downregulation of expression in Xp3 cells. Secondly, A03/360 and B04/400 were expressed in early X cell populations, confirming the close relationship of these cells to ES cells (3.3.3). Thirdly, the downregulation of A03/360 and B04/400 expression was more rapid in X cell cultures passaged in the absence of LIF, than in cultures with LIF at the equivalent passage. This demonstrates that LIF may inhibit the developmental progression of X cells, as has previously been suggested (Thomas 1994, Lake 1996).

Finally, the transcripts identified by the A03/360 and B04/400 probes appeared to be regulated between cell types in a similar manner (Fig. 9B). This may indicate that these transcripts are coordinately regulated during the formation of X cells.

3.4.2 Sequence of A03/360 and B04/400 Markers

The nucleotide sequence and putative open reading frames of A03/360 and B04/400 cDNA fragments are shown in figure 3.10. The sizes of the A03/360 and B04/400 cDNA fragments were 360 bp and 458 bp respectively, excluding bases removed during restriction digestion and cloning, with original names retained for simplicity.

Database search (2.3.36) indicated that A03/360 did not exhibit significant similarity to described sequences, in any reading frame. The only reading frame without stop codons was 108 residues long and in the expected orientation, as is indicated in figure 3.10A.

The B04/400 clone (Fig. 3.10B) revealed no significant homology at the nucleotide level by database search (2.3.36). However, when all six reading frames were translated and compared, a region of similarity to a predicted *C. elegans* open reading frame was detected. A 96 amino acid sequence in the B04/400 open reading frame (ORF), spanning positions 25 to 120, contained 44.8% amino acid identity (43/96 residues) to positions 348 to 443 of the predicted ORF CLEB0336.3 10 on the *C. elegans* cosmid B0336 10 (Wilson *et al.* 1994; accession no. U32305), as indicated in figure 3.9B.

The 121 amino acid B04/400 ORF encoding this similarity was on the opposite strand from that predicted by the primer orientation, with 3'-hp at the 5' end of the clone and at the amino terminal end of the predicted ORF (Fig. 3.10B). Northern blot hybridisation of ES cell RNA was used to determine the orientation of the transcripts recognised by the B04/400 product (Fig. 3.10C). A riboprobe generated by *Hind*III digestion of the B04/400 cDNA clone (2.2.7ii, 3.4) and transcription with T7 RNA polymerase (2.3.26), representing the "sense" strand primed by 3'-hp, did not identify any antisense transcripts expressed in ES cells (lane 1). An "antisense" B04/400 strand, generated by *Bam*HI digestion of the B04/400 cDNA clone (2.2.7ii, 3.4) and transcription with T3 RNA polymerase (2.3.26), detected the 3.5 kb, 3.7 kb and 5.5 kb transcripts expressed by ES cells (lane 2). A double stranded probe, isolated by *Eco*RI/*Hind*III digestion of the B04/400 cDNA (3.4), detected the same

Figure 3.10 Sequence Analysis of A03/360 and B04/400 cDNA Fragments.

The nucleotide sequence and conceptual translation of the A03/360 and B04/400 3'-hp DDPCR clones and the verification of B04/400 transcript orientation are shown. Primer sequences are labelled and in bold, *EcoRI* restriction sites are underlined. Sequence external to the 3'-hp *EcoRI* restriction site and Taq polymerase adenosine additions (not shown) were not included in the nucleotide size. Primer sequences were not translated and the open reading frames are numbered from the first non-primer codon.

(A) The nucleotide sequence of the 360 bp A03/360 cDNA fragment and the conceptually translated 108 residue open reading frame (ORF) are indicated. The A03/360 cDNA (3.4) contained an internal *EcoRI* site at nucleotide position 252-257. The 256 bp 5' A03/360 fragment (OPA-03 end) was cloned by ligation to pBluescript *EcoRI*/T (2.3.6ii) and the 108 bp 3' A03/360 fragment (3'-hp end) was cloned by ligation to pBluescript KS *EcoRI* (2.3.6i). The 5' and 3' A03/360 cDNA clones were sequenced in both directions with T3 and T7 primers. Other reading frames contained numerous stop codons. The A03/360 ORF did not contain extended regions of similarity to database sequences.

(B) The nucleotide sequence of the 458 bp B04/400 cDNA fragment and the conceptually translated 121 residue open reading frame (ORF) are indicated. The B04/400 cDNA fragment (3.4) was cloned by ligation to pBluescript *EcoRI*/T (2.3.6ii). The depicted sequence was derived from overlapping sequence from each end of the B04/400 DDPCR cDNA clone with T3 and T7 primers and in both directions over this region by the sequencing of B04/400 transcripts (see Fig. 5.1). The B04/400 ORF was encoded on the cDNA strand primed by the 3'-hp primer and was termed the "sense" strand. The B04/400 ORF contained 43/121 (35.5%) identical residues (boxed) to positions 325-445 of a predicted *C. elegans* ORF, CLEB0336.3 10 (accession no. U32305). Amino acid identity was 44.8% in a 96 residue region (43/96) between B04/400 positions 25-102 and CLEB0336.3 10 positions 348-443.

(C) Northern blot verification of the orientation of B04/400 transcripts. 5 μ g (A)_n⁺ RNA isolated from D3 ES cells was electrophoresed on a 1.3% agarose gel (2.3.23) and blotted (2.3.24). Parallel northern filters were probed with either a sense riboprobe (1), or an antisense riboprobe (2), prepared from the B04/400 cDNA (2.2.7ii, Fig. 3.10B). The filters were washed twice in 2 X SSC, 1% SDS at 65°C for 15 min. and once in 0.2 X SSC, 1% SDS at 75°C for 20 min. The northern blots were exposed in a phosphorimager cassette overnight. The sense riboprobe northern blot (1) was stripped and reprobed with a double stranded B04/400 probe (3), isolated by *EcoRI*/*HindIII* digestion of the B04/400 cDNA (3.4, Fig. 3.10B). The filter was washed twice in 2 X SSC, 0.1% SDS at 42°C for 20 min and once at 65°C for 20 min. The filter was then washed in 0.2 X SSC, 0.1% SDS at 65°C for 20 min. and exposed in a phosphorimager cassette overnight. The sizes of the transcripts recognised by the B04/400 probe are indicated.

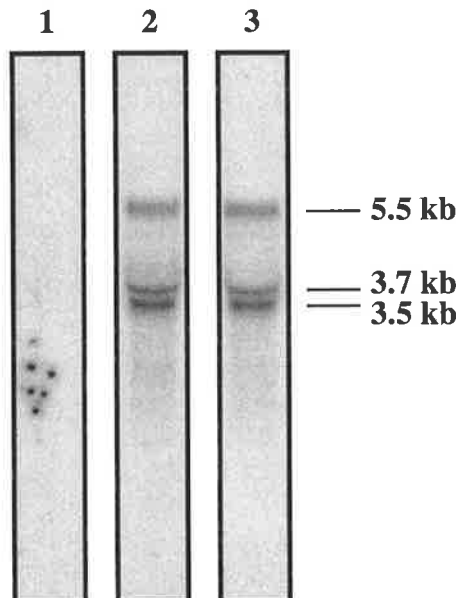
A

A03/360 ORF *OPA-03*
AGTCAGCCACCA AGA CAC CAT CGT GGA CCC CCA CAG AGA GAC TGG AAA TGG GAA AAG GAT GGC 63
 R H H R G P P Q R D W K W E K D G 17
 TTT AAT AGT ACT AGA AAA AAC AGC TTT CCA CAT TCT TTG AGG AAT AGT GGT GGA CCA AGA 123
 F N S T R K N S F P H S L R N S G G P R 37
 GGA AGT TCT GTC TGG CAT AAA GGT GCT ACA AGA GGC TCT TCA ACT TGG TTC CTT AAT CAT 183
 G S S V W H K G A T R G S S T W F L N H 57
 AGC AAT TCT GGA GGA GGC TGG CAT TCA AAC AAT GGG ATG GTA GAT TGG AAT TAT AAT GGT 243
 S N S G G G W H S N N G M V D W N Y N G 77
 ACA GGA AGG AAT TCC AGT TGG CAT TCT GAA GGA ACA GGT GGC TTT CCC AGC TGG CAC ATG 303
 T G R N S S W H S E G T G G F P S W H M 97
 AAC AAC AGT AAC GGA AAC TGG AAA TCC AGT GTA **CGTGGTTCCAAAACCGTCGAATTC** 360
 N N S N G N W K S S V *3'-hp* 108

B

B04/400 ORF *3'-hp*
GAATTCGACGGTTTTTGGAAACCA GAG AGC CGA AAG AGA GCT ATT AGT GGT TTG GAG GGG CCA CTT 65
 E S R K R A I S G L E G P L 14
 A A Q G V V T Q P G S F R G 338
 ACA AAG AAG CCT TGG CTG GGA AAA CAA GGG AAC AAC AAT CAA AGT AAA CCA GGC TTC CTG 125
 T K K P W L G K Q **G** N N N Q S K P **G F** L 34
 A P R G R G N I R **G** R G G F A R G **G F** T 358
 AGG AAG AAT CAT TAT ACA AAC ACC AAA CTA GAG **GTG** AAA AAA ATC CCC CAG GAA TTG AAC 185
 R K N H Y T **N** T K **L** E **V** K **K I P** Q **E** L **N** 54
 G A I N R D **N** C T **L** Q **V** A **K I P** P **E** M **N** 378
 AAC ATT ACC AAG CTC AAC GAA CAC TTC AGC AAA TTT GGA ACC ATA GTT AAT ATC CAG GTT 245
 N **I** T **K L N E H F** S K **F G T** I V **N I Q V** 74
 T **I** A **K L N E H F** A T **F G T** V D **N I Q V** 398
 GCT TTT AAG GGT GAT CCA GAA GCA GCA CTC ATC CAA TAC CTT ACC AAT GAA GAG GCC AGG 305
 A F K **G** D P E A **A L** I Q Y L T N E E **A** R 94
 R Y N **G** E I D S **A L** V T Y A S K F D **A** G 418
 AAA GCC ATT TCT AGC ACA GAA GCA GTT TTG AAC AAT CGG TTC ATT CGA GTC CTG TGG CAT 365
K A I S **S** T E A **V L N N R F I** R **V** L **W H** 114
K A Y K **S** P T P **V L N N R F I** K **V** F **W H** 438
 AGA GAG AAC AAT GAG CAA CCA **GCACTCCAGTCC** 458
 R E N N **E** Q P *OPB-04* 121
 N P G G **E** G S 445

C



three transcripts as the antisense probe (lane 3). This analysis demonstrated that the only B04/400 transcripts detectable by northern blot in ES cells were sense transcripts.

The sense strand orientation of B04/400 transcripts indicated that the B04/400 DDPCR product was generated by amplification with primers in the opposite orientation to that expected. Sense strand amplification with the opposite primer orientation may have resulted from two different mechanisms. Reverse transcription by Taq polymerase (Jones and Foulkes 1989) may have been primed by OPB-04, on carried over RNA in the PCR reaction, followed by subsequent OPB-04/3'-hp PCR amplification. Alternatively, the same result may have been generated by 3'-hp primed reverse transcription and amplification with primers in the opposite orientation, due to primer binding at nested fortuitous sequences within the first strand cDNA, at low annealing temperatures.

3.4.3 *In situ* Analysis of A03/360 and B04/400 Expression in ES Cell Monolayers

A second validation step was required to confirm that the cells expressing A03/360 and B04/400 by northern blot were pluripotent cells and not sporadic cell types arising by spontaneous differentiation *in vitro*, generally located at the periphery of undifferentiated ES cell colonies.

Direct visualisation of A03/360 and B04/400 expressing cells was carried out by *in situ* hybridisation analysis of cell monolayers. A03/360 expression was detected using a digoxigenin labelled (DIG, 2.3.29) antisense riboprobe generated from the 5' (256 bp) cDNA clone (2.2.7i) and B04/400 transcripts were detected with an antisense DIG riboprobe (2.3.29) generated from the B04/400 DDPCR clone (2.2.7ii). Both A03/360 and B04/400 exhibited similar expression in D3 ES cell cultures (purple staining), with transcripts restricted to the undifferentiated pluripotent cells in the centre of ES cell colonies (Fig. 3.11A, B). Differentiated cells within the culture (for example, Fig. 3.11 A03/360 A, B) and those at the edge of ES cell colonies did not express either A03/360 or B04/400. This demonstrated that A03/360 and B04/400 were expressed by pluripotent cells within ES cell cultures and importantly, that they were specifically downregulated during cellular differentiation. The expression levels of A03/360 and B04/400 appeared to be similar in this analysis, with

Figure 3.11 Localisation of A03/360 and B04/400 Expression in ES Cell Cultures.

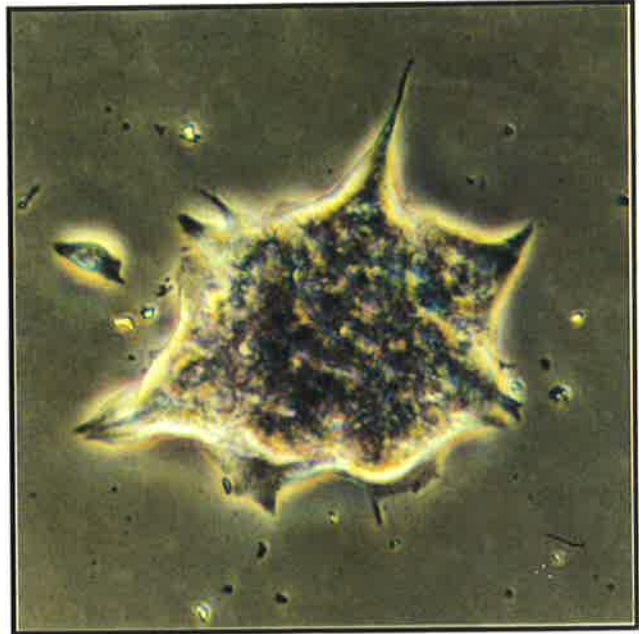
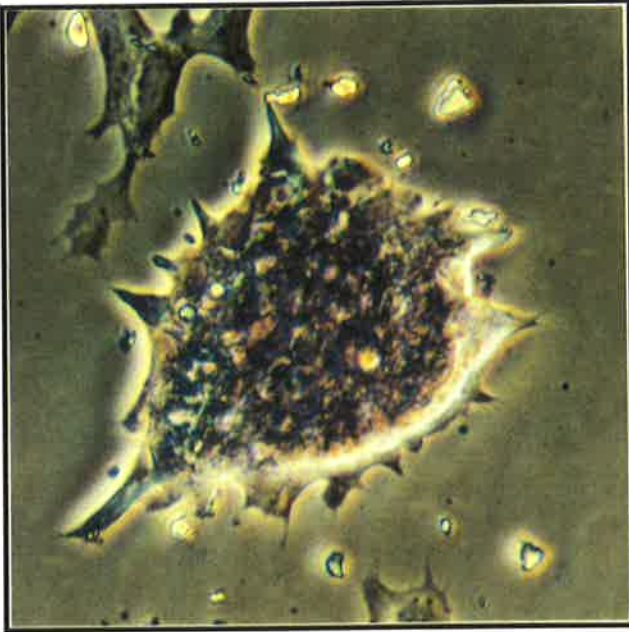
Detection and localisation of A03/360 and B04/400 expression in embryonic stem cell cultures was carried out by *in situ* hybridisation (2.3.30). D3 ES cells were cultured for three days in the presence of LIF (2.4.1). A03/360 expression was detected (1, 2) using a DIG labelled antisense riboprobe (2.3.29) of the 256 bp 5' A03/360 cDNA (2.2.7i). B04/400 expression was detected (1, 2) using a DIG labelled antisense riboprobe (2.3.29) of the 458 bp B04/400 cDNA (2.2.7ii). *In situ* hybridisation was also carried out with A03/360 and B04/400 sense probes (2.2.7i, 2.2.7ii) as negative controls (3). Cells that expressed A03/360 or B04/400 were indicated by purple staining. The A03/360 and B04/400 plates were developed for 7 hours. Photographs were taken under phase contrast (1) and bright field optics (2, 3) at 100 X magnification.

D differentiated cell type

AO3/360

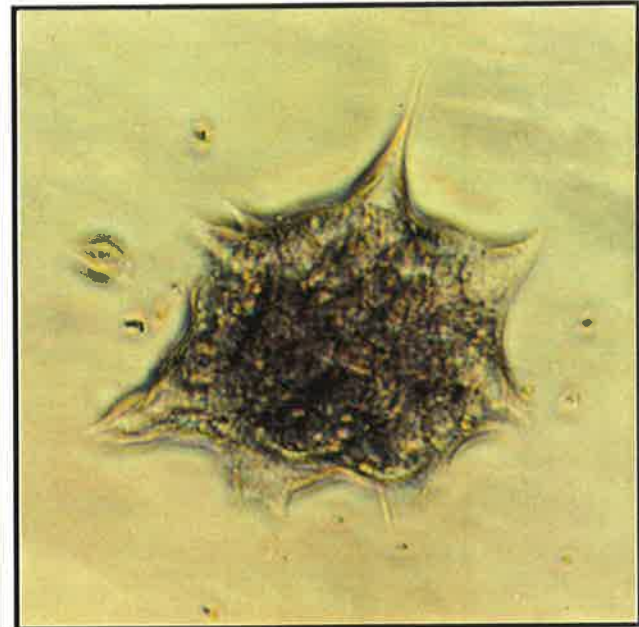
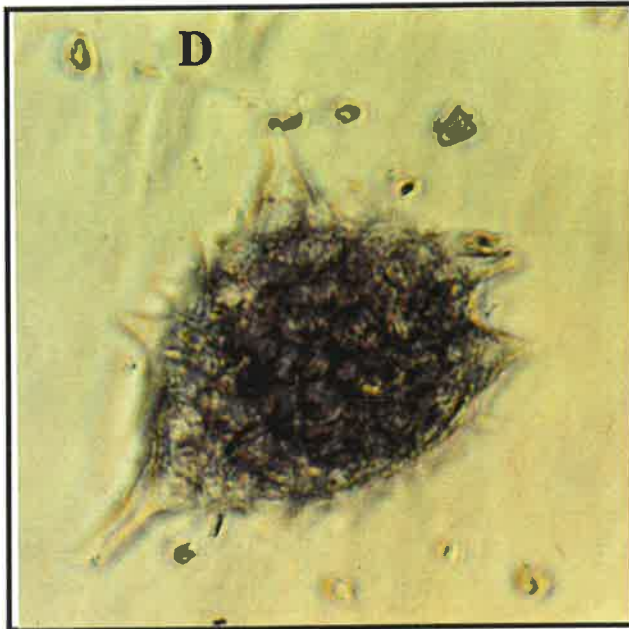
BO4/400

A



B

D



C



equivalent probe and antibody concentrations and exposure times maintained for both markers. Sense control probes for A03/360 and B04/400 did not exhibit any staining (Fig. 3.11C), demonstrating the specificity of this analysis.

3.5 Discussion

This chapter described the investigation of differential display PCR as a method to identify markers for pluripotent cell subpopulations *in vitro*. A modified DDPCR was developed in an attempt to enrich for homeobox gene transcripts, therefore potentially isolating key genes involved in the control of pluripotent cell development (3.1.2, 3.1.3). The major modification was to substitute the poly(A) tail anchored 3' primer normally used in DDPCR with a degenerate complimentary oligonucleotide, 3'-hp, designed to recognise a conserved region within the homeobox consensus. This modification enabled DDPCR profiles to be generated from random positions within transcripts, with potential for the amplification of homeobox gene transcripts, as discussed in 3.2.4.

In a screen for transcripts that were differentially regulated between pluripotent cells, 3'-hp DDPCR identified nine products that were potentially pluripotent specific and restricted to ES cells or early pluripotent cell types within the ES to X cell transition. All these products exhibited downregulation during ES cell differentiation, indicating they were likely to be specific for the pluripotent state, with residual expression in differentiated cultures caused by the persistence of pluripotent cell nests. All nine products were downregulated during X cell formation, specifically marking ES cells and early X cell subtypes within the pool of *Oct-4*⁺ pluripotent cells *in vitro*. The expression of these transcripts appeared to be consistent with that of the ES cell marker *Rex-1* (1.6.8), which is downregulated during the formation of X cells.

Importantly, this demonstrated the presence of at least three distinct pluripotent cell subtypes within the ES to X cell transition. Firstly, ES cells are distinguished from all X cell types on the basis of morphology (1.6.7), differentiation potential (1.6.8) and DDPCR markers for X cells (3.3.3), indicating that subtle differences occur very early during X cell formation (1.6.9). Secondly, an “early” X cell type, in Xp0 and Xp1 cultures, remained closely related to ES cells on the basis of early pluripotent cell marker expression, for

example A03/360. Thirdly, a “late” X cell type, in Xp2 and Xp3 cultures, exhibited downregulation of these markers. The assignment of X cell subpopulations on the basis of these early pluripotent cell transcripts is supported by the expression of *Rex-1* in ES cells and “early” X cells, and downregulation in “late” X cells (1.6.8); and expression of *FGF-5*, which is induced in “late” X cells (1.6.8). As consistent X cell preparations were not used throughout this analysis and different X cell cultures have the potential to develop at slightly variable rates, the precise assignment of pluripotent cell subtypes to passage number was not possible. It may be possible to achieve this by marker expression analysis in matched X cell preparations. For example, the expression of A02/295, A02/320 and A02/410 (Fig. 3.5) indicated that the transition from the “early” to “late” X cell type occurred between the first and second X cell passage. Further analysis of X cell formation may also reveal if a distinct transition point occurs, where expression of these markers is downregulated simultaneously, or if X cells pass through several intermediate stages in their development, as defined by the downregulation of individual markers in a temporal order.

The distinct morphological, developmental and gene expression characteristics that are acquired during the formation of X cells (1.6.8) suggest that differential gene expression should be detected in specific X cell populations, with the induction of a range of transcripts at specific time points, for example *FGF-5* (1.6.8). Numerous X cell specific transcripts have been detected using the original DDPCR technique (3.3.3, S. Sharma, unpublished data). Specifically, a transcript that is induced in Xp0 cultures but is downregulated by Xp3 cells, and numerous transcripts that are induced in Xp2 and Xp3 cultures, have been identified. These transcripts should provide markers for multiple distinct X cell subpopulations and support the assignment of “early” and “late” X cell subpopulations. These markers could enable identification and analysis of molecular heterogeneity (1.7) during the development of the primitive ectoderm *in vivo* (1.6.8).

An important feature of this DDPCR screen was the availability of multiple, closely related cell types for transcript comparison. This bypassed the requirement to compare duplicate reactions to detect artefactual banding (Liang *et al.* 1993), by demonstrating consistent amplification of common class I transcripts in multiple different cell types, as compared to class II, III, VI, V and VI differential transcripts in the same cell types (3.2.1 iii,

3.3.2). Furthermore, analysis of multiple closely related cell types increases the likelihood of detecting relevant molecular differences, a strategy that was also used by Guimarães *et al.* (1995) during DDPCR analysis of haematopoietic development in the yolk sac and embryoid bodies.

Of the nine early pluripotent cell markers identified, four were cloned and two, A03/360 and B04/400, were characterised in detail. The identity of the A03/360 and B04/400 cDNA clones was validated by northern blot. Interestingly, A03/360 and B04/400 expression appeared to be regulated in a similar manner during the ES to X cell transition, with similar levels of downregulation found in individual cell populations (Fig 3.9B). This indicates that these transcripts may be coordinately regulated as these pluripotent cell subtypes develop. The coordinated regulation of transcripts during the ES to X cell transition may also be suggested by the similarity of A03/360 and B04/400 expression to that of *Rex-1* (Lake 1996). Furthermore, the downregulation of A03/360 and B04/400 expression was more distinct in X cells cultured in the absence of additional LIF, suggesting faster rates of pluripotent cell development in these cultures. This is consistent with previous analyses of *Rex-1* and *FGF-5* gene expression during the ES to X cell transition (Bettess 1993, Thomas 1995, Lake 1996) and may be related to the capability of LIF to inhibit the formation of primitive ectoderm *in vitro* (Shen and Leder 1992).

Importantly, *in situ* hybridisation of A03/360 and B04/400 to ES cell cultures demonstrated that marker transcripts were restricted to pluripotent cell types, with no detectable expression in differentiated cells. This indicated that within the ES to X cell transition A03/360 and B04/400 were pluripotent cell specific, and that expression did not result from contaminating differentiated cells.

The restriction of A03/360 and B04/400 to early pluripotent cell populations suggests that the gene products recognised by these markers may have functional roles during pluripotent cell development *in vitro*. Sequence analysis revealed that A03/360 was a novel gene. The sequence of B04/400 demonstrated 44.8% amino acid identity over a 96 residue region to a *C. elegans* predicted open reading frame, CLEB0336.3 10. As B04/400 and CLEB0336.3 10 exhibit no significant similarity to other proteins, this region may represent a novel functional domain in these proteins. Analysis of CLEB0336.3 10 that gives insight into

expression or function during the development of *C. elegans* has not been documented. The isolation of full length cDNAs for A03/360 and B04/400 would potentially give information about the presence of functional protein domains in these transcripts and as a result, information on the molecular activities of these genes. This approach is also likely to yield important tools for studying A03/360 and B04/400 function, such as transcript specific probes, or constructs to perturb expression levels or to enable the production of specific antibodies for protein analysis.

This analysis demonstrated the capacity of DDPCR to identify markers for pluripotent cell subpopulations, based on a cell progression *in vitro*. A screen for differential gene expression detected a series of markers that subdivided the *Oct-4*⁺ pool of cells and identified an ES and early pluripotent cell subpopulation *in vitro*. Expression analysis of A03/360 and B04/400 during early embryogenesis therefore has the potential to identify restricted subpopulations or “states” within the *Oct-4*⁺ pluripotent cell pool *in vivo*. In particular, differential expression between the pluripotent cells of the ICM and primitive ectoderm would be expected. Furthermore, given the ES cell restricted nature of A03/360 and B04/400 expression, *in vivo* analysis has the potential to define the embryonic equivalent of ES cells to a precise window of development. This could direct the development of alternative isolation strategies for ES cells from commercially important species (1.6.1, 3.3). It may also be possible to also correlate the expression of A03/360 and B04/400 with specific developmental events, highlighting potential functions within the pluripotent cells of early development.

The investigation of DDPCR techniques outlined in this chapter therefore demonstrated the capacity to identify and isolate specific markers for pluripotent cell subpopulations *in vitro*. The further application of this approach is likely to generate a comprehensive array of temporally expressed markers within the ES to X cell transition. These markers have the potential to refine the complexity of known pluripotent cell subtypes *in vitro*, the equivalent stem cell populations or “states” *in vivo*, and to clarify the molecular basis of pluripotent cell development during early embryogenesis.

CHAPTER 4

*Expression Analysis of Early Pluripotent Cell Markers
in vitro and During Early Embryogenesis*

CHAPTER 4

EXPRESSION ANALYSIS OF EARLY PLURIPOTENT CELL MARKERS *IN VITRO* AND DURING EARLY EMBRYOGENESIS

4.1 Introduction

Investigation of gene expression in pluripotent cell populations generated from the culture of ES cells enabled the isolation of two novel markers with interesting features. Expression analysis of these markers, termed A03/360 and B04/400, indicated that they were likely to be restricted to pluripotent cell types *in vitro* and more importantly, their expression was specific to early stem cell types within the ES to X cell progression. A03/360 and B04/400 were expressed by ES cells and were downregulated during X cell formation, in a similar manner to *Rex-1* expression (1.6.8). A03/360 and B04/400 therefore distinguished an *Oct-4*⁺ “early” X cell type that retained high levels of expression, from a *Oct-4*⁺ “late” X cell type with low transcript levels. These expression profiles suggested that A03/360 and B04/400 were expressed in pluripotent cell populations during embryogenesis and may be capable of identifying transient stem cell subtypes or “states” in the early murine embryo.

Clarification of A03/360 and B04/400 expression *in vitro* would allow more detailed analysis of pluripotent cell development. Firstly, it may clarify the transition between “early” and “late” X cell types, potentially subdividing X cells into timed subpopulations. Secondly, this analysis could enable the comparison of pluripotent cell types found in different culture regimes.

The definition of A03/360 and B04/400 expression patterns during early embryogenesis was important for several reasons. Marker expression would be expected at the ICM stage of development, with downregulation of expression occurring during the formation of primitive ectoderm. This analysis may identify subpopulations that express A03/360 and B04/400 within the *Oct-4*⁺ pool of pluripotent cells. Identification of transient pluripotent cell subpopulations may enable more precise definitions of the embryonic equivalents of ES and X cell types. The regulation of marker expression may also be correlated with critical developmental events, highlighting potential functional roles for A03/360 or B04/400 and

demonstrating the potential of using these markers to probe functional aspects of pluripotent cell development (1.7).

This chapter describes the analysis of A03/360 and B04/400 expression by *in situ* hybridisation. Analysis of the ES to X cell transition and ES cell differentiation demonstrated that A03/360 and B04/400 were pluripotent cell specific and indicated that they may be coordinately regulated *in vitro*. Investigation of expression during murine embryogenesis revealed that A03/360 and B04/400 were expressed in overlapping but distinct pluripotent cell populations. A03/360 was expressed in pluripotent cells during preimplantation development and was renamed *Icm1*, for Inner cell mass 1. B04/400 expression was induced in the late stage ICM and was expressed in the early postimplantation epiblast, prior to the formation of the proamniotic cavity. B04/400 was therefore renamed *Psc1*, for Peri-implantation stem cell 1. Expression of both *Icm1* (A03/360) and *Psc1* (B04/400) was downregulated by the primitive ectoderm stage of embryogenesis. Expression of *Psc1* was also detected in the extraembryonic ectoplacental cone.

4.2 Expression of A03/360 and B04/400 *in vitro*

Direct visualisation of the cell types that express A03/360 and B04/400 *in vitro* would enable a more detailed analysis of pluripotent cell development and comparison of pluripotent cell subpopulations. To ascertain the distribution, regulation and comparison of A03/360 and B04/400 expression, *in situ* hybridisation analysis was conducted on induced ES cell cultures, cell monolayers representative of different stages during the ES to X cell transition, and reverted X cells (1.6.8). To assist with the comparison of marker expression levels between different cultures, riboprobe batch and exposure times for individual probes were consistent throughout this analysis. Detailed comparison of expression levels between cell types was complicated by differences in colony morphology, which varied from domed and multilayered "ES cell" like colonies, to monolayer growth exhibited by X cells and differentiated cell types. A higher intensity of staining in domed colonies, compared to monolayer colonies, can be misleading due to gene expression in a three dimensional array of cells.

4.2.1 *In situ* Hybridisation Analysis of A03/360 and B04/400 Expression During ES Cell Differentiation

D3 ES cell cultures were differentiated by LIF withdrawal (SP, 2.4.2ii), or by induction with retinoic acid (RA, 2.4.2iii), 3-methoxybenzamide (MBA, 2.4.2iv) or dimethylsulfoxide (DMSO, 2.4.2v), as described in section 1.6.5.

DIG labelled antisense riboprobes from the 5' (256 bp) A03/360 (2.2.7i), the B04/400 (2.2.7ii) and *Oct-4* (2.2.7iii) cDNA clones were used to detect expression in SP, RA, MBA and DMSO induced ES cell cultures. The pluripotent cell specific marker *Oct-4* (1.5.2) was used to monitor the distribution and morphology of pluripotent cell types and residual pluripotent cell niches present in these cultures (Fig. 4.1). The morphology of these pluripotent cell niches varied from rounded and "ES cell like" colonies (not shown), to monolayer colonies of small cells with structurally characteristic nuclei (purple staining, Fig. 4.1A-D). Monolayer pluripotent cells were more frequent in MBA and DMSO induced cultures. Examples of *Oct-4*⁺ cells that differed from "ES like" colonies are indicated in figure 4.1. Differentiated cell types were morphologically distinct from pluripotent cell types and did not express *Oct-4* (Fig. 4.1 A-D).

A03/360 and B04/400 exhibited similar expression patterns during this analysis (Figs. 4.2, 4.3 respectively). The expression of A03/360 or B04/400 was not detected in any differentiated cell types (*Oct-4*⁻ cells, Fig. 4.1) resulting from any of the ES cell induction regimes (Figs. 4.2, 4.3). This confirmed that A03/360 and B04/400 expression was downregulated during cellular differentiation. A03/360 and B04/400 expression was tightly restricted to nests of residual pluripotent cells, often present as colony niches surrounded by differentiated cell types (purple staining, Figs. 4.2, 4.3 respectively). A03/360 and B04/400 were expressed by both "ES cell like" (Figs. 4.2A, B. Fig. 4.3B respectively) and monolayer (Fig 4.2C, D. Figs. 4.3A, C, D respectively) colonies of residual pluripotent cells recognised by *Oct-4* expression (Fig. 4.1). The expression of A03/360 and B04/400 in two morphologically distinct pluripotent cell colonies indicates that these cells are closely related. No staining was observed in the same batch of induced ES cell cultures probed with A03/360, B04/400 or *Oct-4* sense control riboprobes (data not shown).

Figure 4.1 *In situ* Hybridisation Analysis of *Oct-4* Expression During ES Cell Differentiation.

Detection and localisation of *Oct-4* expression in induced embryonic stem cell cultures was carried out by *in situ* hybridisation (2.3.30). D3 ES cells were differentiated by: (A) SP, spontaneous (-LIF) induction (2.4.2ii); (B) RA, retinoic acid induction (2.4.2iii); (C) MBA, 3-methoxybenzamide induction (2.4.2iv); (D) DMSO, dimethylsulfoxide induction (2.4.2v). *Oct-4* expression was detected using a 394 bp DIG labelled antisense riboprobe (2.2.7iii) and indicated the prevalence and morphology of residual pluripotent cell types in differentiated ES cell cultures. The *Oct-4* plates were developed for 2 hours and 30 min. Cells that expressed *Oct-4* were indicated by purple staining. Photographs were taken under phase contrast (1) and bright field optics (2) at 100 X magnification. *In situ* hybridisation of induced ES cell cultures with an *Oct-4* sense control probe (2.2.7iii) did not exhibit any stain (not shown).

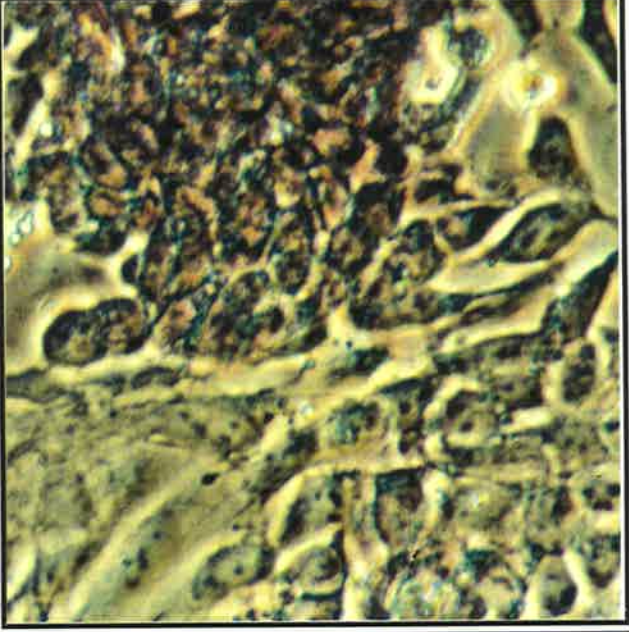
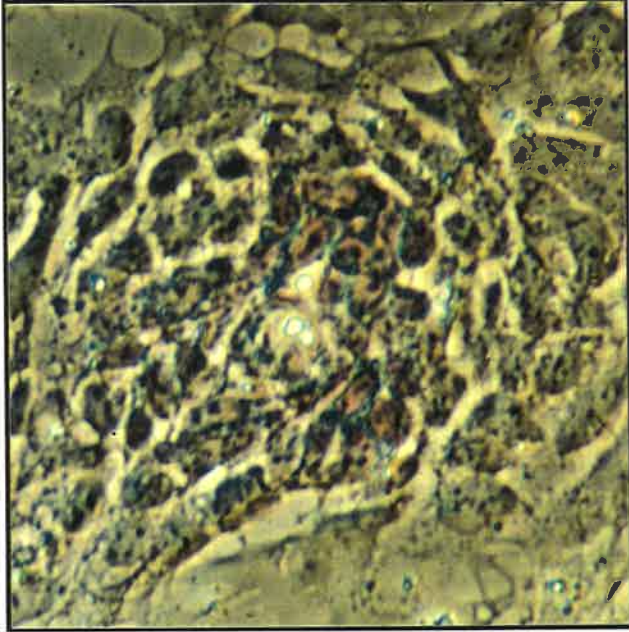
- R residual pluripotent cell types
- D differentiated cell types

(A). SP

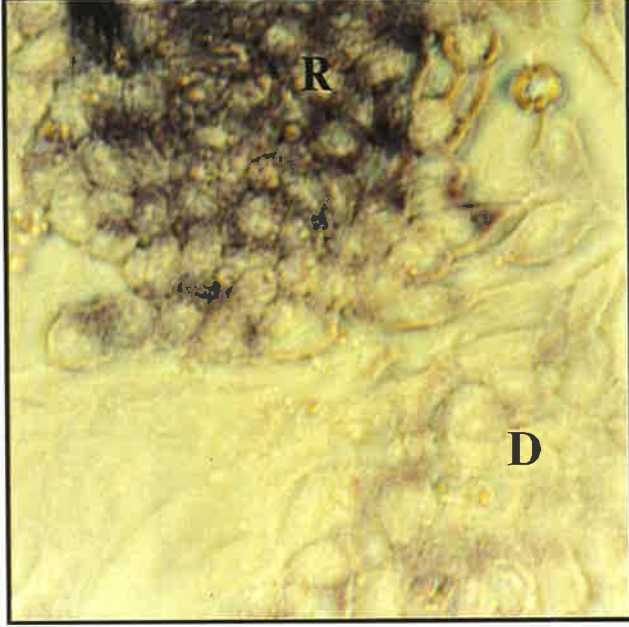
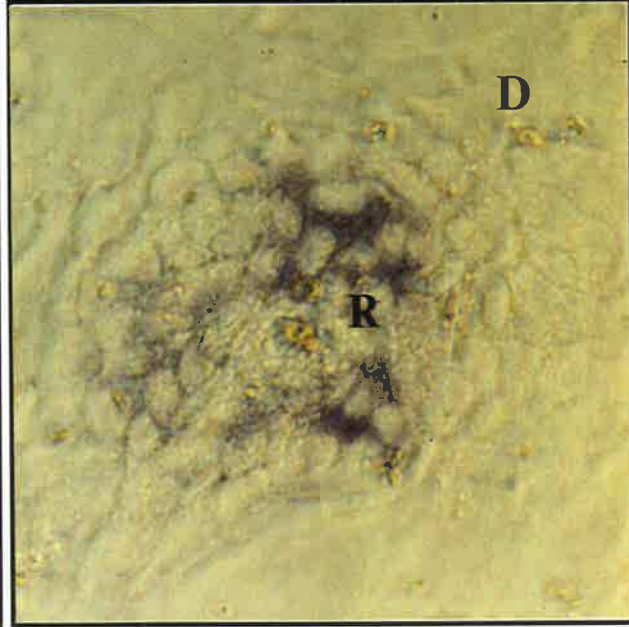
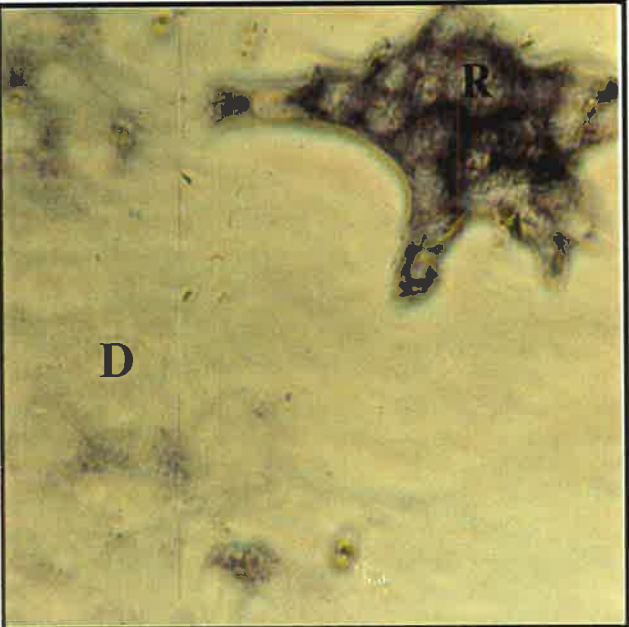
(B). RA

(C). MBA

1

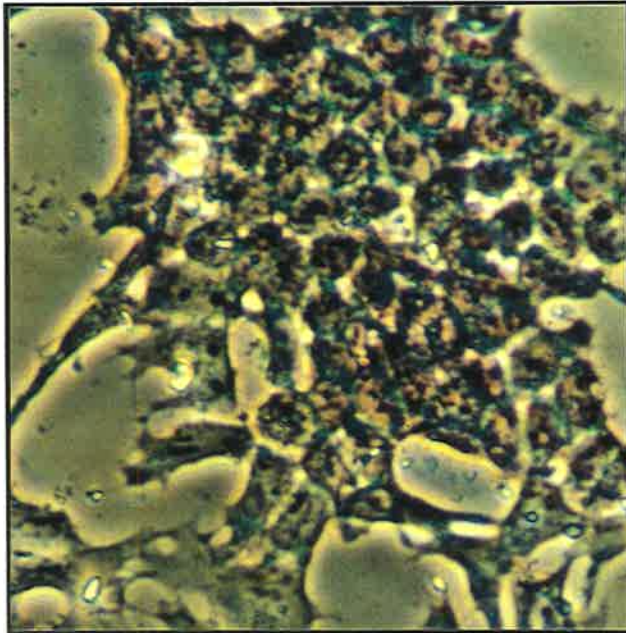


2



(D). DMSO

1



2

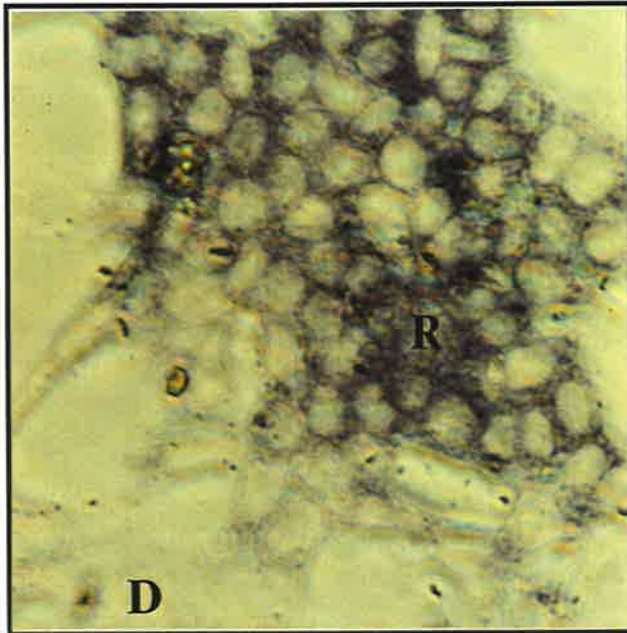


Figure 4.2 *In situ* Hybridisation Analysis of *Icm1* (A03/360) Expression During ES Cell Differentiation.

Detection and localisation of *Icm1* (A03/360) expression in induced embryonic stem cell cultures was carried out by *in situ* hybridisation (2.3.30). D3 ES cells were differentiated by: (A) SP, spontaneous (-LIF) induction (2.4.2ii); (B) RA, retinoic acid induction (2.4.2iii); (C) MBA, 3-methoxybenzamide induction (2.4.2iv); (D) DMSO, dimethylsulfoxide induction (2.4.2v). *Icm1* expression was detected using a DIG labelled antisense riboprobe of the 256 bp 5' A03/360 clone (2.2.7i). The *Icm1* plates were developed for 9 hours. Cells that expressed *Icm1* were indicated by purple staining. The light staining cells in (D) were potentially residual pluripotent cells undergoing differentiation. Photographs were taken under phase contrast (1) and bright field optics (2) at 100 X magnification. *In situ* hybridisation of induced ES cell cultures with an *Icm1* sense control probe (2.2.7i) did not exhibit any stain (not shown).

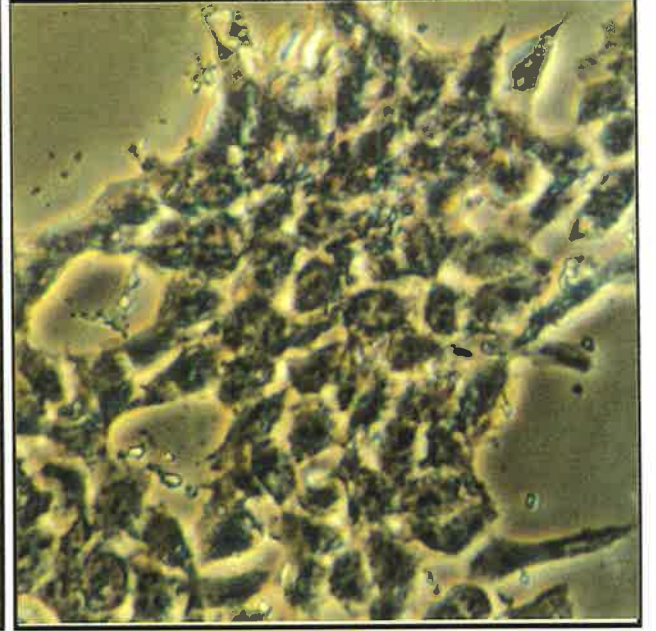
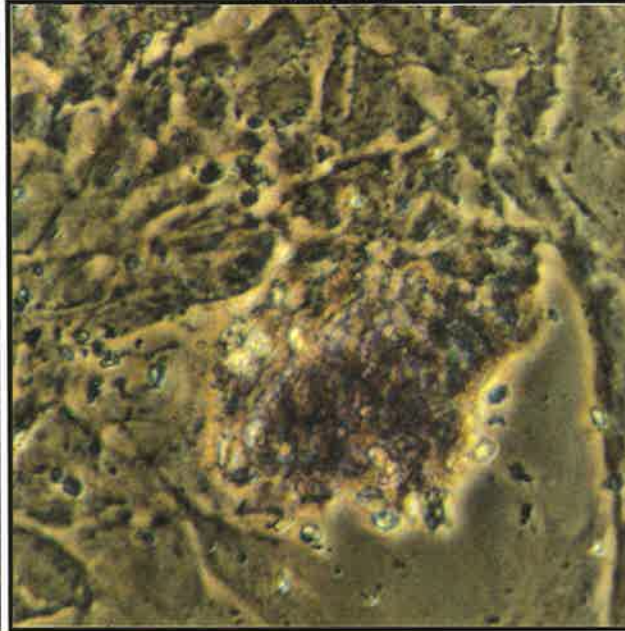
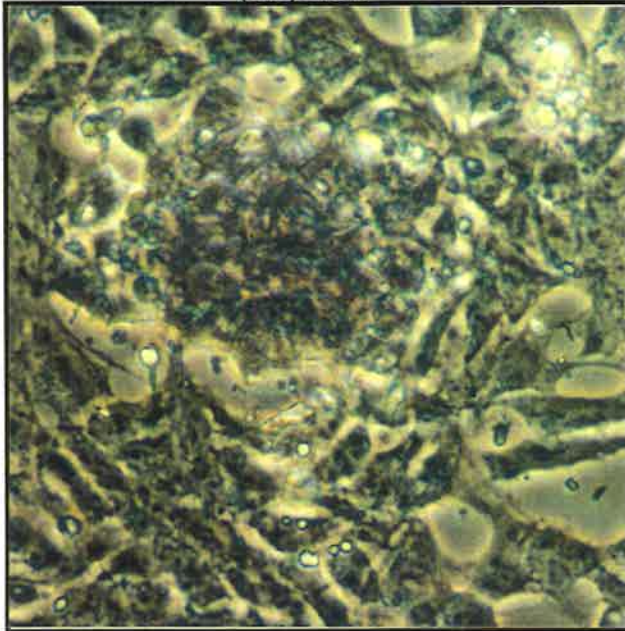
- R residual pluripotent cell types
- D differentiated cell types

(A). SP

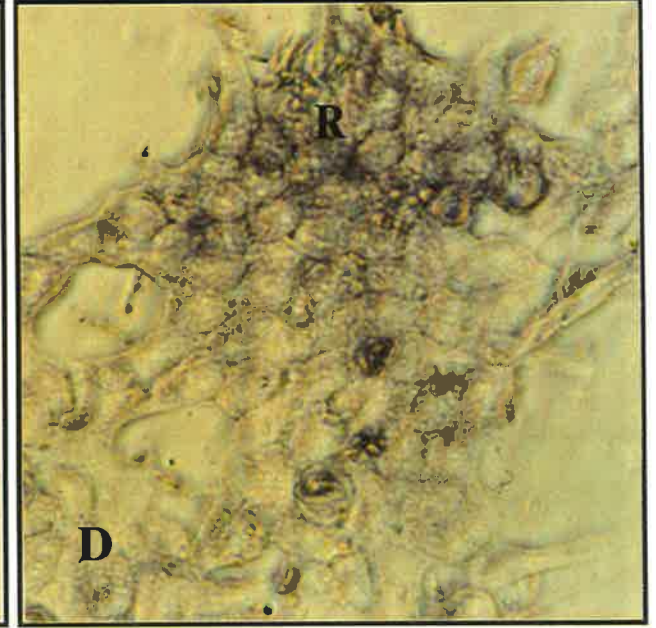
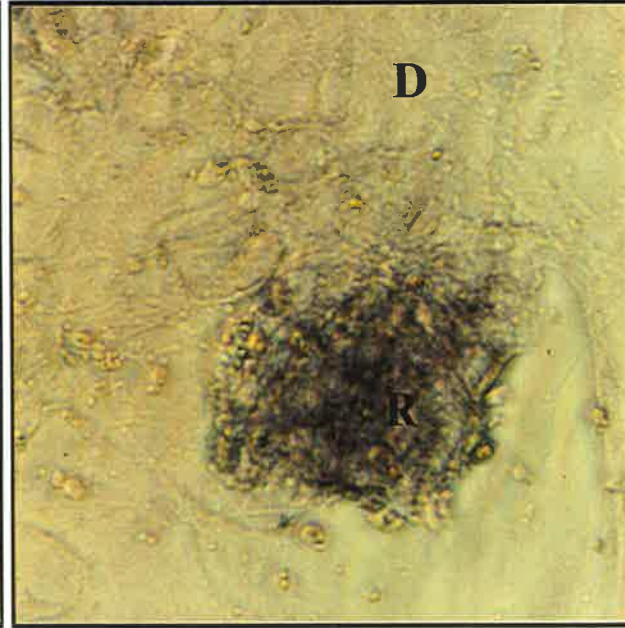
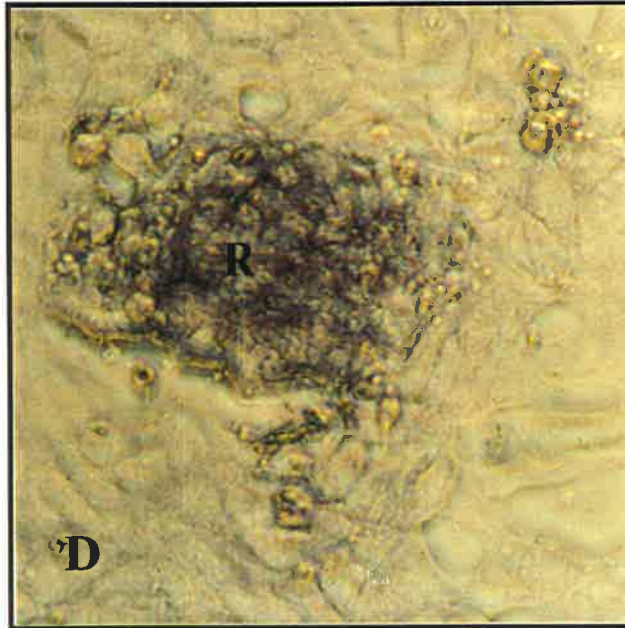
(B). RA

(C). MBA

1

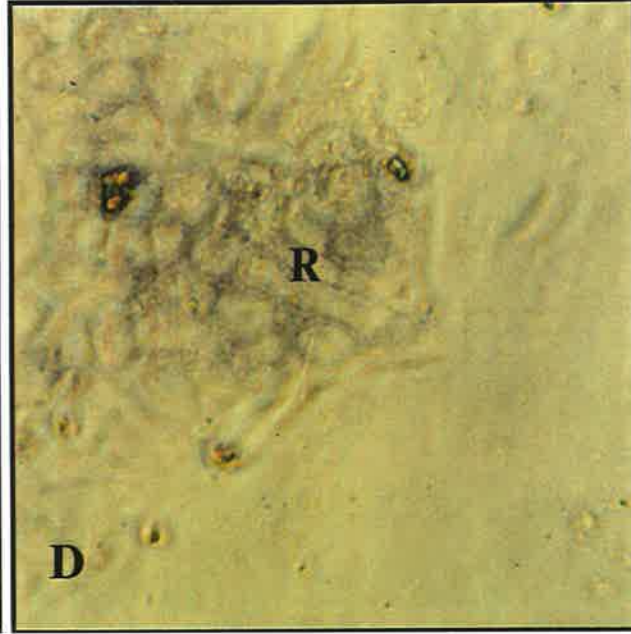
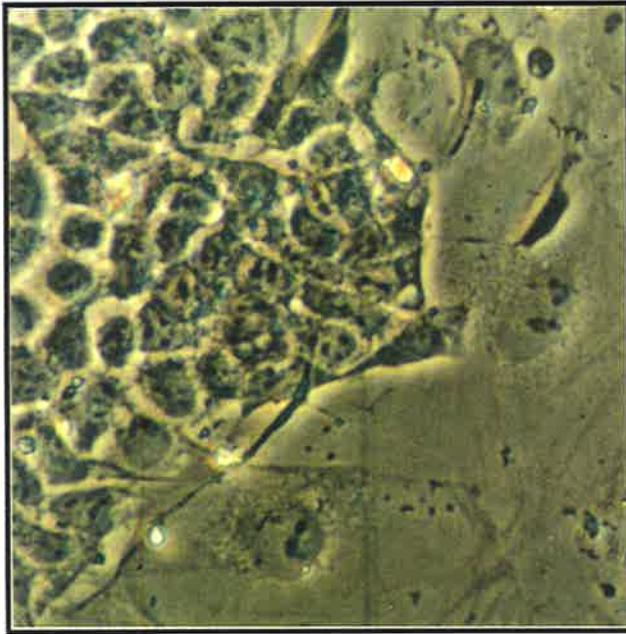


2



1

2



(D). DMSO

Figure 4.3 *In situ* Hybridisation Analysis of *Psc1* (B04/400) Expression During ES Cell Differentiation.

Detection and localisation of *Psc1* (B04/400) expression in induced embryonic stem cell cultures was carried out by *in situ* hybridisation (2.3.30). D3 ES cells were differentiated by: (A) SP, spontaneous (-LIF) induction (2.4.2ii); (B) RA, retinoic acid induction (2.4.2iii); (C) MBA, 3-methoxybenzamide induction (2.4.2iv); (D) DMSO, dimethylsulfoxide induction (2.4.2v). *Psc1* expression was detected using a DIG labelled antisense riboprobe of the 458 bp B04/400 clone (2.2.7ii). The *Psc1* plates were developed for 9 hours. Cells that expressed *Psc1* were indicated by purple staining. Photographs were taken under phase contrast (1) and bright field optics (2) at 100 X magnification. *In situ* hybridisation of induced ES cell cultures with an *Psc1* sense control probe (2.2.7ii) did not exhibit any stain (not shown).

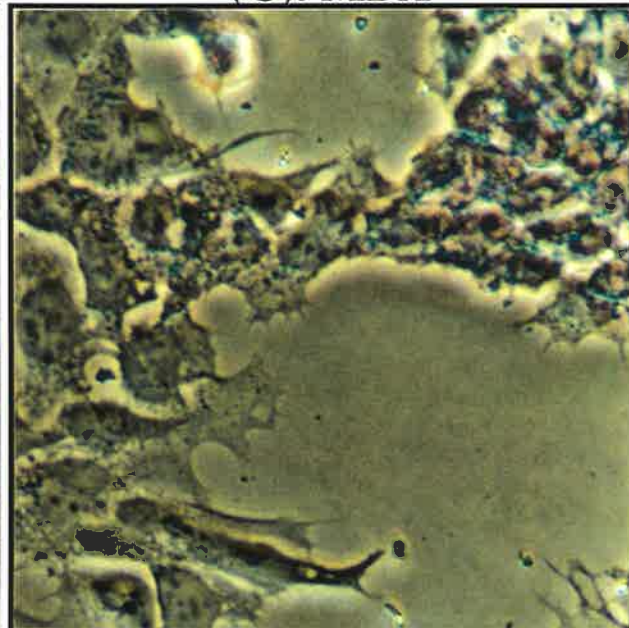
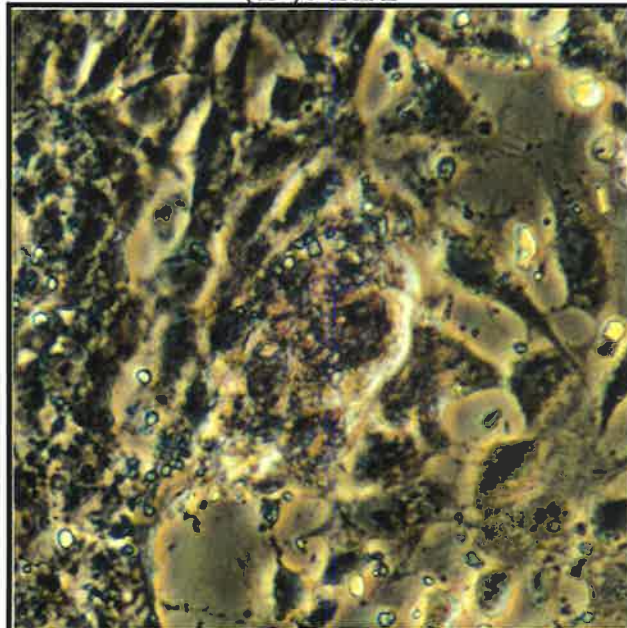
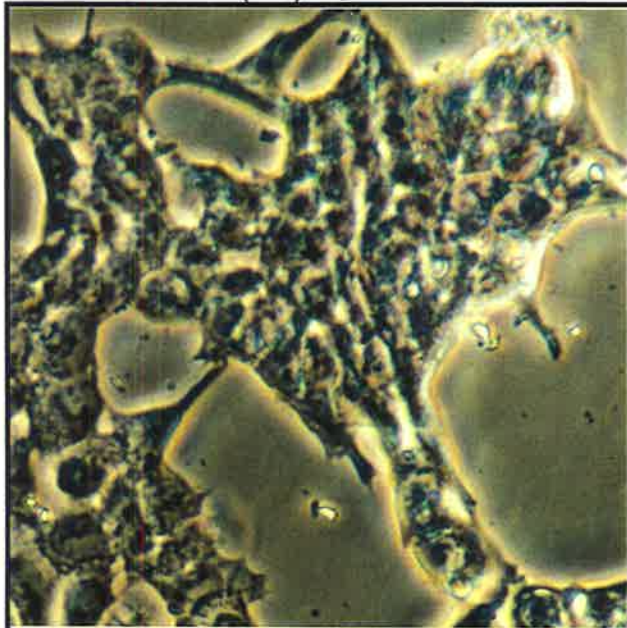
- R residual pluripotent cell types
- D differentiated cell types

(A). SP

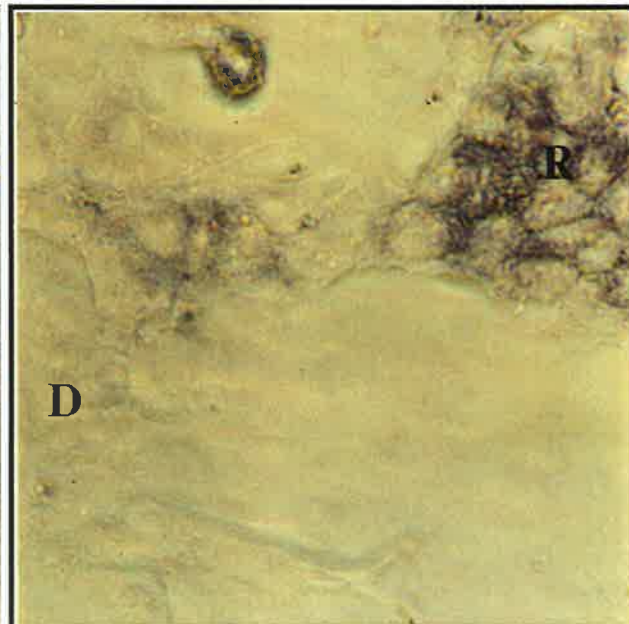
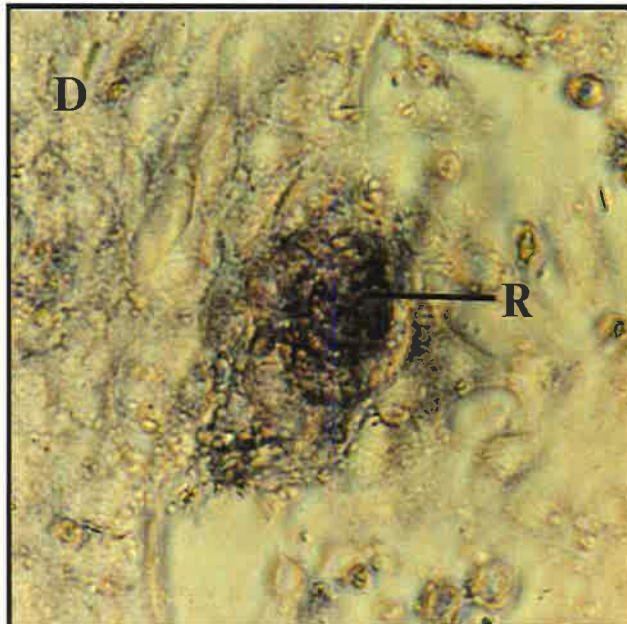
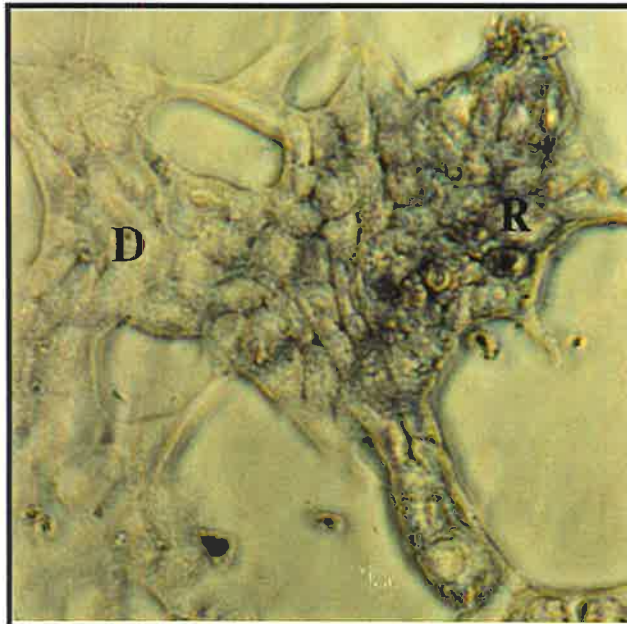
(B). RA

(C). MBA

1

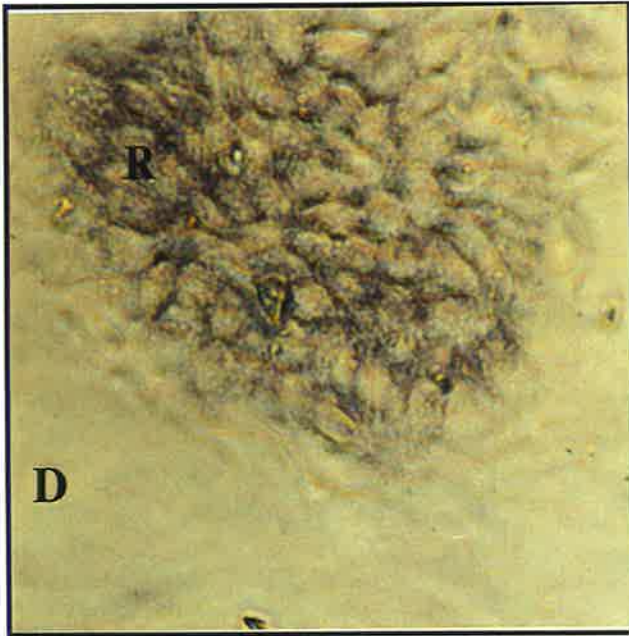
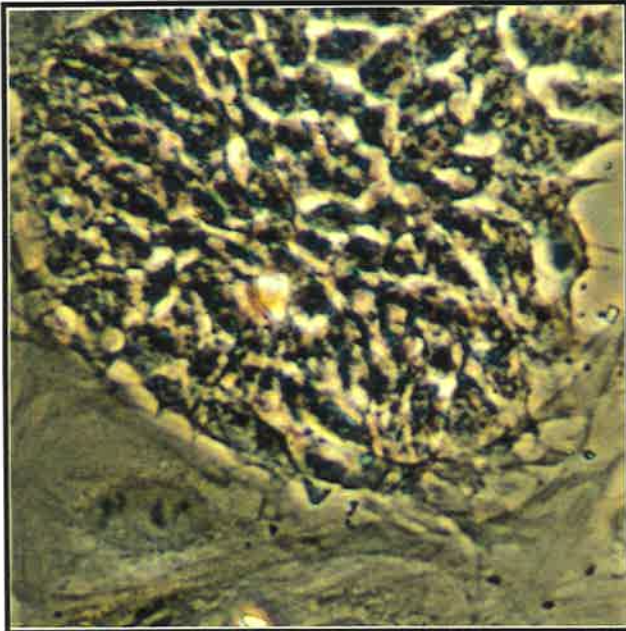


2



1

2



(D). DMSO

In situ hybridisation analysis during ES cell differentiation therefore demonstrated that A03/360 and B04/400 expression was downregulated during cellular differentiation and the transcripts recognised by these markers were specifically restricted to pluripotent cell populations *in vitro*.

4.2.2 *In situ* Hybridisation Analysis of A03/360 and B04/400 Expression During the ES to X Cell Transition

In situ hybridisation analysis of *Oct-4*, A03/360 and B04/400 expression during the development of ES cells to X cells and in reverted X cells (1.6.8) was carried out using the same probes as above (4.2.1) and is indicated in figures 4.4, 4.5 and 4.6. Cell cultures were generated as described in section 2.4.2i.

Morphology and the expression of *Oct-4* in the ES to X cell transition demonstrated the presence and morphology of pluripotent cells in these cultures (Fig. 4.4A-E). The transcripts identified by A03/360 and B04/400 appeared to be regulated in a cognate manner during this transition. A03/360 and B04/400 were expressed by D3 ES cells (Figs. 4.5A, 4.6A respectively) and were detected in X+0 cells (Figs. 4.5B, 4.6B) and at a low level in X+1 cells (Figs. 4.5C, 4.6C) when D3 ES cells were passaged in MedII and LIF. Expression was downregulated in X+2 cells (Figs. 4.5D, 4.6D) and was not detected in X+3 cells (Figs. 4.5E, 4.6E). A03/360 and B04/400 expression was not detected in spontaneously differentiated cell types present in X cell cultures, which may differ from those present in ES cell cultures (Lake 1996).

Interconversion between pluripotent cell “states” is predicted to be a functional characteristic of pluripotent cell types (Rossant 1993). X cell reversion generates a pluripotent cell type that is similar ES cells on the basis of morphology, marker gene expression, *in vitro* differentiation potential and contribution to chimera formation (1.6.8). The expression of A03/360 and B04/400 was rapidly upregulated when X+3 cells were reverted to an ES cell phenotype by culture in LIF without MedII. Reverted X cells with ES cell colony morphology expressed *Oct-4* (Fig. 4.4F), A03/360 and B04/400 at comparable levels to ES cells. A03/360 and B04/400 expression was specifically restricted to reverted colonies with “ES cell like” domed morphology (X+3/Rp0, data not shown) and was significant in X+3

Figure 4.4 *In situ* Hybridisation Analysis of *Oct-4* Expression During the ES to X Cell Transition.

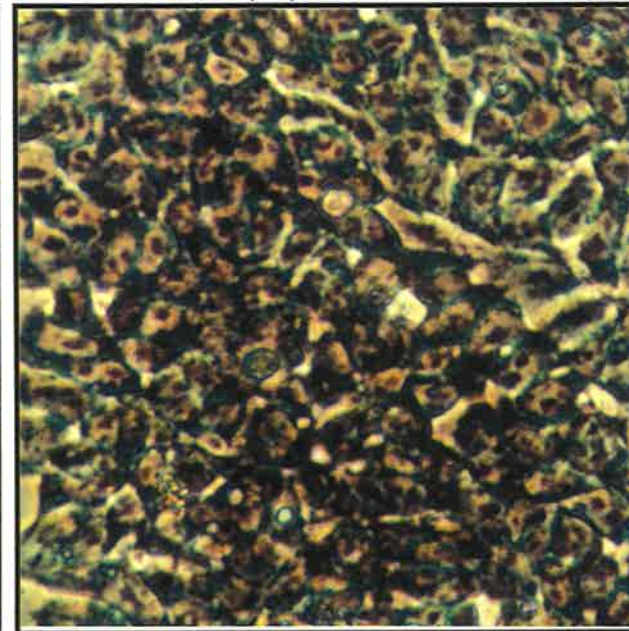
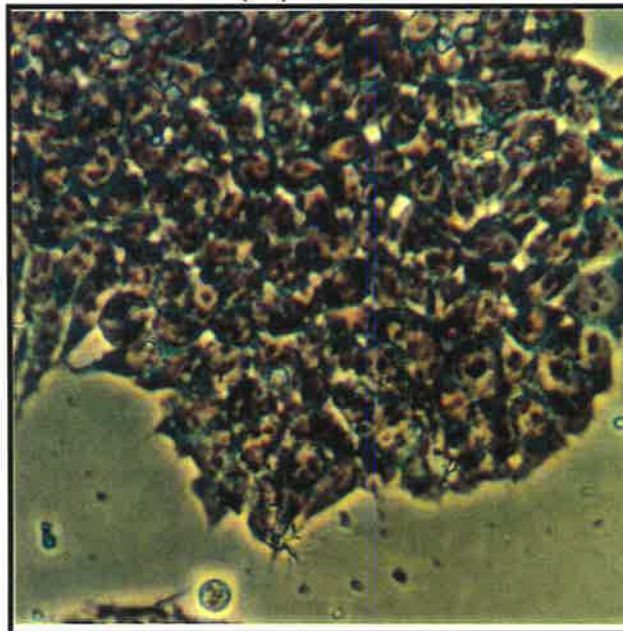
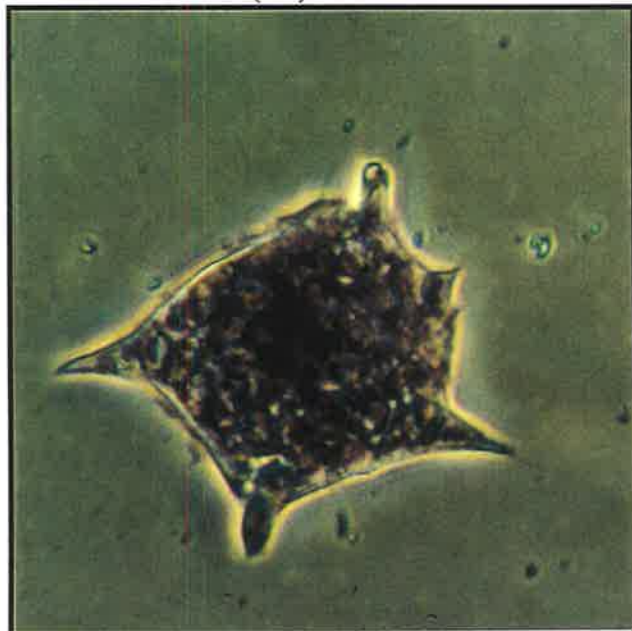
Detection and localisation of *Oct-4* expression during the ES to X cell transition was carried out by *in situ* hybridisation (2.3.30). A sequential series of D3 ES cells, X+0 cells (X cells +LIF / 0 passages), X+1 cells (X cells +LIF / 1 passage), X+2 cells (X cells +LIF / 2 passages), X+3 (X cells +LIF / 3 passages) and X+3/R1 cells (X cells +LIF / 3 passages, reverted +LIF alone / 1 passage) were cultured as described in section 2.4.1 and 2.4.2. *Oct-4* expression was detected using a 394 bp DIG labelled antisense riboprobe (2.2.7iii). The *Oct-4* plates were developed for 1 hour and 30 min. Cells that expressed *Oct-4* were indicated by purple staining. Photographs were taken under phase contrast (1) and bright field optics (2) at 100 X magnification.

(A). ES

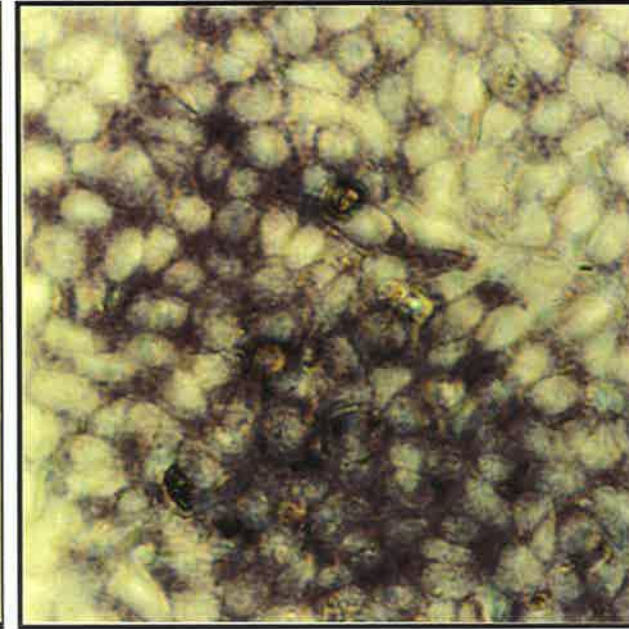
(B). X+0

(C). X+1

1



2

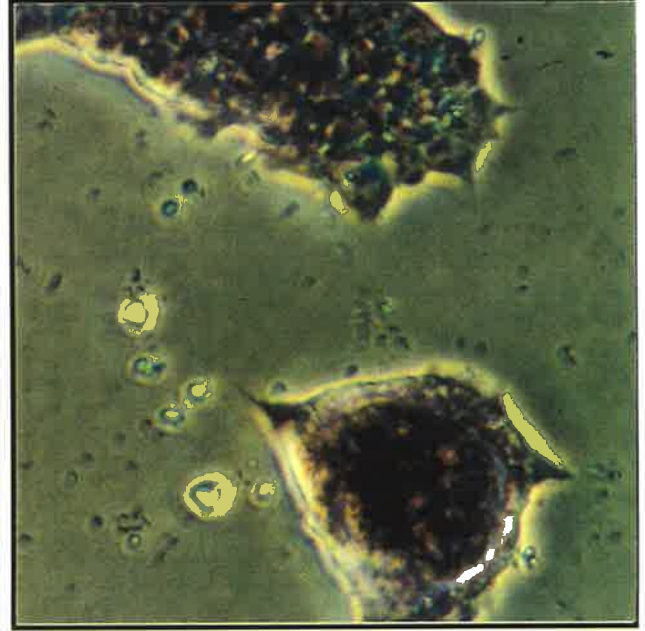
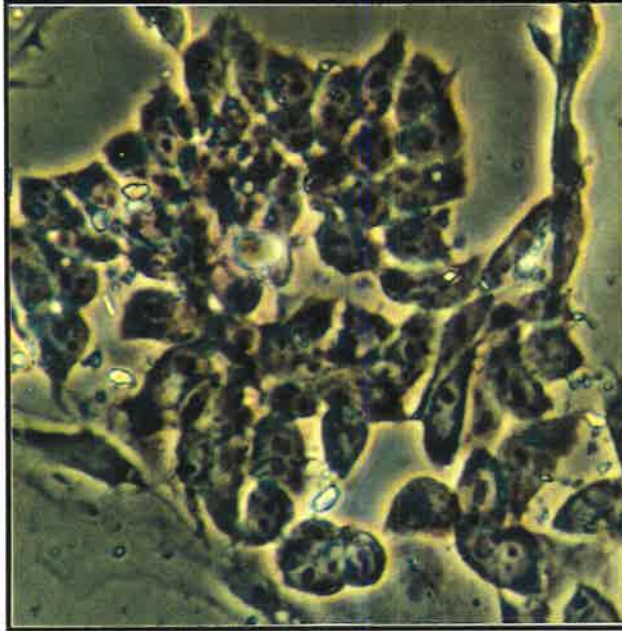
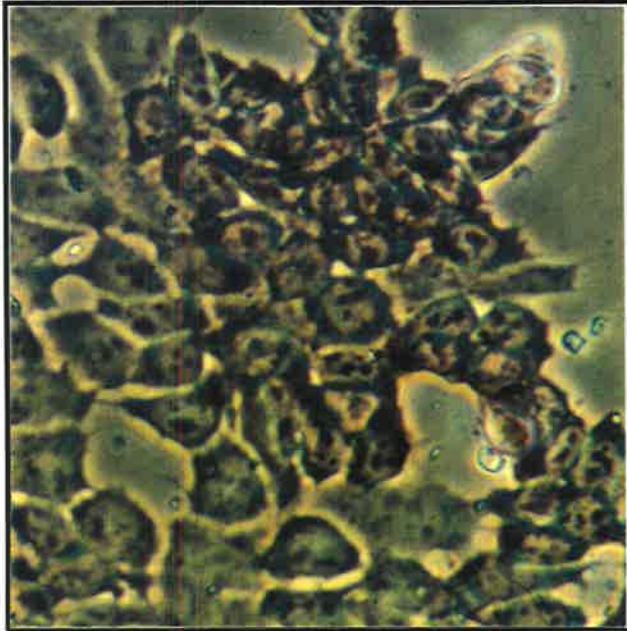


(D). X+2

(E). X+3

(F). X+3/R1

1



2

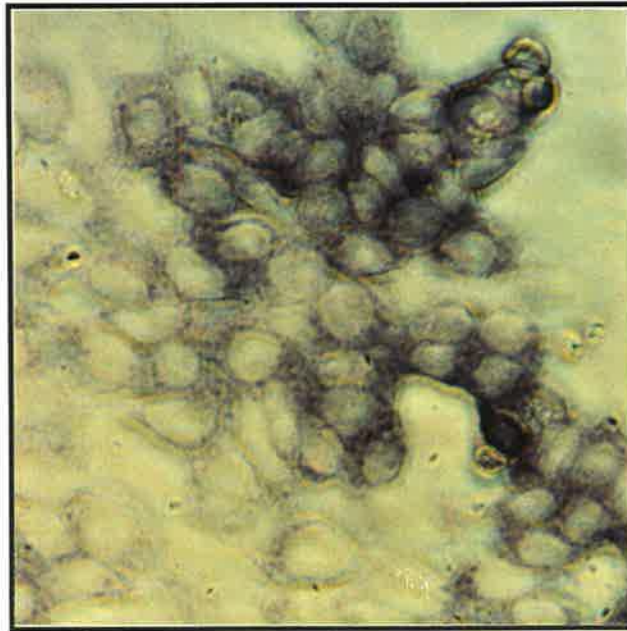


Figure 4.5 *In situ* Hybridisation Analysis of *Icm1* (A03/360) Expression During the ES to X Cell Transition.

Detection and localisation of *Icm1* (A03/360) expression during the ES to X cell transition was carried out by *in situ* hybridisation (2.3.30). A sequential series of D3 ES cells, X+0 cells (X cells +LIF / 0 passages), X+1 cells (X cells +LIF / 1 passage), X+2 cells (X cells +LIF / 2 passages), X+3 (X cells +LIF / 3 passages) and X+3/R1 cells (X cells +LIF / 3 passages, reverted +LIF alone / 1 passage) were cultured as described in section 2.4.1 and 2.4.2. *Icm1* expression was detected using a DIG labelled antisense riboprobe of the 256 bp 5' A03/360 clone (2.2.7i). The *Icm1* plates were developed for 8 hours and 30 min. Cells that expressed *Icm1* were indicated by purple staining. Photographs were taken under phase contrast (1) and bright field optics (2) at 100 X magnification.

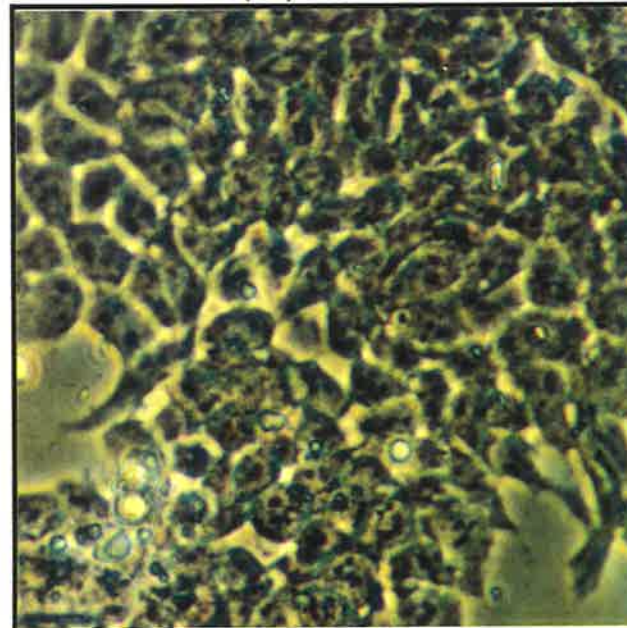
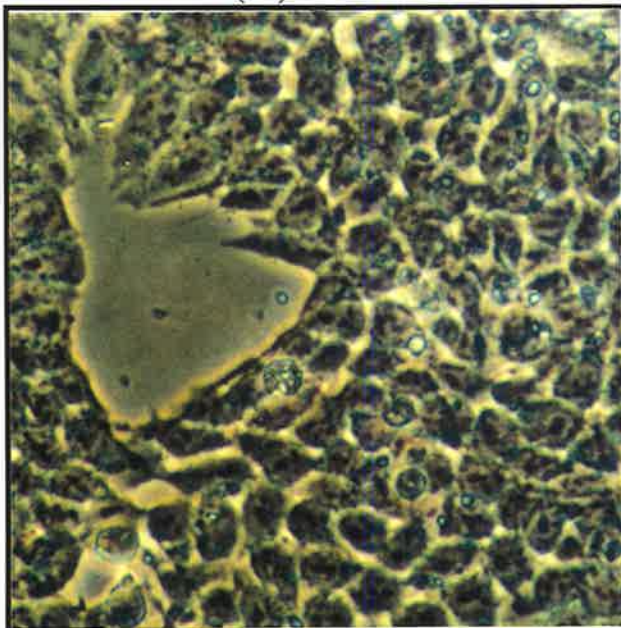
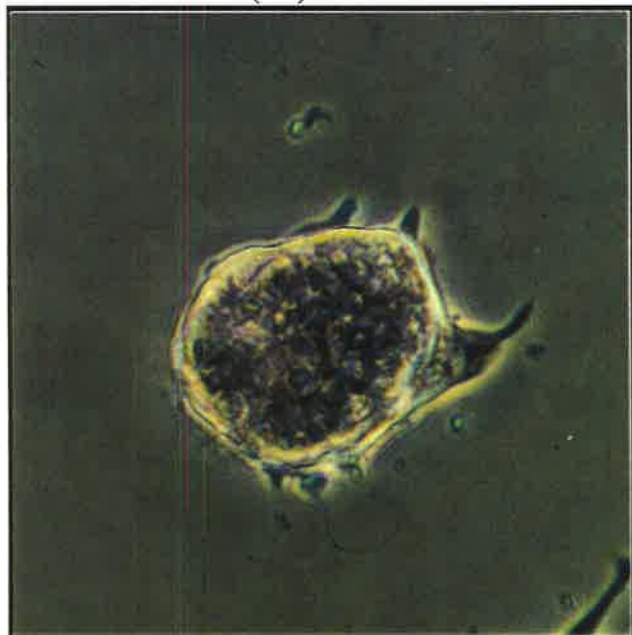
X *Oct-4*⁺ (Fig. 4.4)/*Icm1*⁻ “late” X cell types in X+2, X+3 cultures

(A). ES

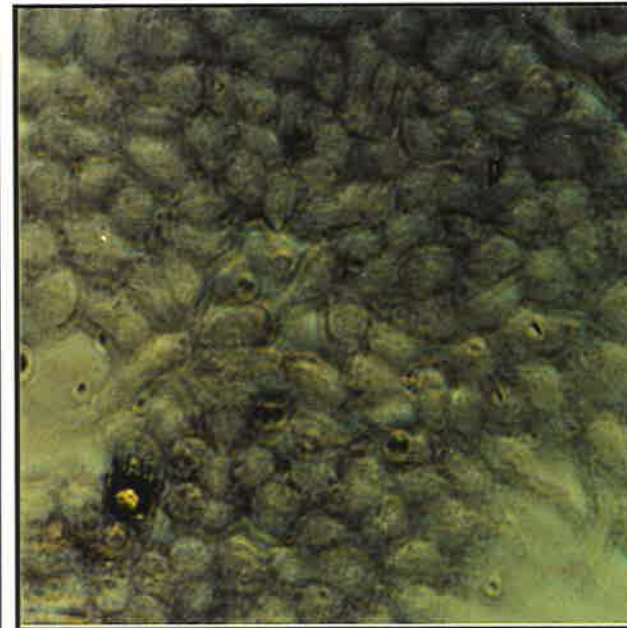
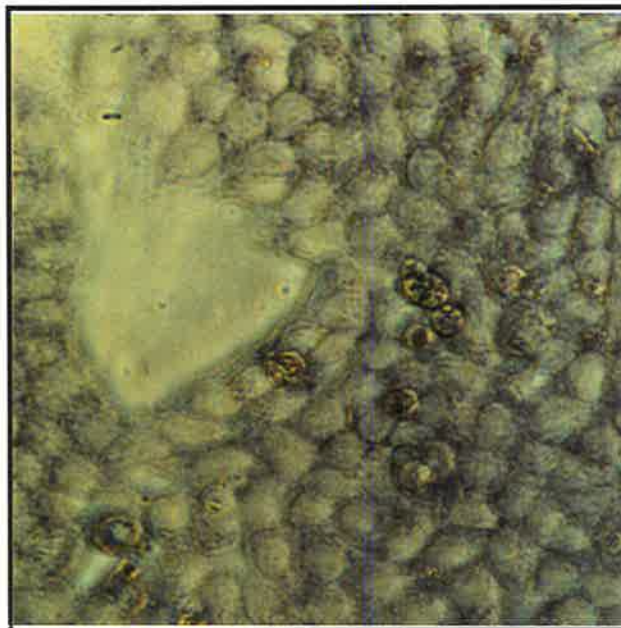
(B). X+0

(C). X+1

1



2

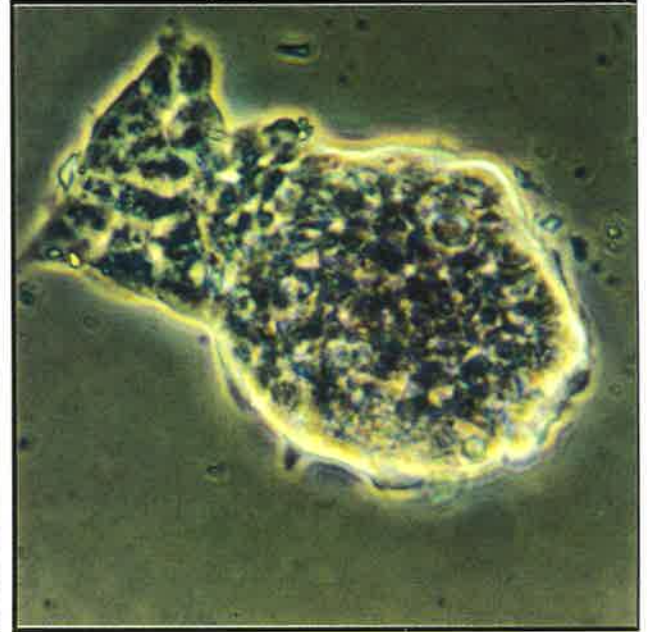
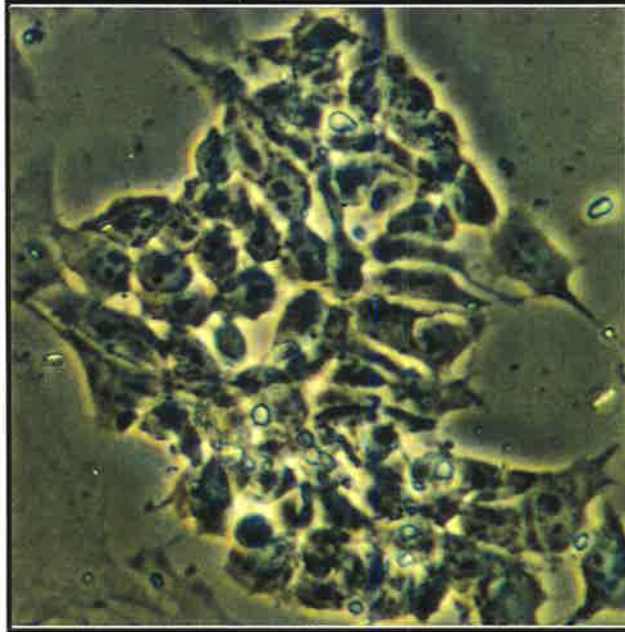
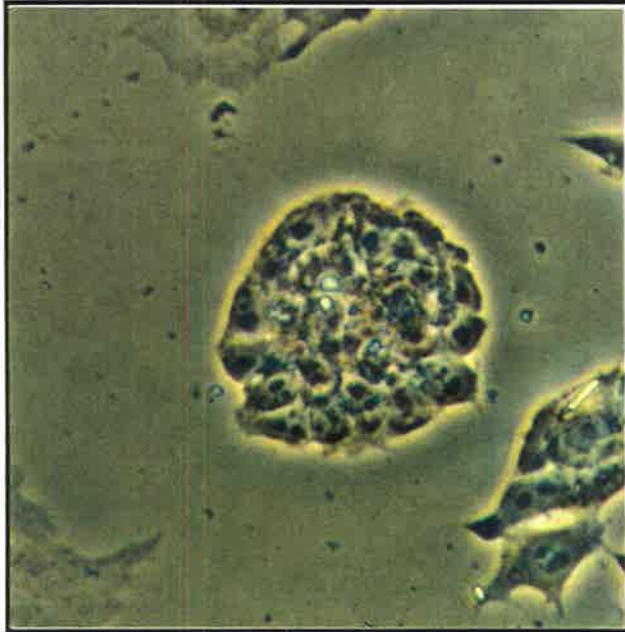


(D). X+2

(E). X+3

(F). X+3/R1

1



2

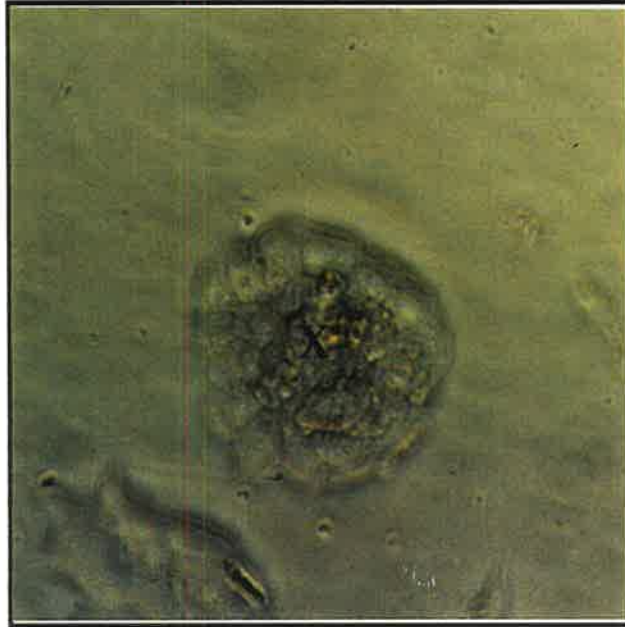


Figure 4.6 *In situ* Hybridisation Analysis of *Psc1* (B04/400) Expression During the ES to X Cell Transition.

Detection and localisation of B04/400 (*Psc1*) expression during the ES to X cell transition was carried out by *in situ* hybridisation (2.3.30). A sequential series of D3 ES cells, X+0 cells (X cells +LIF / 0 passages), X+1 cells (X cells +LIF / 1 passage), X+2 cells (X cells +LIF / 2 passages), X+3 (X cells +LIF / 3 passages) and X+3/R1 cells (X cells +LIF / 3 passages, reverted +LIF alone / 1 passage) were cultured as described in section 2.4.1 and 2.4.2. *Psc1* expression was detected using a DIG labelled antisense riboprobe of the 458 bp B04/400 clone (2.2.7ii). The *Psc1* plates were developed for 8 hours and 30 min. Cells that expressed *Psc1* were indicated by purple staining. Photographs were taken under phase contrast (1) and bright field optics (2) at 100 X magnification. *In situ* hybridisation of X+3 cells with an *Psc1* sense control probe (2.2.7ii) did not exhibit any stain (not shown).

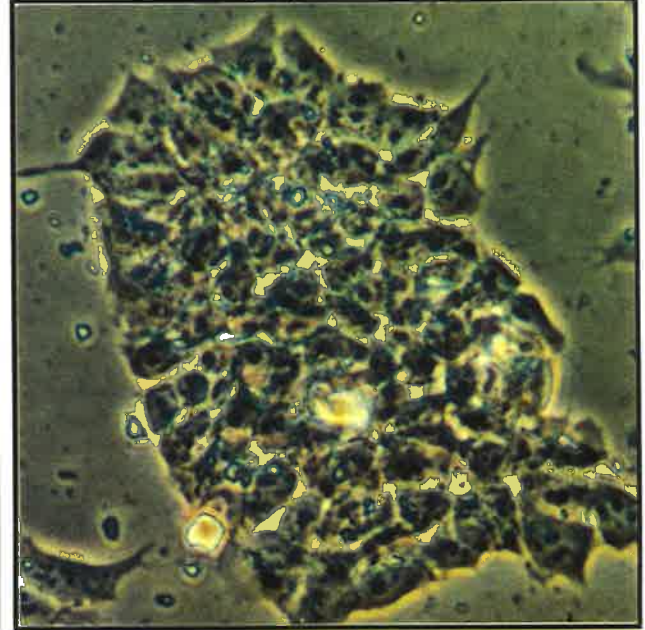
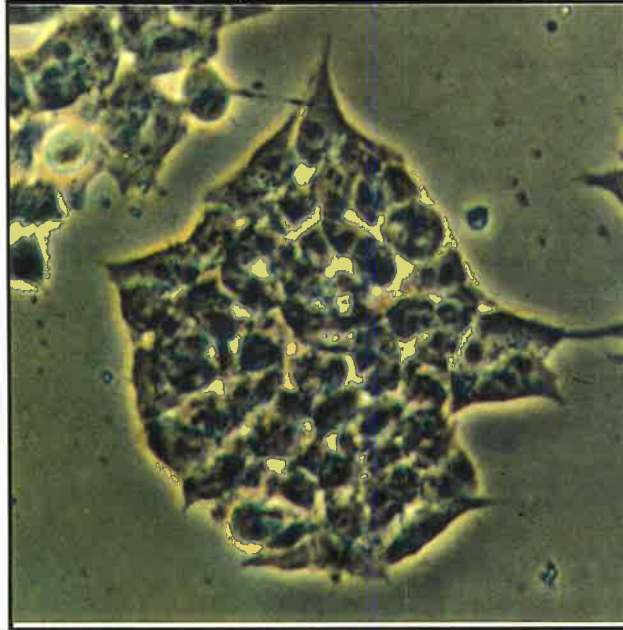
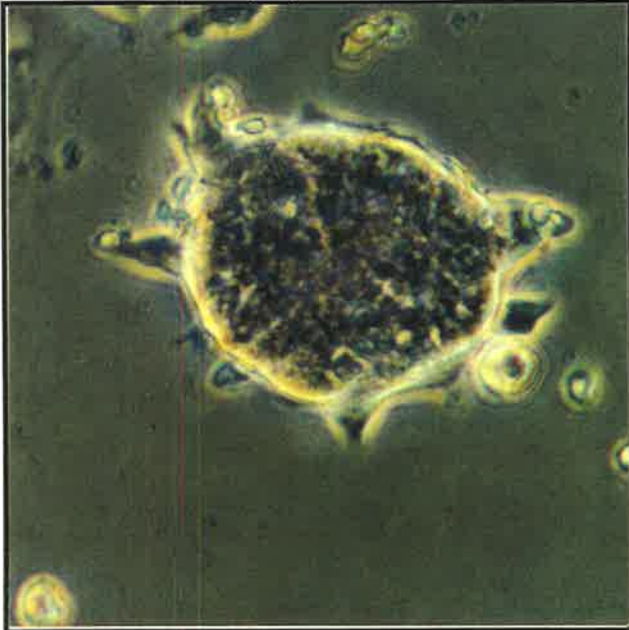
X *Oct-4*⁺ (Fig. 4.4)/*Psc1*⁻ “late” X cell types in X+2, X+3 cultures

(A). ES

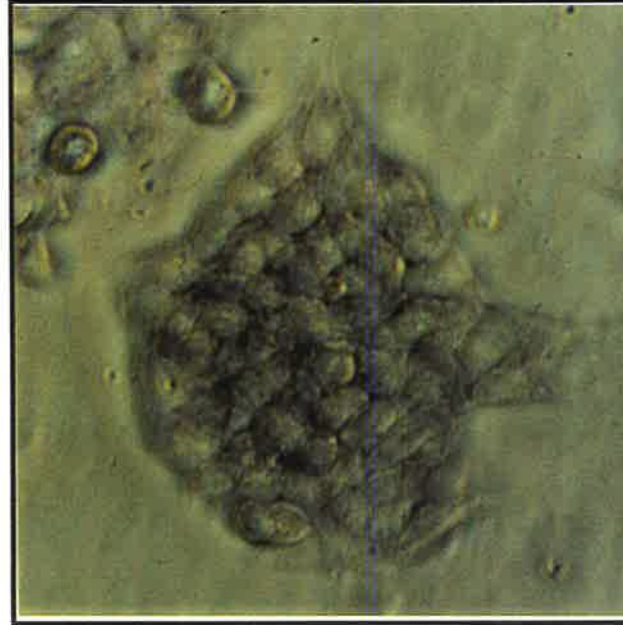
(B). X+0

(C). X+1

1



2

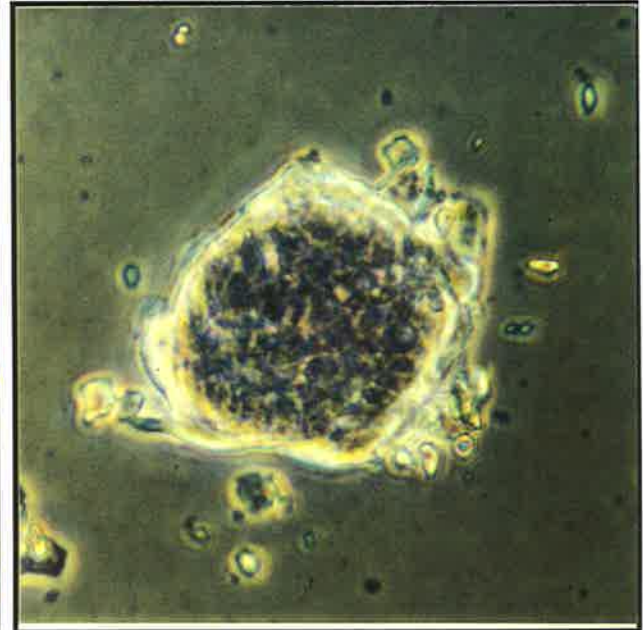
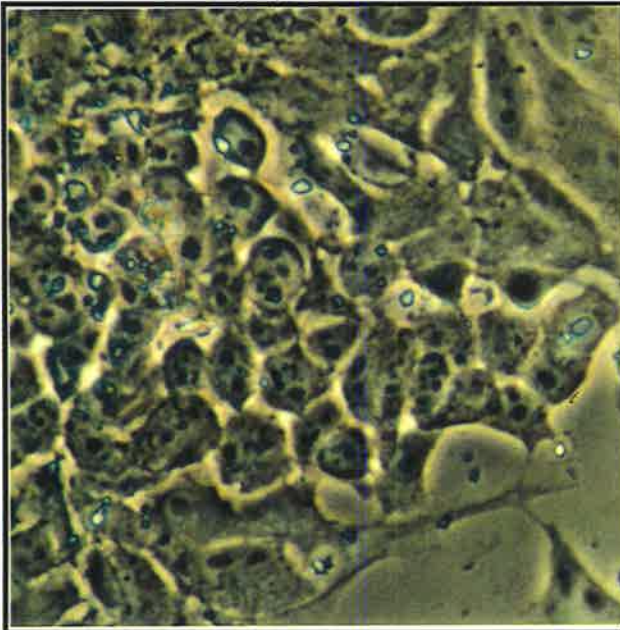
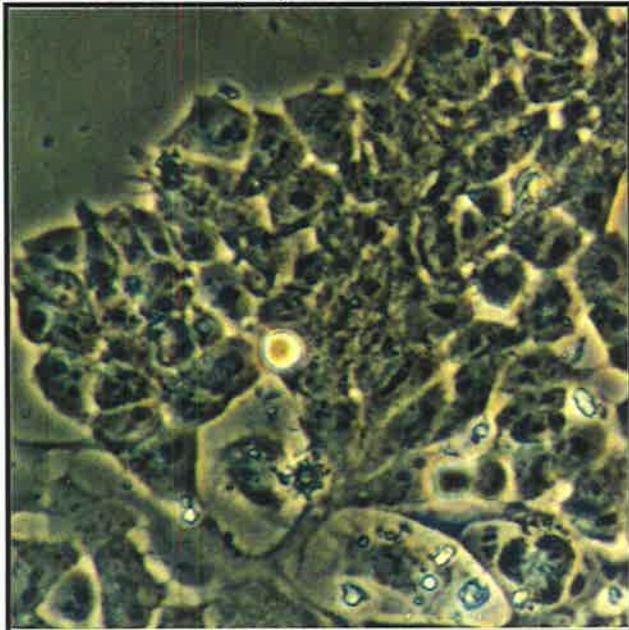


(D). X+2

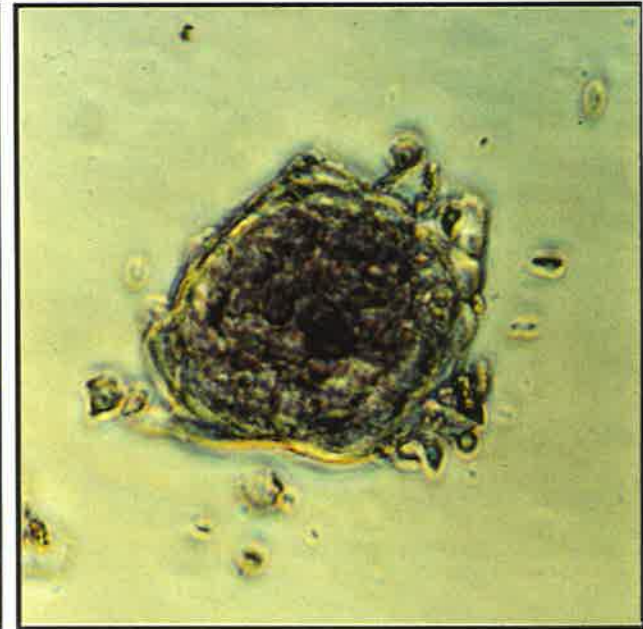
(E). X+3

(F). X+3/R1

1



2



cells reverted for one passage, termed X+3/Rp1 (Figs. 4.5F, 4.6F). Sense control probes of X cell cultures did not exhibit any staining (data not shown).

This analysis demonstrated several important features of the regulation of A03/360 and B04/400 expression during the ES to X cell transition. Firstly, expression of A03/360 and B04/400 was restricted within the *Oct-4*⁺ pool of pluripotent cells. These genes were expressed only by ES cells or closely related early pluripotent cell subpopulations, such as X+0 and X+1 cell types. Secondly, this indicated that in X cell cultures passaged in the presence of LIF, a transition point between “early” and “late” X cell types may occur prior to the X+2 stage. Thirdly, expression of A03/360 and B04/400 was specifically upregulated in “ES like” colonies upon reversion from a “late” X cell type. This indicated that the regulation of A03/360 and B04/400 was likely to be tightly associated with the ES/early subpopulation of pluripotent cells and potentially required for, or a direct consequence of, the generation of the ES/early pluripotent cell “state”. Finally, the similarity in expression profiles of A03/360 and B04/400 during ES cell differentiation, the ES to X cell transition and X cell reversion, suggested that these genes may be coordinately regulated *in vitro*.

This analysis confirmed that A03/360 and B04/400 were markers for ES and early subpopulations of pluripotent cells *in vitro*. Expression analysis of A03/360 and B04/400 during early embryogenesis could therefore identify pluripotent cell subpopulations *in vivo*.

4.3 Analysis of A03/360 and B04/400 Expression During Early Mouse Embryogenesis

Defining the cell types that express ES cell and early pluripotent cell specific markers in the pre-gastrulation mouse embryo would potentially reveal information on the nature and development of embryonic pluripotent cell subpopulations. A03/360 and B04/400 could potentially recognise hitherto poorly defined pluripotent cell subpopulations, with the predicted range of their expression indicated by the ES cell model of pluripotent cell development. Expression was expected to be restricted to pluripotent cell types during early embryogenesis, encompassing the blastocyst ICM and pluripotent ICM derivatives, prior to the generation of primitive ectoderm. Importantly, expression was therefore expected to subdivide the *Oct-4*⁺ pluripotent cell population. This analysis also had potential for

validation of the embryonic stem cell *in vitro* model of development, by identifying the embryonic equivalent of ES cells.

4.3.1 Whole Mount *in situ* Hybridisation Analysis of *Oct-4* Expression During Early Mouse Development

The expression of the POU domain homeobox gene *Oct-4* during early embryogenesis is well documented and it is the best characterised pluripotent cell specific transcript (Rosner *et al.* 1990, Schöler *et al.* 1990a, Schöler *et al.* 1990b, Yeom *et al.* 1991). *Oct-4* is expressed by the fertilised egg, morula, inner cell mass, primitive ectoderm, primordial germ cells and germ cells of the testis and ovary. *Oct-4* is downregulated during terminal differentiation or lineage restriction of these cell types. Analysis of *Oct-4* expression during early embryogenesis therefore provided the most appropriate positive control for A03/360 and B04/400 expression. Furthermore this analysis enabled the development of *in situ* hybridisation methodologies prior to investigation of A03/360 and B04/400 expression.

Non-radioactive whole mount *in situ* hybridisation was used to detect *Oct-4* expression during early embryogenesis. Modifications to existing whole mount techniques were required to enable the analysis of preimplantation and peri-implantation embryos (2.3.31). Briefly, to minimise loss and stress on fragile embryos, flushed preimplantation or dissected postimplantation embryos were processed in “baskets” constructed from eppendorf tubes and nylon mesh (2.5.2), and samples were permeabilised with RIPA buffer rather than proteinase K. Expression was detected using a DIG labelled *Oct-4* antisense riboprobe (2.2.7iii).

The morphological and developmental changes that occur within pluripotent cell populations during early murine embryogenesis have been described in sections 1.2 to 1.4.3. The pluripotent cell specific expression of *Oct-4* during four stages of early embryogenesis is indicated in figure 4.7. At the blastocyst stage of development (3.5 d.p.c.), *Oct-4* expression was restricted to the pluripotent cells of the ICM (Fig. 4.7A). Polar or mural trophectoderm and the visceral layer of primitive endoderm did not express *Oct-4*. During early postimplantation embryogenesis, the pluripotent cells of the ICM proliferate and reorganise to form the early primitive ectoderm. By 5.5 d.p.c. programmed cell death has contributed to the formation of the early proamniotic cavity. *Oct-4* expression in embryos of this stage was

Figure 4.7 Expression of *Oct-4* in Pluripotent Cell Populations During Early Murine Development.

Detection and localisation of *Oct-4* expression during early murine development was examined by whole mount *in situ* hybridisation (2.3.31) using a 394 bp DIG labelled antisense riboprobe (2.2.7iii). Embryo stains were developed for 1-2 hours. Cells that expressed *Oct-4* were indicated by purple staining. Photographs were taken under phase contrast with a Nikon inverted microscope (A, C), or with bright field illumination under a Leica MZ-8 dissecting microscope (B, D). Scale bars: 50 μm (A, B, C), 100 μm (D).

(A) 3.5 d.p.c. blastocyst.

1. polar trophectoderm
2. inner cell mass
3. visceral layer of primitive endoderm
4. blastocoel
5. mural trophectoderm

(B) 5.5 d.p.c. early primitive ectoderm stage embryo.

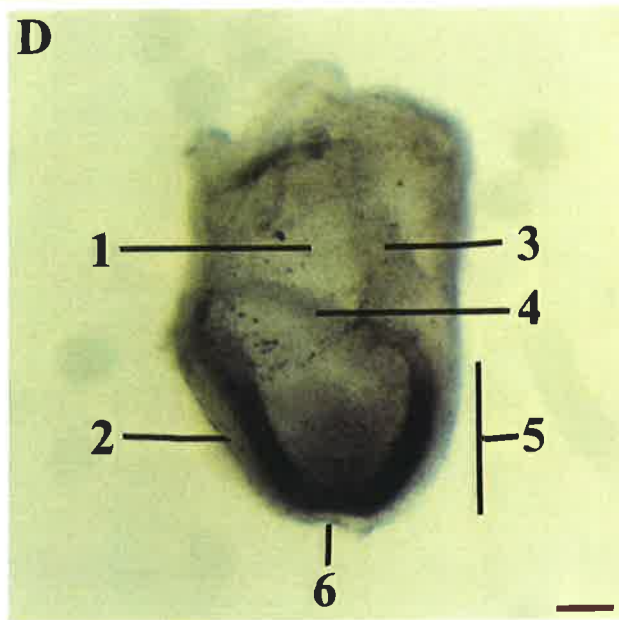
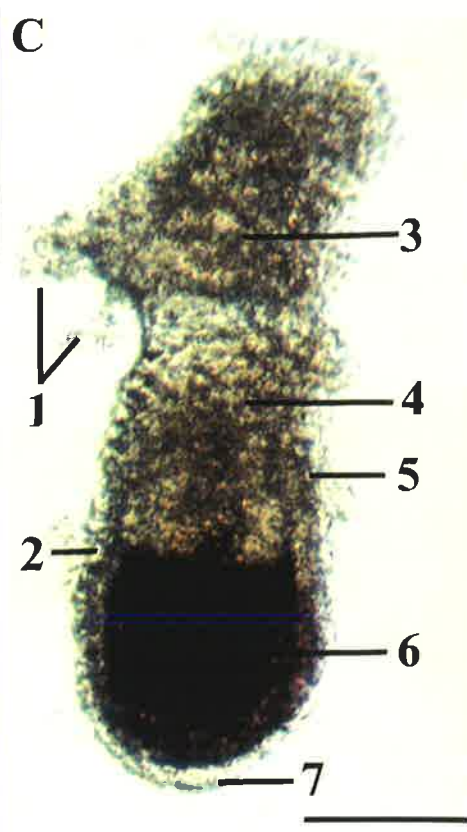
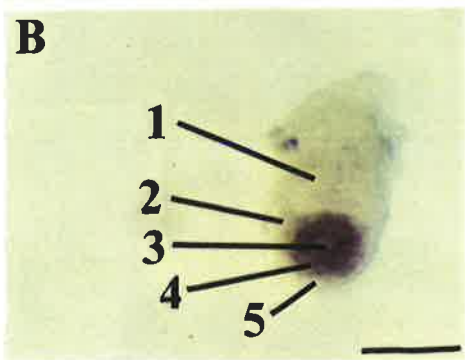
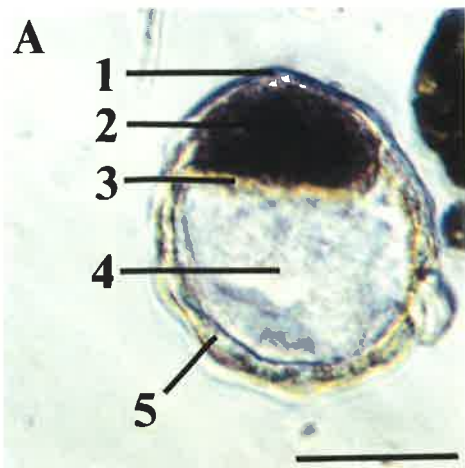
1. extraembryonic ectoderm
2. line of demarcation between embryonic and extraembryonic regions
3. early proamniotic cavity
4. early primitive ectoderm
5. visceral layer of extraembryonic endoderm

(C) 6.0 d.p.c. pregastrulation stage embryo.

1. remnants of Reichert's membrane
2. line of demarcation between embryonic and extraembryonic regions
3. ectoplacental cone
4. extraembryonic ectoderm
5. visceral layer of extraembryonic endoderm
6. primitive ectoderm
7. visceral layer of extraembryonic endoderm

(D) 7.5 d.p.c. midgastrulation stage embryo.

1. exocoelom
2. visceral layer of extraembryonic endoderm
3. allantois
4. amnion
5. region of the primitive streak
6. region of the node



restricted to the pluripotent cells of the early primitive ectoderm, which is a pseudostratified epithelium by this stage (Fig. 4.7B). No expression was exhibited by the extraembryonic ectoderm or visceral endoderm at 5.5 d.p.c. (Fig. 4.7B). At 6.5 d.p.c. the proamniotic cavity has enlarged and the primitive ectoderm and extraembryonic ectoderm have proliferated. *Oct-4* expression at 6.0 d.p.c. was also restricted to the primitive ectoderm, with extraembryonic cell types not staining (Fig. 4.7C). Expression during gastrulation, in the early headfold stage embryo at approximately 7.5 d.p.c., was restricted to the residual primitive ectoderm (Fig. 4.7D). No expression was detectable in other embryonic or extraembryonic cell types. These expression profiles were consistent with previously documented analyses of *Oct-4* using ³⁵S-labelled riboprobes and sectioned embryos (Rosner *et al.* 1990, Schöler *et al.* 1990).

This analysis clearly demonstrated the distribution of *Oct-4*⁺ pluripotent cell types during early development and therefore indicated the viability of using whole mount *in situ* hybridisation to investigate gene expression in pluripotent cells during these stages of embryogenesis.

4.3.1i Specificity of Embryonic *in situ* Hybridisation Analysis

Sense control probes were used to demonstrate the specificity of hybridisation during the establishment of conditions for whole mount *in situ* hybridisation. Blastocysts probed with an *Oct-4* sense control probe (2.2.7iii) did not exhibit any stain after development for two hours (not shown). Early postimplantation embryos (5.25-5.5 d.p.c.) probed with a *Psc1* (B04/400) sense control probe (2.2.7ii) did not exhibit any stain after development for six hours (not shown). The specificity of *Oct-4*, *Icm1* (A03/360) and *Psc1* (B04/400) *in situ* hybridisation analysis was also demonstrated by their unique expression profiles, the presence of negative and positive cell populations within individual embryos, and distinct stages of embryonic development that were negative for expression of these markers.

4.3.2 *In situ* Hybridisation Analysis of *Icm1* (A03/360) Expression During Early Mouse Embryogenesis

To determine the expression profile of A03/360 in pluripotent cell populations during early embryogenesis, preimplantation and early postimplantation embryos were analysed by whole mount *in situ* hybridisation, as shown in figure 4.8. Transcripts were detected using a DIG labelled antisense riboprobe generated from the 5' (256 bp) A03/360 cDNA clone (2.2.7i). During the morula to blastocyst stages of preimplantation development, the expression of A03/360 was found to be specifically restricted to pluripotent cell populations (purple staining, Fig. 4.8). The earliest stage of development analysed was the 16 to 32 cell morula, at approximately 2.5 d.p.c. (Fig. 4.8A). A03/360 was expressed by every cell, or blastomere, at this early stage of embryogenesis. Within the early cavitating blastocyst at approximately 3.0 d.p.c., A03/360 expression was tightly restricted to the pluripotent cells that had been reorganised into the presumptive ICM (Fig. 4.8B, C). A03/360 expression was not detected in polar or mural trophoblast and is presumably downregulated during the differentiation of these cell types from outer blastomeres of the morula. The view of this blastocyst results in a portion of the early ICM lying below the plane of focus, which appears as regions of fainter staining adjacent to the blastocoel (Fig. 4.8B, C). The presence of primitive endoderm within this embryo was not determined. Specific ICM staining of A03/360 expression was also shown during blastocyst development at approximately 3.5 d.p.c. (data not shown). Blastocyst expression of A03/360 at 4.0 d.p.c. is indicated in figures 4.8D and 4.8E. A03/360 was specifically expressed by the pluripotent cells of the ICM. Importantly, the expression of A03/360 was downregulated in three differentiated cell lineages. Like early blastocysts, polar and mural trophoblast did not express A03/360. At this stage of embryogenesis, ICM cells lining the blastocoel have differentiated to primitive endoderm. A03/360 was downregulated during the differentiation of the visceral layer of primitive endoderm.

The analysis of A03/360 during early postimplantation development is indicated in figures 4.8F and 4.8H. Following implantation, the conceptus undergoes rapid proliferation and the ICM becomes reorganised into the early primitive ectoderm. A03/360 expression was downregulated during these events, with little or no expression detected in early primitive

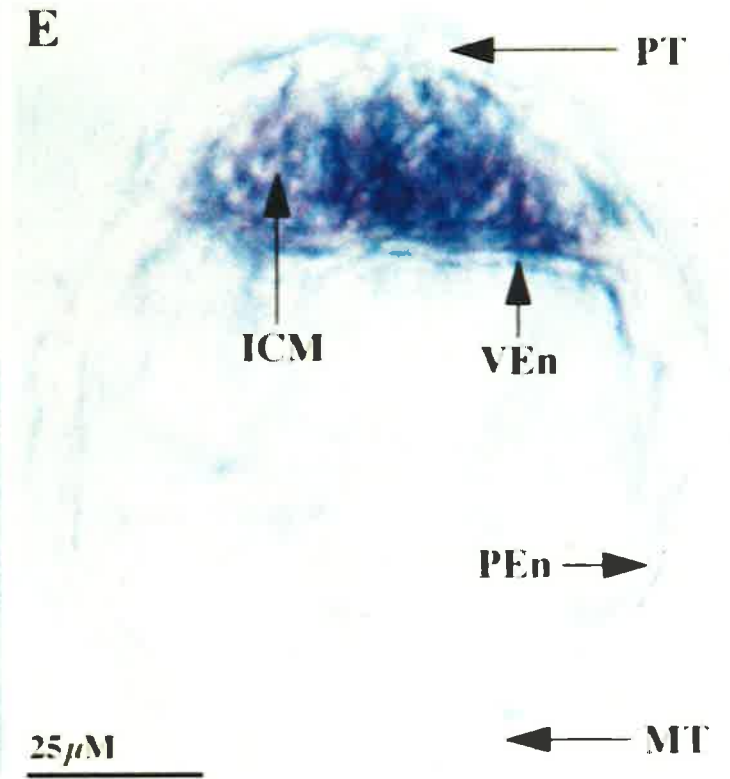
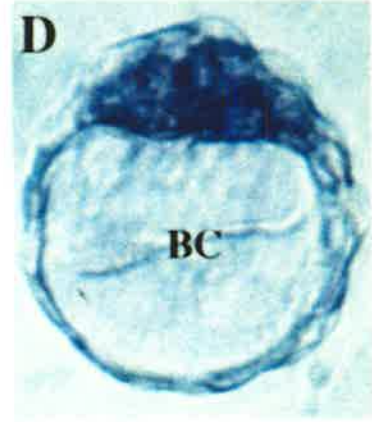
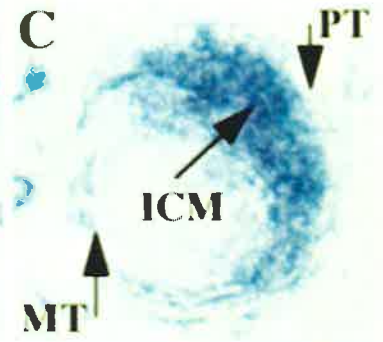
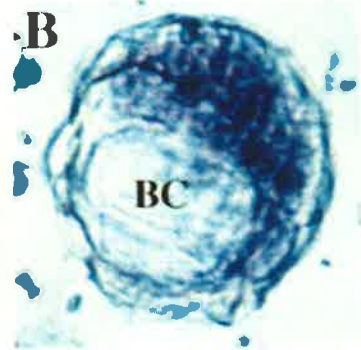
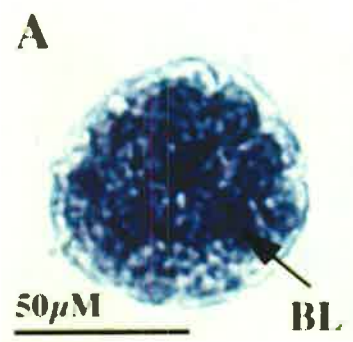
Figure 4.8 Expression of *Icm1* (A03/360) During Early Murine Embryogenesis.

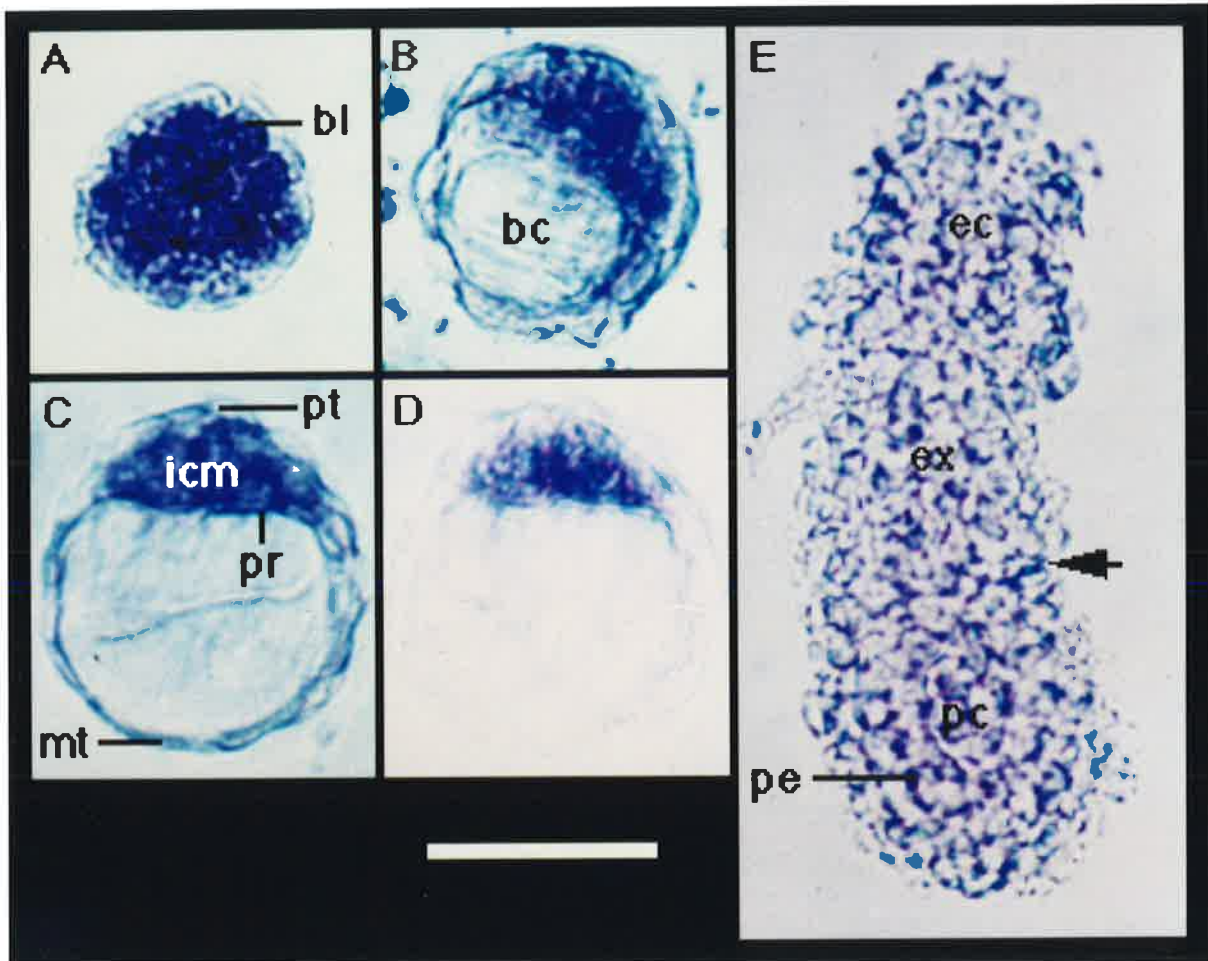
Detection and localisation of *Icm1* expression during early murine development was examined by *in situ* hybridisation (2.3.31) using a 256 bp DIG labelled antisense riboprobe (2.2.7i). Embryo stains were developed for 3-6 hours. Cells that expressed *Icm1* were indicated by purple staining. Photographs were taken under phase contrast (A, B, D, F) or bright field optics (C, E) with a Nikon inverted microscope, or under bright field illumination with a Leica MZ-8 dissecting microscope (H). Scale bars: 50 μm (A-D), 25 μm (E), 50 μm (F, G), 50 μm (H, I).

| | | | |
|-----------------------|--|-----|------------------------------|
| BL | compacted blastomere | RM | residual Reichert's membrane |
| BC | blastocoel | EPC | ectoplacental cone |
| MT | mural trophoctoderm | EX | extraembryonic ectoderm |
| PT | polar trophoctoderm | PE | primitive ectoderm |
| ICM | inner cell mass | PC | proamniotic cavity |
| VE _n | visceral layer of primitive endoderm | | |
| <u>VE_n</u> | visceral endoderm | | |
| PE _n | future migratory surface of parietal endoderm | | |
| De | non-specific debris | | |
| LD | approximate line of demarcation between embryonic and extraembryonic regions | | |

- (A) 2.5 d.p.c. 16-32 cell compacted morula.
- (B, C) 3.0 d.p.c. early blastocyst.
- (D) 4.0 d.p.c. blastocyst.
- (E) 4.0 d.p.c. blastocyst (enlarged bright field view of (D)).
- (F) 5.5 d.p.c. early primitive ectoderm stage embryo.
- (G) Schematic representation of (F). Dashed lines suggest ectoplacental cone and visceral endoderm regions.
- (H) 5.75 d.p.c. early primitive ectoderm stage embryo.
- (I) Schematic representation of (H). Dashed lines suggest proamniotic cavity, ectoplacental cone and visceral endoderm regions.

An amendment to Figure 4.8, termed Amended Figure 4.8.2, is presented on the reverse of the following page.



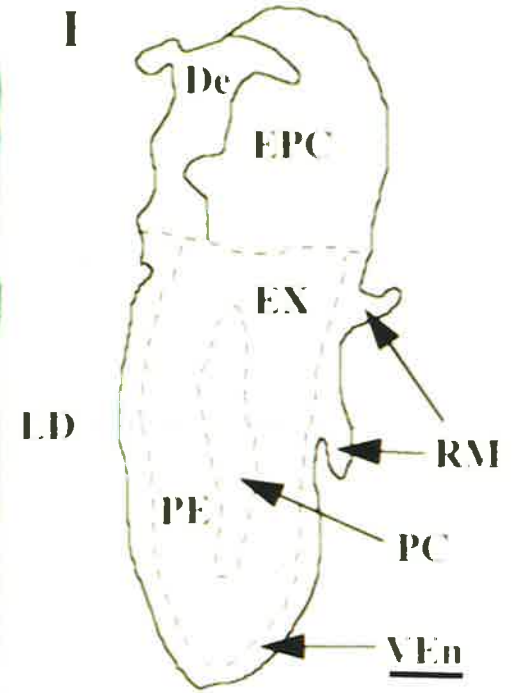
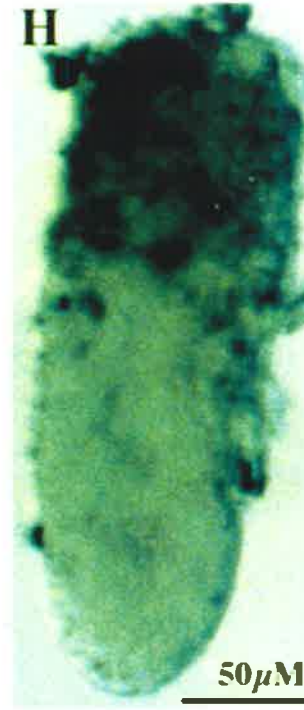
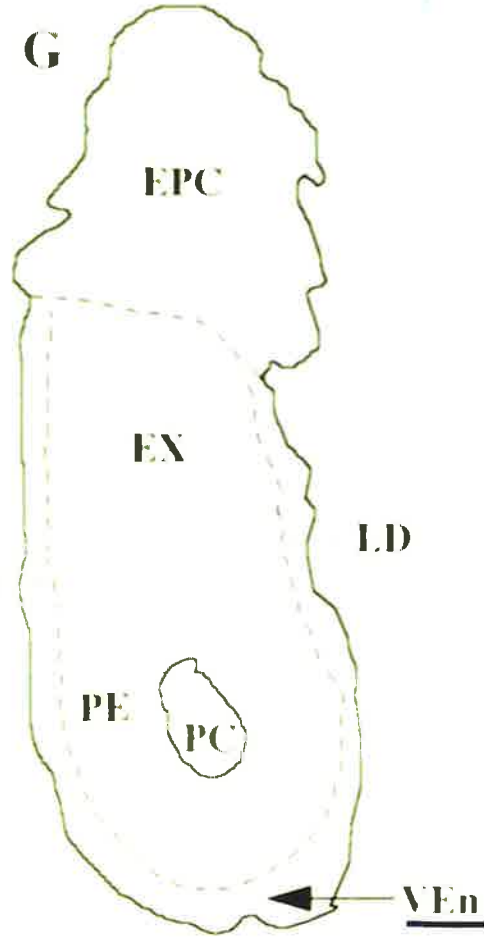
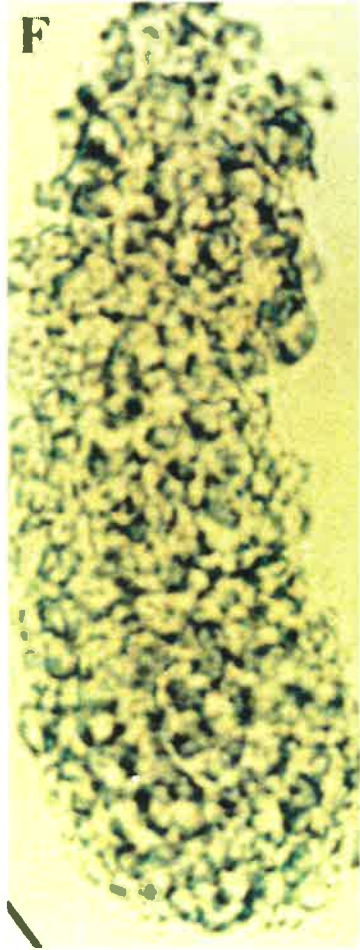


Amended Figure 4.8.2 Expression of *Icm1* (A03/360) During Early Murine Embryogenesis.

This figure depicts the same embryos as Figure 4.8, without panels 4.8 C, G, H & I.

Panels 4.8.2 C, D & E represent panels 4.8 D, E & F, respectively.

Scale bar: 50 μm . The labels are the same as for Figure 4.8 except: ec, ectoplacental cone; arrowhead, approximate line of demarcation between embryonic and extraembryonic regions.



ectoderm stage embryos with proamniotic cavities at 5.5 d.p.c. (Fig. 4.8F). Critically, the primitive ectoderm, the pluripotent *Oct-4*⁺ descendant of the ICM, did not express A03/360. That A03/360 was downregulated prior to the formation of primitive ectoderm was confirmed by the lack of expression in embryos at approximately 5.75 d.p.c. of development (Fig. 4.8H). Dark areas in the ectoplacental cone region of the embryo shown in figure 4.8H are a result of non-specific debris. The distribution of cell types in the embryos shown in figures 4.4F and 4.4H are schematically represented in figures 4.4G and 4.4I respectively. Postimplantation embryos were staged on the basis of size and morphology of the early proamniotic cavity.

In situ hybridisation analysis during early embryogenesis therefore demonstrated that A03/360 expression was restricted to the pluripotent cells of the preimplantation embryo. While stages of development between the ICM and early primitive ectoderm are relatively inaccessible and were not analysed, A03/360 clearly subdivided the *Oct-4*⁺ population, marking a pluripotent cell type present in the morula, and in the early and late blastocyst ICM. The gene and transcripts detected by the A03/360 probe were therefore named *Icm1* (Inner cell mass 1).

4.3.3 *In situ* Hybridisation Analysis of *Psc1* (B04/400) Expression During Early Mouse Embryogenesis

The expression of B04/400 during early embryogenesis was detected with a DIG labelled antisense riboprobe generated from the B04/400 cDNA clone (2.2.7ii). The expression of B04/400 was overlapping but different from that of *Icm1* (A03/360) within the pluripotent stem cell pool and is shown in figure 4.9. B04/400 expression was not detectable during early blastocyst development, between approximately 3.0 and 4.0 d.p.c. (Fig 4.9A, B), even with prolonged exposures. Low level B04/400 expression was first observed in the late stage blastocyst, at approximately 4.0 d.p.c. and is indicated in figures 4.9C and 4.9D. B04/400 was specifically expressed in the pluripotent cells of the late stage ICM. Interestingly, the differentiation of primitive endoderm had occurred by this stage, with no B04/400 expression detected in the visceral layer of primitive endoderm. Mural trophectoderm did not stain and it appeared that polar trophectoderm did not express

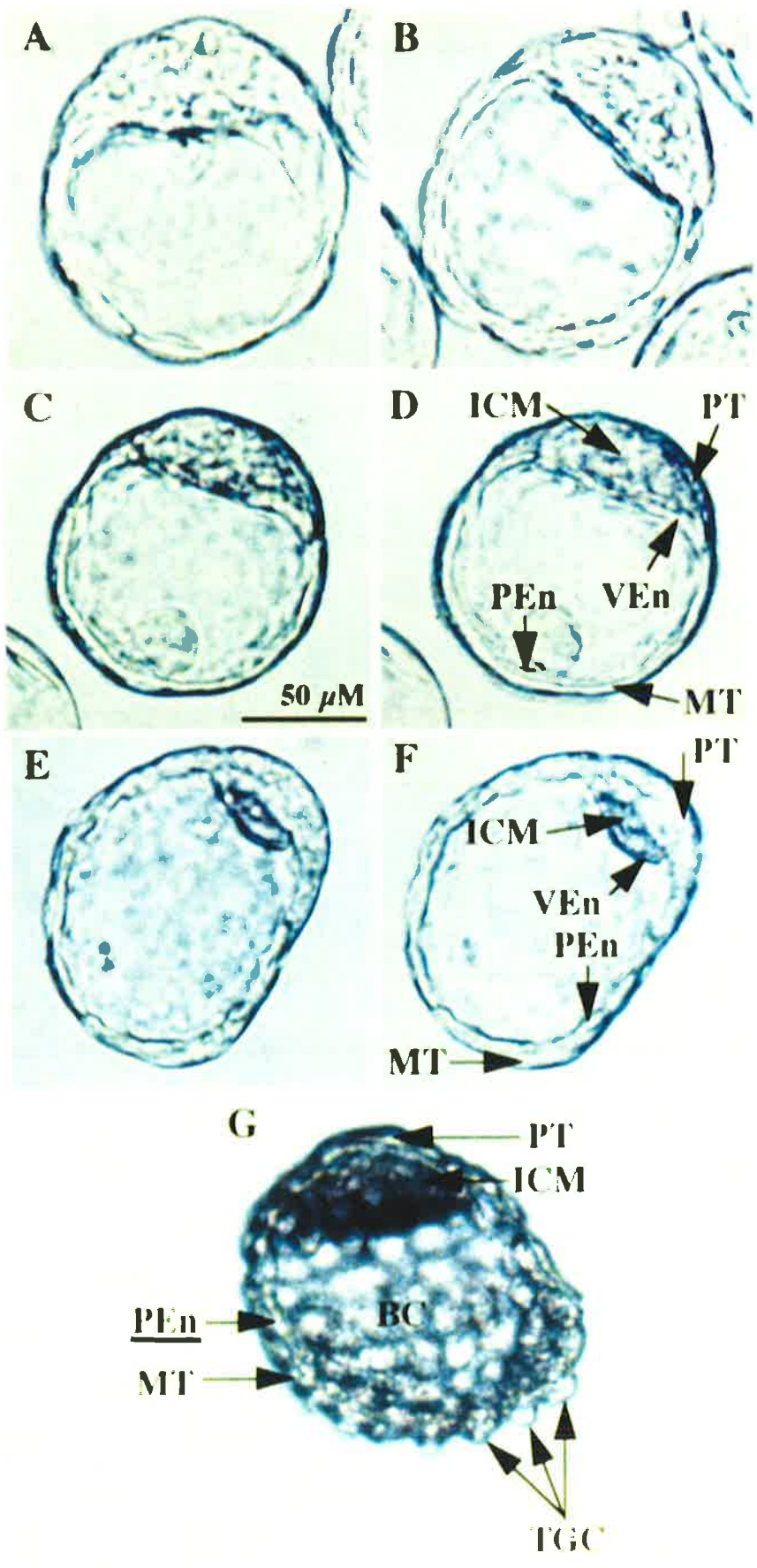
Figure 4.9 Expression of *Psc1* (B04/400) During Early Murine Embryogenesis.

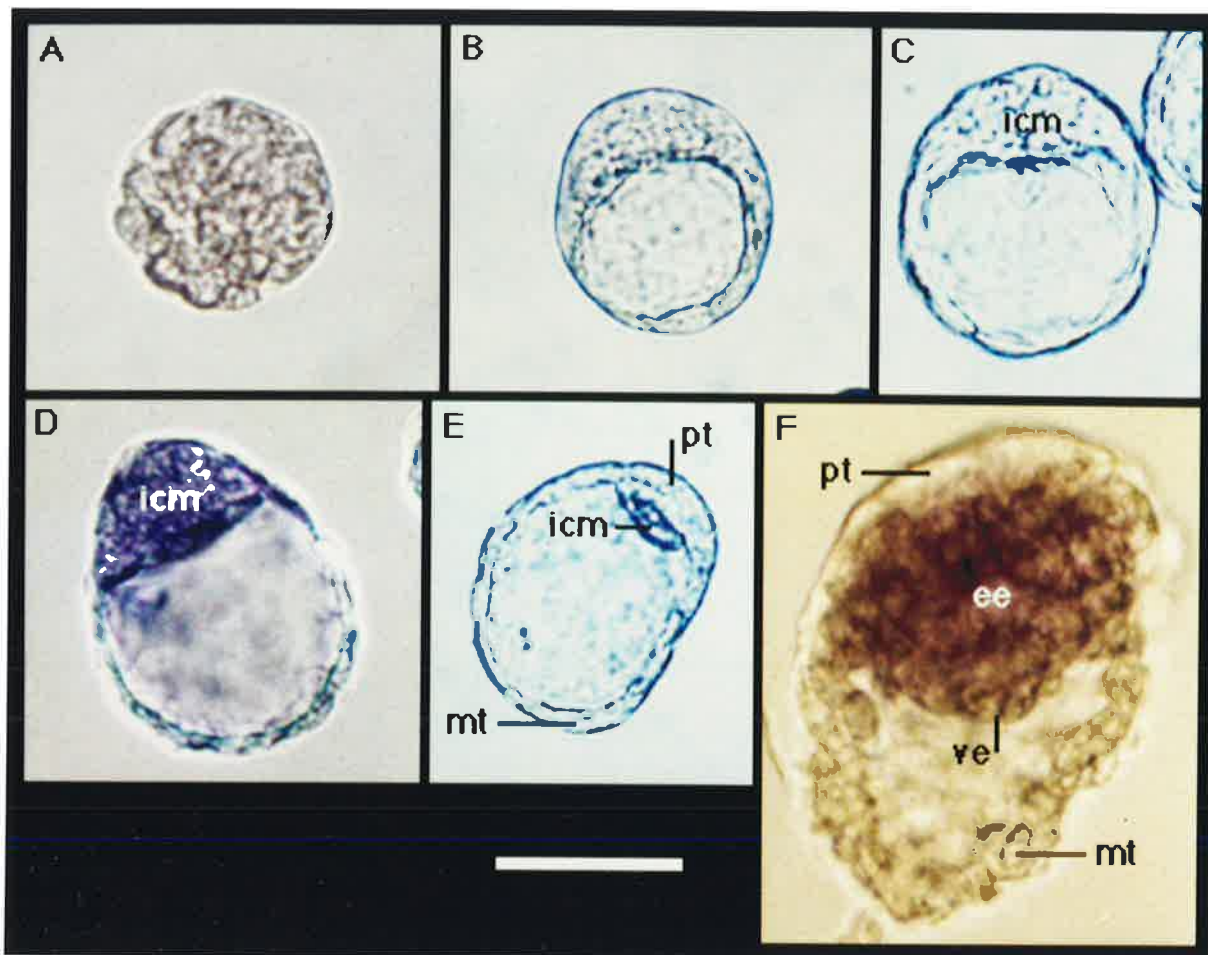
Detection and localisation of *Psc1* expression during early murine development was examined by whole mount *in situ* hybridisation (2.3.31) using a 458 bp DIG labelled antisense riboprobe (2.2.7ii). Embryo stains were developed for 3-6 hours. Cells that expressed *Psc1* were indicated by purple staining. Photographs were taken under phase contrast (C, E, G, H, K, N, P) or bright field optics (A, B, D, F, I, L) with a Nikon inverted microscope. Scale bar: 50 μ m (A-Q same scale).

| | | | |
|-------------|--|-----|----------------------------------|
| BC | blastocoel | TGC | primary trophoblastic giant cell |
| MT | mural trophectoderm | EPC | ectoplacental cone |
| PT | polar trophectoderm | EX | extraembryonic ectoderm |
| ICM | inner cell mass | PE | primitive ectoderm |
| EE | embryonic ectoderm | RM | residual Reichert's membrane |
| PC | proamniotic cavity | De | non-specific debris |
| <u>VE</u> n | visceral endoderm | | |
| VEn | visceral layer of primitive endoderm | | |
| PEn | future migratory surface of parietal endoderm | | |
| <u>PE</u> n | parietal endoderm | | |
| LD | approximate line of demarcation between embryonic and extraembryonic regions | | |

- (A) 3.5 d.p.c. blastocyst.
 (B) 3.5 d.p.c. blastocyst.
 (C, D) 4.0-4.5 d.p.c. blastocyst.
 (E, F) 4.5 d.p.c. fully expanded blastocyst.
 (G) 4.75 d.p.c. peri-implantation blastocyst.
 (H, I) 5.0 d.p.c. peri-implantation embryo, "epiblast bud" stage.
 (J) Schematic representation of (H, I). Dashed lines suggest early ectoplacental cone and visceral endoderm regions.
 (K, L) 5.0-5.25 d.p.c. peri-implantation embryo, very early proamniotic cavity stage.
 (M) Schematic representation of (K, L). Dashed lines suggest early ectoplacental cone, residual Reichert's membrane and visceral endoderm regions.
 (N) 5.5 d.p.c. early primitive ectoderm stage embryo.
 (O) Schematic representation of (N). Dashed lines suggest early ectoplacental cone and visceral endoderm regions.
 (P) 5.5 d.p.c. early primitive ectoderm stage embryo.
 (Q) Schematic representation of (P). Dashed lines suggest early ectoplacental cone and visceral endoderm regions.

Amendments to Figure 4.9, termed Amended Figure 4.9.2 and Figure 4.9.3, are presented on the reverse of the following two pages.





Amended Figure 4.9.2 Expression of *Psc1* (B04/400) During Early Murine Embryogenesis.

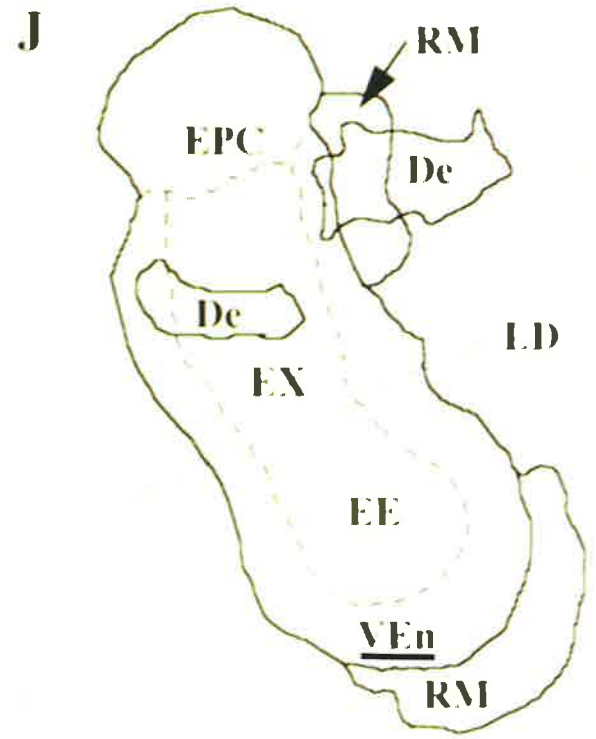
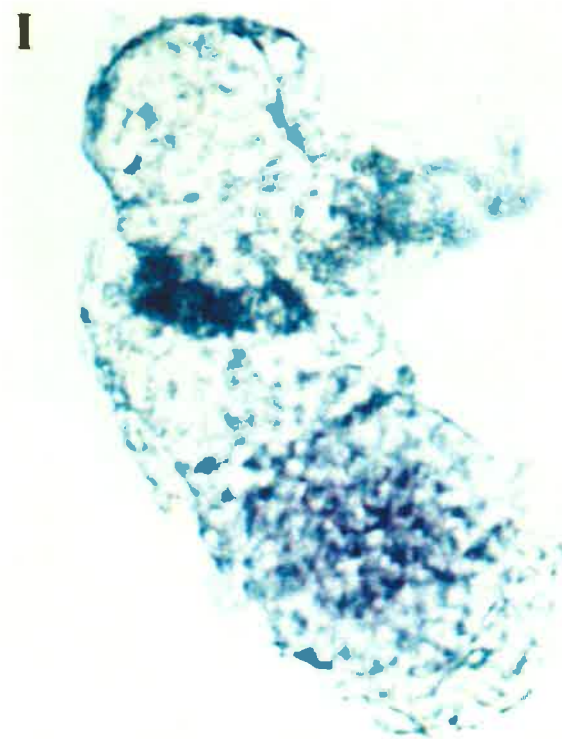
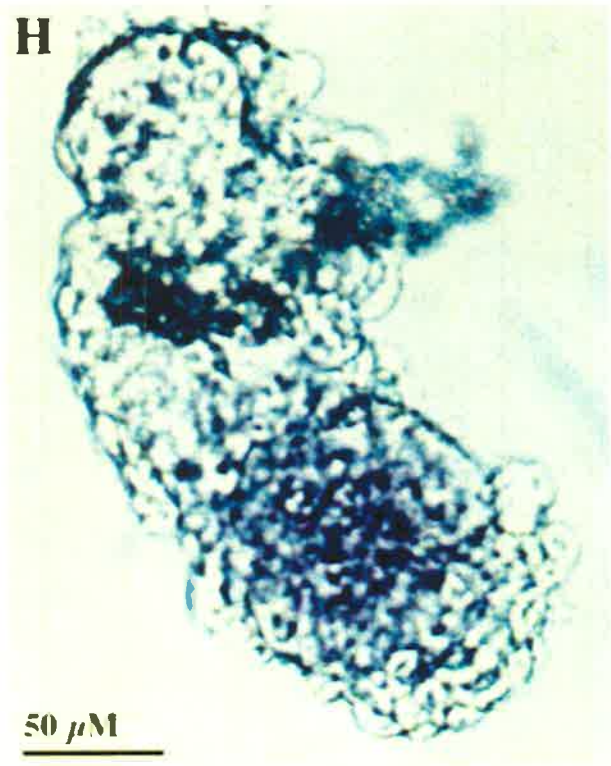
This figure depicts the same embryos as Figure 4.9 panels A & E, in panels C & E, respectively and additional embryos using the same experimental conditions.

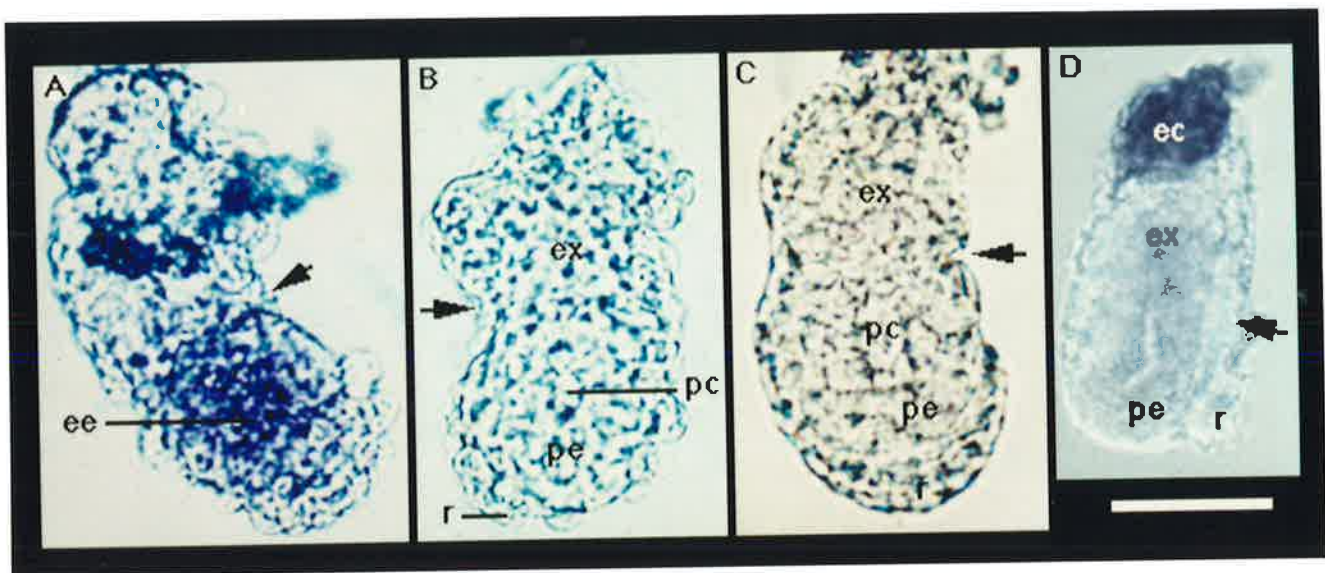
(A) 2.5 d.p.c. 16-32 cell morula. (B) 3.0 d.p.c. early blastocyst.

(D) 4.0-4.5 d.p.c. blastocyst. (F) 5.0 d.p.c. peri-implantation embryo.

Scale bar: 50 μm . The labels are the same as for Figure 4.9 except:

ee, embryonic ectoderm; ve, visceral endoderm. Photographs were taken under phase contrast (B) with a Nikon inverted microscope, or with differential interference optics (A, D), or bright field optics (F) with a Zeiss microscope.

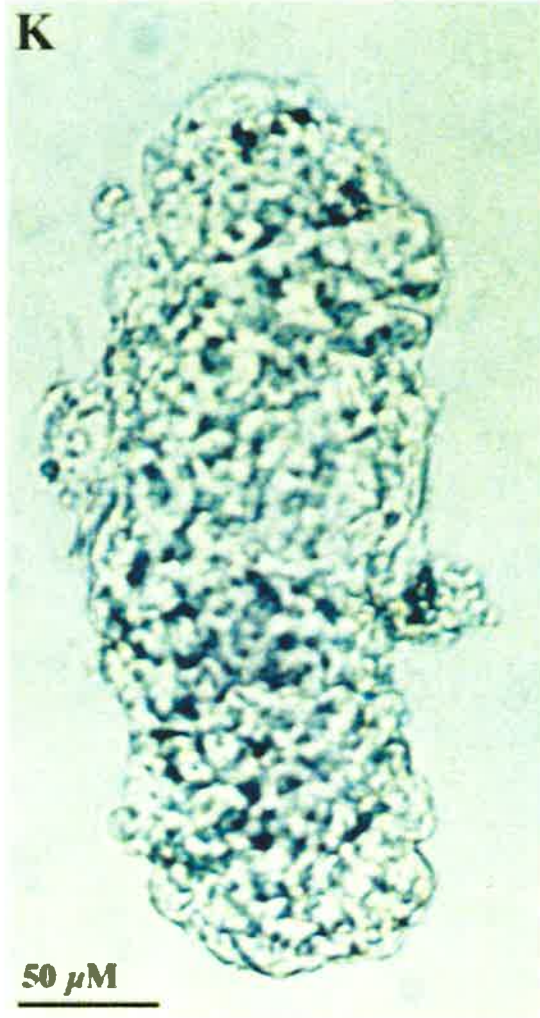
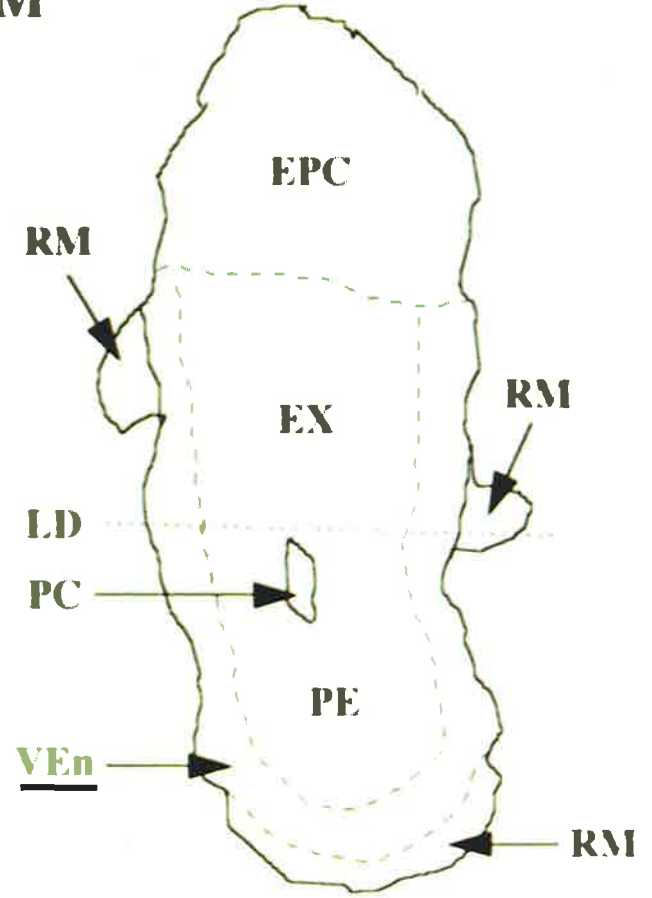


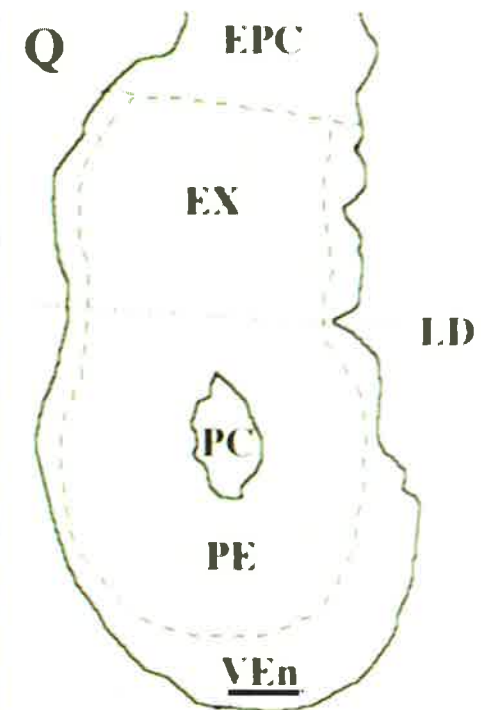
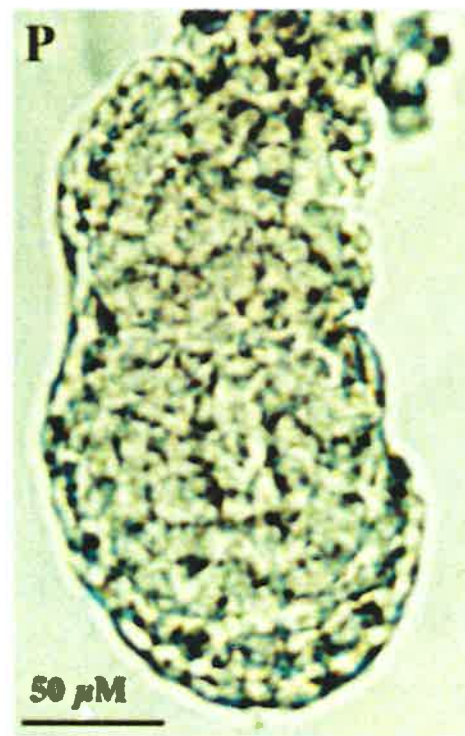
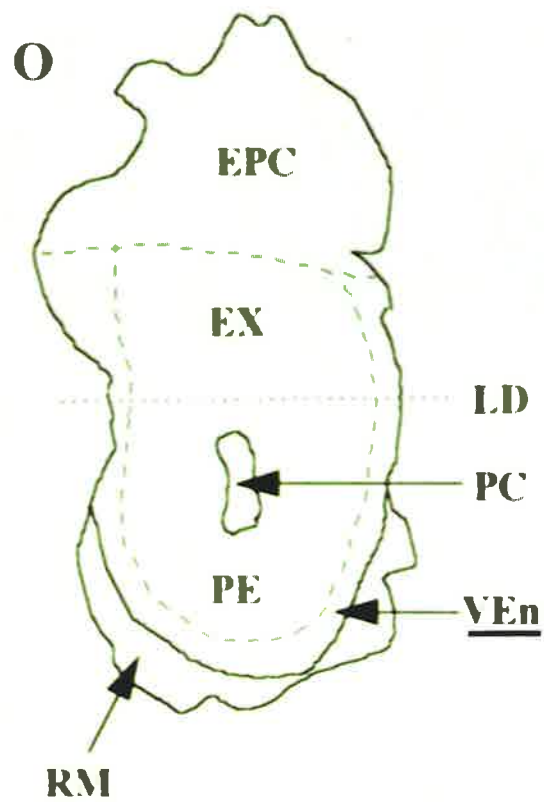
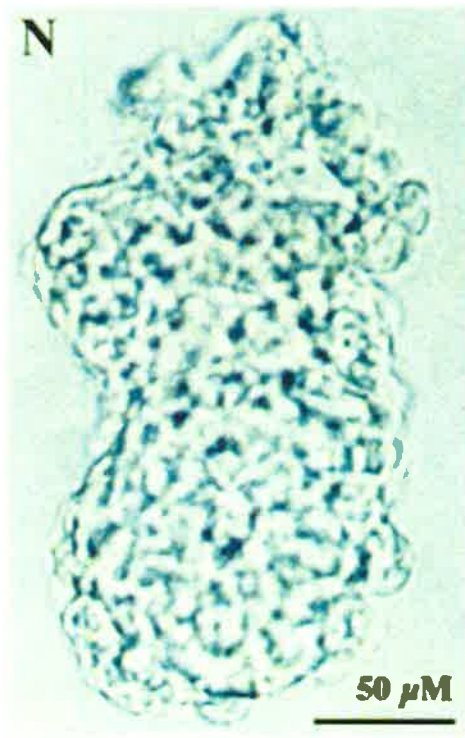


Amended Figure 4.9.3 Expression of *Psc1* (B04/400) During Early Murine Embryogenesis.

This figure depicts the same embryos as Figure 4.9 panels H, N & P in panels A, B & C, respectively. Panel D depicts the embryo in Figure 4.10 A, photographed with differential interference optics with a Zeiss microscope.

Scale bar: 50 μm (A-C), 100 μm (D). The labels are the same as for Figure 4.9 except: ee, embryonic ectoderm; ec, ectoplacental cone; r, residual Reichert's membrane; arrowhead, approximate line of demarcation between embryonic and extraembryonic regions.

K**L****M**



B04/400, although this cell type was not clearly distinguished (Fig. 4.9D). Analysis of the fully expanded blastocyst at 4.5 d.p.c, supported the restriction of B04/400 expression to the undifferentiated cells of the ICM (Fig. 4.9E, F). At this stage of development the embryo has hatched from the zona pellucida and has assumed an ovoid structure. The ICM has flattened and abuts the polar trophoderm. B04/400 was clearly not expressed by the polar or mural trophoderm, with the visceral layer of primitive endoderm not clearly distinguishable. This analysis suggested that B04/400 was not expressed in the early blastocyst and was not expressed by any of the differentiated lineages of the blastocyst. B04/400 was first expressed during the development of the late stage inner cell mass at approximately 4.0 to 4.5 d.p.c.

The expression of B04/400 during peri-implantation and early postimplantation development is indicated in figures 4.9G to 4.9Q. Uterine implantation of the blastocyst occurs at approximately 4.5 d.p.c. The expression of B04/400 in a peri-implantation embryo, flushed from uteri at approximately 4.75 d.p.c., is demonstrated in figure 4.9G. This embryo still exhibits the basic blastocyst structure, although mural trophodermal cell types not in contact with the epiblast have ceased proliferation and become primary trophoblastic giant cells (Hogan *et al.* 1994). The expression of B04/400 was clearly restricted to the pluripotent cells of the epiblast. Trophodermal derivatives and parietal endoderm did not express B04/400, but expression in visceral endoderm could not be ascertained due to the close proximity of the highly stained epiblast.

The proliferation and proamniotic cavitation of the epiblast are major developmental events involved with the conversion of the ICM to the primitive ectoderm (1.3, 1.4.1). At 5.0 d.p.c. the pluripotent cells of the embryonic ectoderm; the derivatives of the ICM, have proliferated and protruded into the former blastocoel, forming a solid structure without a proamniotic cavity. In an embryo isolated at this time point, B04/400 was specifically expressed by a bud of pluripotent embryonic ectoderm (Figs. 4.9H, I). Cell types of residual Reichert's membrane, visceral endoderm, extraembryonic ectoderm and the early ectoplacental cone did not express B04/400 (Fig. 4.9I). The distribution of cell types in this embryo is schematically represented in Figure 4.9J.

At approximately 5.25 d.p.c. a small central cavity begins to form in the epiblast, caused by the programmed cell death of central epiblast cells (Coucovanis and Martin 1995).

The proamniotic cavity is enlarged by 5.5 d.p.c., with the pluripotent cells becoming reorganised into the cup-shaped epithelial layer of the primitive ectoderm. In an embryo staged at approximately 5.25 d.p.c., specific B04/400 expression in the embryonic ectoderm was significantly downregulated (Figs. 4.9K, L). This embryo was staged on the basis of the first detectable evidence of a proamniotic cavity (Fig. 4.9L, schematically depicted in 4.9M) and indicated the embryological event that correlated with B04/400 downregulation. Postimplantation embryos at the early primitive ectoderm stage (5.5 d.p.c.) contained small, but distinct proamniotic cavities and did not express B04/400. Examples of this are indicated in figures 4.9N and 4.9P and schematically represented in figures 4.9O and 4.9Q respectively.

From 5.5 d.p.c. to 6.0 d.p.c. proamniotic cavitation extends through the extraembryonic ectoderm and the ectoplacental cone (EPC) becomes a more defined structure (Kaufman 1992, Hogan *et al.* 1994). The expression of B04/400 in embryos with intact EPC regions, staged at approximately 5.5 d.p.c. to 5.75 d.p.c., is indicated in figure 4.10. B04/400 was not expressed in primitive ectoderm, visceral endoderm, or extraembryonic ectoderm at this stage of embryogenesis (Fig. 4.10. Additional analyses not shown). However, B04/400 was expressed at a high level in the EPC at approximately 5.5 d.p.c. to 5.75 d.p.c. The ectoplacental cone is derived from the trophoblast and forms a major component of the placenta during later gestation (Hogan *et al.* 1994). This indicated that B04/400 could have a role during the formation of early placental structures during postimplantation development. B04/400 expression was not detected in the trophectoderm of the blastocyst (Figs. 4.9E-G) and in early EPC regions during peri-implantation development (Figs. 4.9H-M). This indicated that expression of *Psc1* in the EPC was induced during early postimplantation development at approximately 5.5 d.p.c. The precise timing of the induction of B04/400 expression in the EPC was not determined, as this region was generally removed from postimplantation embryos during dissection (compare Fig. 4.10 to Fig. 4.9P).

During the analysis of B04/400 expression, all the depicted embryos excluding those in figures 4.9G and 4.10, were simultaneously compared with consistent probe and detection regimes.

Figure 4.10 Expression of *Psc1* (B04/400) in the Ectoplacental Cone at 5.5-5.75 d.p.c.

Detection and localisation of *Psc1* expression during early murine development at 5.5-5.75 d.p.c. was examined by whole mount *in situ* hybridisation (2.3.31) using a 458 bp DIG labelled antisense riboprobe (2.2.7ii). Embryo stains were developed for 3 hours. Cells that expressed *Psc1* were indicated by purple staining. Photographs were taken under bright field optics with a Nikon inverted microscope, using Ectachrome 100 slide film. Scale bar: 50 μ m.

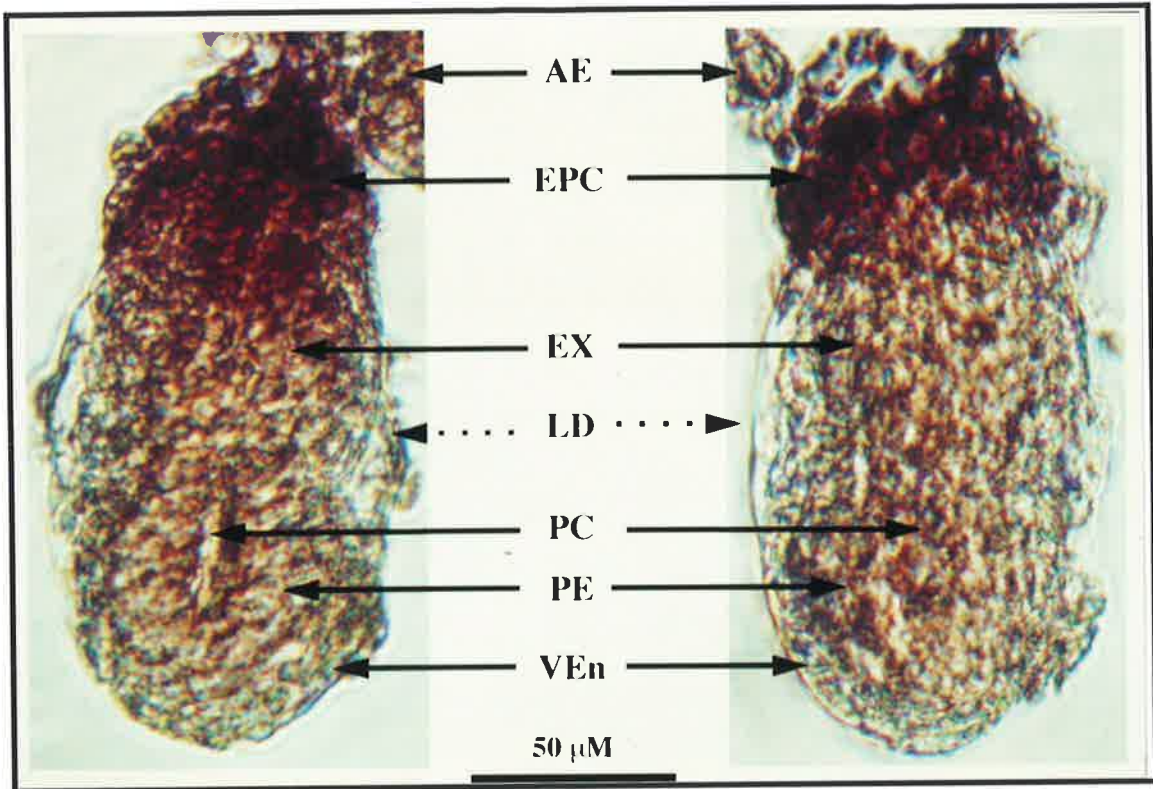
| | |
|-----|--|
| EPC | ectoplacental cone |
| EX | extraembryonic ectoderm |
| PE | primitive ectoderm |
| PC | proamniotic cavity |
| VEn | visceral endoderm |
| LD | approximate line of demarcation between embryonic and extraembryonic regions |
| AE | tissue from an adjacent embryo within the photograph |

- (A) 5.5 d.p.c. early primitive ectoderm stage embryo with intact ectoplacental cone.
(B) 5.5-5.75 d.p.c. early primitive ectoderm stage embryo with intact ectoplacental cone.

An amendment to Figure 4.10 is included in Amended Figure 4.9.2.

A

B



In situ hybridisation analysis during early embryogenesis therefore indicated that B04/400 was expressed from the pluripotent cells of the late and fully expanded blastocyst ICM, to the peri-implantation embryonic ectoderm, prior to proamniotic cavitation. B04/400 was also expressed in the ectoplacental cone at approximately 5.5 d.p.c. The gene identified by B04/400 was therefore named *Psc1* (**p**eri-**i**mplantation **s**tem **c**ell **1**). *Psc1* (B04/400) clearly identified a pluripotent cell subpopulation, distinct from both *Oct-4*⁺ and *Icm1*⁺ (A03/360) pluripotent cell populations.

4.4 Discussion

The development of early embryonic pluripotent cells through various distinct stages has been demonstrated by transplantation, differentiation and gene expression analysis (1.5.1). This information indicates the existence of heterogeneity and specific pluripotent cell subpopulations within the developing stem cell pool. The molecular basis of pluripotent cell subpopulations has not been addressed in detail, hence the scarcity of markers that define their existence and boundaries. By targeting gene transcripts as potential markers and using an ES cell model of development, a system was designed to detect and isolate expression differences between pluripotent cells *in vitro* (Chapter 3). This chapter described the more detailed analysis of two isolated ES cell and early pluripotent cell markers, *Icm1* (A03/360) and *Psc1* (B04/400).

4.4.1 The Expression of *Icm1* and *Psc1* in Pluripotent Cell Subpopulations *in vitro*

The examination of *Icm1* and *Psc1* expression during ES cell differentiation and the formation of X cells provided insight into the nature of stem cell maintenance and gene expression *in vitro*. Both markers were regulated in a precise manner, being restricted to subpopulations of *Oct-4*⁺ pluripotent cells. Importantly, *Icm1* and *Psc1* expression was restricted to pluripotent cell populations, with rapid downregulation observed during terminal differentiation. The transcription of *Icm1* and *Psc1* was clearly maintained within “early” X cell types, as predicted by DDPCR (3.3.2) and northern analysis (3.4.1), as expression was detected in proliferating colonies after passaging had split cultures to single cells. The loss of

Icm1 and *Psc1* transcripts prior to X+2 cultures was consistent with the downregulation of *Rex-1* expression and induction of *FGF-5* expression in X cells passaged in the absence of LIF (Lake 1996) and can now be associated with the regulated progression of X cells from an “early” to a “late” pluripotent cell state. Low level expression of *Icm1* or *Psc1* was not detected in X+2 and X+3 cells, compared to DDPCR and northern analysis of these markers (3.3.2, 3.4.1). This may be caused by differences in the relative sensitivity of these techniques, or differing rates of developmental progression in individual X cell batches. Interestingly, pluripotent X+3 cells, or “late” X cells, that had downregulated *Icm1* and *Psc1*, rapidly induced expression of these genes when reverted to an ES cell morphology by passaging in LIF in the absence of MedII. This information, coupled with the general similarity of *Icm1* and *Psc1* expression, demonstrates that these genes might be coordinately regulated in these cell types. The restriction of *Icm1* and *Psc1* to ES cells, “early” X cells and reverted X cells may indicate that they are functionally linked to the maintenance of the undifferentiated state of “early” pluripotent cell subtypes *in vitro*. Conversely, *Icm1* and *Psc1* may be expressed as a direct consequence of the cell state defined by ES or “early” X cells. The roles of *Icm1* and *Psc1* during the development of pluripotent cell populations may be functionally assayed through the use of mutational approaches in ES cells, combined with pluripotent cell progressions or cellular differentiation regimes *in vitro*. Inactivation or modification of a single or both alleles of *Icm1* or *Psc1* by homologous recombination, or ectopic expression of *Icm1* or *Psc1*, could enable the functional analysis of these genes during the development of pluripotent cell subpopulations *in vitro*.

Two types of *Icm1*⁺/*Psc1*⁺/*Oct-4*⁺ residual pluripotent cells were identified in induced ES cell cultures (Figs. 4.1-4.3). One type exhibited a domed ES cell colony morphology, whereas the other consisted of cell monolayers with specific nuclear characteristics. This analysis enabled the comparison of these pluripotent cell types to those found during the ES to X cell transition (Figs. 4.4-4.6). On the basis of morphological similarities and expression of *Icm1* and *Psc1*, the domed residual colonies found in differentiated cultures could be equivalent to undifferentiated ES cells, while the flattened colony type could represent a population akin to “early” X cells, such as those present in X+0 or X+1 cultures. The identification of *Icm1*⁺/*Psc1*⁺/*Oct-4*⁺ “early” X cell-like cells during ES cell differentiation is

consistent with the proposition that X cells are a normal intermediate during the differentiation of ES cells (Thomas 1995). The existence of an *Icm1*⁻/*Psc1*⁻/*Oct-4*⁺ pluripotent cell type, similar to “late” X cells, would therefore also be expected during ES cell differentiation, but was not identified by this approach. The development of double labelling *in situ* hybridisation techniques to compare the expression of *Oct-4* to *Icm1* or *Psc1* within a single cell may be required to identify this cell type during ES cell differentiation. Further analysis and comparison of predicted pluripotent cell subpopulations during ES cell differentiation may depend on the future isolation of X cell subtype specific markers by DDPCR (3.3.3).

4.4.2 *Icm1* and *Psc1* Expression and Pluripotent Cell Subpopulations in the Mouse Embryo

In situ hybridisation analysis of *Oct-4*, *Icm1* and *Psc1* in the early mouse embryo demonstrated the development of appropriate techniques for mapping gene expression during the early stages of pluripotent cell development *in vivo*. These approaches, in combination with the use of DDPCR and controlled pluripotent cell progression *in vitro*, have clearly demonstrated the capacity to develop markers that address the question of molecular heterogeneity within the developing pluripotent cell pool *in vitro* and *in vivo*, which was the major objective of this thesis.

Icm1 and *Psc1* were specifically expressed in overlapping but distinct stages of pluripotent cell development in the early mouse embryo. *Icm1* was expressed by the pluripotent blastomeres of the morula and expression was downregulated during trophectoderm differentiation, becoming restricted to the early ICM cells of the cavitating blastocyst. In the late stage blastocyst, *Icm1* was specifically expressed by the pluripotent cells of the ICM. *Icm1* was not detected in early postimplantation stages of embryogenesis, with expression of the gene being downregulated in the early primitive ectoderm. Therefore, the expression of *Icm1* identified a pluripotent cell subpopulation within the *Oct-4*⁺ pool, from the blastomeres of the morula to the ICM of the late stage blastocyst. Further analysis of cleavage stage and peri-implantation development may refine the temporal boundaries of *Icm1* expression and the distribution of this pluripotent cell subpopulation.

Psc1 was not expressed during early blastulation but was first detected in the ICM of the late stage blastocyst. *Psc1* was expressed during peri-implantation development in ICM derivatives and in the embryonic ectoderm, or epiblast bud. Most interestingly, *Psc1* appeared to be downregulated at the onset of proamniotic cavitation and was not expressed by the early primitive ectoderm. While additional analysis of *Psc1* during peri-implantation development may address embryonic intermediates between the proliferated blastocyst and epiblast bud stages, it is unlikely that further refinement of the timing of *Psc1* downregulation could easily be achieved.

Psc1 expression therefore highlighted the presence of a transient stem cell type within the *Oct-4⁺/Icm1⁺* pluripotent cell pool. This stem cell subpopulation was specific to peri-implantation development and differed at a molecular level from both preceding and following pluripotent cell types. This represents the first identification of a specific peri-implantation stem cell subtype and demonstrates the power of ES cell approaches to the analysis of developmental heterogeneity within the pluripotent cell pool. This suggests that specific transcripts and markers for other stem cell types could also be identified through these approaches, for example markers of heterogeneity during the development of the primitive ectoderm.

The endpoint of *Icm1* expression was not analysed during peri-implantation development, which is difficult to study due to the small size and inaccessibility of this stage of embryogenesis. However, the similarity of *Icm1* and *Psc1* regulation *in vitro*, in cell types representative of lineages between the ICM and primitive ectoderm, suggests that *Icm1* may also be expressed in the peri-implantation epiblast bud and potentially downregulated at a similar point during proamniotic cavitation. Interestingly, *Psc1* was not co-expressed with *Icm1* in pluripotent cell types prior to the late stage ICM, indicating that coordinated regulation of these genes does not occur during these stages. This was not predictable from the expression of *Icm1* and *Psc1* *in vitro*, which was only relevant to lineages downstream of ICM/ES cells.

4.4.3 Correlation of *Icm1* and *Psc1* Expression to Embryonic Development

The expression and regulation of *Icm1* and *Psc1* in the pluripotent cells of early embryogenesis may give information about the functional roles of these genes. The cellular processes that *Icm1* and *Psc1* gene products are involved with may be indicated by correlating gene expression with specifically timed embryonic events.

The expression of *Icm1* was tightly restricted to pluripotent cells during preimplantation development. Critically, *Icm1* expression was downregulated upon cellular differentiation, to both the polar and mural trophoctoderm and to the visceral layer of primitive endoderm. Therefore, *Icm1* function may be required for pluripotent cell “maintenance versus differentiation” decisions. In this way, *Icm1* may specify pluripotency in the ICM and in ICM primordia, by maintaining the undifferentiated state of these cells. Alternatively, downregulation of this gene may be associated with lineage restriction during trophoctoderm or primitive endoderm differentiation. These postulated functions correlate well with *Icm1* expression *in vitro*.

The expression of *Psc1* correlated with several fundamental embryological events and potential functions during pluripotent cell development. The timing of primitive endoderm differentiation, from ICM cells lining the blastocoel, has not been precisely defined but occurs during mid to late stage blastulation (1.2.3). Thus the induction of *Psc1* in the late stage ICM may indicate a pluripotent cell type that has acquired endoderm differentiation potential and differentiates if contacting the blastocoel. Conversely, *Psc1* could be induced in ICM cells once primitive endoderm differentiation has occurred, restricting further differentiation and maintaining pluripotency. It may be possible to use the expression of *Psc1* as a marker to identify and follow the developmental signals that ICM cells are responding to during this stage of embryogenesis. Examples of these may be inductive signals from the differentiated cells of the trophoctoderm or primitive endoderm that trigger key events in pluripotent cells during peri-implantation development.

The expression of *Psc1* during blastulation demonstrated a clear difference between the early (*Psc1*⁻) and late (*Psc1*⁺) stage ICM. This is consistent with previous reports of a developmental maturation of ICM cells from the mid to late stage blastocyst (1.5.1, 1.5.2), and could indicate a general alteration in the state of ICM cells associated with progression to

the peri-implantation subpopulation. Furthermore, a poorly defined transition occurs during these stages of embryogenesis, where development shifts from a cleavage based system to active proliferation of the epiblast and growth of the embryo. Thus *Psc1* expression may mark pluripotent cell types that have acquired first, the capacity to proliferate and consequently, the defining stem cell property of self renewal. *Psc1* could therefore be critical for generation of the first true pluripotent stem cell type during embryogenesis, with potential involvement in the induction of stem cell proliferative capacity. Massive proliferation of the epiblast accompanies implantation, and it protrudes into the former blastocoel as a solid epiblast bud. Therefore, along with potential induction or response to induction of proliferative capacity, *Psc1* may have functions involved with the active proliferation of the peri-implantation epiblast.

Finally, *Psc1* was downregulated during the earliest stages of proamniotic cavitation. This may indicate functional roles in the processes that convert the solid epiblast bud into the cellular monolayer of the early primitive ectoderm, a transition that establishes structural precursors for gastrulation. These functions may include involvement with selective apoptosis and survival mechanisms that are thought to operate during cavitation (1.4.1). Indeed, the expression of *Psc1* in the peri-implantation embryonic ectoderm was consistent with a cellular response to a predicted visceral endoderm derived “death signal”, that potentially induces apoptosis during proamniotic cavitation (Coucovanis and Martin 1995).

In addition to potential functions during pluripotent cell development, the expression of *Psc1* in the ectoplacental cone at approximately 5.5 d.p.c. indicates that this gene could have a role during the generation of the placenta.

Analysis of predicted *Icm1* and *Psc1* functions during pluripotent cell development are likely to involve homologous recombination approaches in ES cells. In this way, single allele modifications of *Icm1* or *Psc1* may be reintroduced into mice and bred to homozygosity to assess developmental defects. Double allele modification of ES cells may enable the analysis of function during pluripotent cell development *in vitro*. Approaches to the investigation of *Icm1* and *Psc1* function are discussed further in Chapter 6.

4.4.4 Embryonic Equivalents of ES Cells and X Cells

The identification of a peri-implantation pluripotent cell subpopulation by *Psc1* enables more precise definition of the embryonic equivalents of ES cells and X cell types and validates the ES cell model of pluripotent cell development *in vitro*. Previous definition of the embryonic equivalent of ES cells has depended on the developmental window within which ES cell lines could be derived, from the late stage blastocyst or early epiblast (1.5.1, 1.6.2), and limited gene expression information (1.6.2). *Psc1* expression indicates that the embryonic equivalents of ES cells reside from the late stage ICM to the epiblast bud, from approximately 4.0/4.5 d.p.c. to 5.0 d.p.c (Fig. 4.11). The embryonic equivalent of *Psc1*⁺ “early” X cells therefore also lies within this pluripotent subpopulation (Fig. 4.11). Further subdivision of ES and “early” X cell equivalents within these stages will depend on markers that are capable of individually recognising these cell types (S. Sharma, unpublished data).

The feature associated with the downregulation of *Psc1* expression was identified as the onset of proamniotic cavitation. This could therefore mark the developmental point at which the embryonic equivalents of ES and “early” X cells develop to a downstream stem cell subpopulation, the early primitive ectoderm. The embryonic equivalents of *Psc1*⁻ “late” X cell types may therefore also be redefined, as existing from approximately 5.25 d.p.c., or immediately following the onset of proamniotic cavitation (Fig. 4.11). This refinement of X cell equivalents also correlates with the expression of *Rex-1* and *FGF-5* during X cell development (1.6.8) and during the ICM to primitive ectoderm transition *in vivo* (1.5.2). *Rex-1* is expressed by ES cells, “early” X cells and in the ICM, but *Rex-1* expression is downregulated in “late” X cells and in the primitive ectoderm. *FGF-5* expression is induced in “late” X cells and is first detected in the early primitive ectoderm at 5.25-5.5 d.p.c. Importantly, these results refine the embryonic equivalents of ES and X cells and validate the *in vitro* model of pluripotent cell development (Fig. 4.11).

ES cells have become a valuable academic resource for the introduction of specific alterations into the mouse germ line (1.6.1). However, the ability to isolate efficiently germ line competent pluripotent stem cells from mice appears to be strain-dependent, and possibly limited to the 129/Sv strain (reviewed in Smith 1992). This suggests that a specific genetic component is critical to generate ES cell lines, or that variation in the developmental timing of

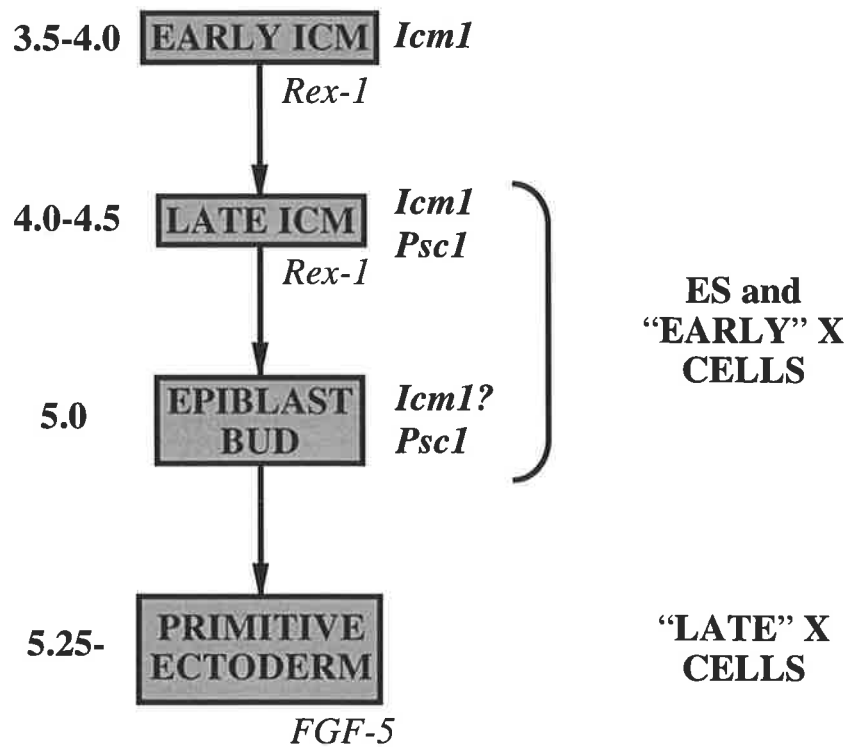


Figure 4.11 Embryonic Equivalents of ES Cells and X Cells.

Oct-4⁺ pluripotent cell populations are boxed. The embryonic equivalents of ES cells and "early" X cells in the late stage ICM and epiblast bud are indicated. The embryonic equivalents of "late" X cells in the primitive ectoderm following proamniotic cavitation are indicated. Expression patterns of *Icm1* (4.3.2), *Psc1* (4.3.3), *Rex-1* (1.5.2) and *FGF-5* (1.5.2) are indicated in italics. Timing of embryonic pluripotent cell populations and gene expression are shown to the left in days post coitum (d.p.c.). *FGF-5* is not expressed in the epiblast bud. The expression of *Icm1* and *Rex-1* in the epiblast bud stage has not been investigated.

different mouse strains may account for strain-dependency. Therefore, the observation that ES cell equivalents exist during peri-implantation development as well as in the late stage blastocyst may have important consequences for ES cell isolation methodologies. The exact embryological mapping of this peri-implantation stem cell subpopulation may assist the isolation of ES cell lines from non-129Sv strains. As such, improved ES cell isolation methods may require harvesting of peri-implantation embryos instead of blastocysts. Additionally, the precise manipulation of the genome of commercially important species, such as rats or livestock, has been hampered by the inability to isolate pluripotent stem cell lines from these species. Definition of the temporal boundaries of pluripotent cell subpopulations in the mouse and characterisation of related stem cell subpopulations in other mammals, perhaps via the isolation of species specific *Psc1* homologues, may play a vital role in the isolation of stem cell lines from these species. The isolation of pluripotent stem cell lines could enable the advanced genomic manipulation of these species, with obvious industrial, commercial or medical benefits.

4.4.5 Nomenclature of Pluripotent Cell Subpopulations

The expression of *Icm1* and *Psc1* indicates that the pluripotent cell pool has a much higher level of complexity than is currently understood from differentiation, transplantation and gene expression analysis (1.5.1, 1.5.2). Future identification and analysis of pluripotent cell subtypes is likely to provide information which will enhance the prevailing embryologic nomenclature. For example, the term “epiblast” is used for the pluripotent cells of the fully expanded blastocyst to the late egg cylinder stage embryo and does not take into account the described heterogeneity within these populations. The term “periblast” is therefore suggested on the basis of results described in this thesis, to offer a distinction between epiblast cell types, specifically describing peri-implantation *Psc1*⁺ cell types prior to proamniotic cavitation.

4.4.6 Potential for Identifying Pluripotent Cell Subpopulations

The analysis of expression during early embryogenesis has therefore validated the approaches taken to generate markers for pluripotent cell subpopulations or stem cell “states”

(1.5.1). This indicates that further analysis of pluripotent cell types *in vitro* may generate a comprehensive series of markers, capable of dissecting the molecular events and processes that govern pluripotent cell development (1.7). Such markers are likely to identify further restricted pluripotent cell populations *in vivo* and significantly enhance our understanding of pluripotent cell maintenance and proliferation in the pre-gastrulation mouse embryo.

CHAPTER 5

Isolation of Psc1 cDNA Clones and Analysis of Psc1 Expression During Later Development

CHAPTER 5

ISOLATION OF *Psc1* cDNA CLONES AND ANALYSIS OF *Psc1* EXPRESSION DURING LATER DEVELOPMENT

5.1 Introduction

Icm1 (A03/360) and *Psc1* (B04/400) were two novel cDNA markers isolated for ES and early pluripotent cell populations *in vitro*. Expression analysis of *Icm1* and *Psc1* during early embryogenesis indicated that they were restricted to overlapping but different pluripotent cell populations, suggesting potential functional roles for these genes during the development of pluripotent cell populations *in vivo*. Further characterisation of *Icm1* or *Psc1* by molecular cloning and more extensive expression analysis therefore had the potential to indicate the cellular or developmental roles of these genes, the protein domains that they encode, or their relationship to other genes.

Psc1 was selected for further characterisation on the basis of several interesting features. Firstly, *Psc1* expression indicated several possible functional activities during pluripotent cell development, ranging from involvement in stem cell maintenance and differentiation *in vitro*, to primitive endoderm differentiation, stem cell proliferation, or proamniotic cavitation *in vivo* (4.4.3). Secondly, analysis of the *Psc1* DDPCR product (B04/400) indicated a potential relationship to an uncharacterised *C. elegans* protein CLEB0336.3 10. Additional analysis of *Psc1* transcripts was required to define these genes as members of a novel gene class or to identify common functional domains.

Psc1 expression during early embryogenesis was restricted to pluripotent cells during peri-implantation development and extraembryonic cell types of the ectoplacental cone. The investigation of expression during later development would indicate if *Psc1* function was specific to early development, or whether expression could be detected at other embryonic or extraembryonic sites. RNase protection was used as a sensitive screen for *Psc1* expression in specific tissue samples isolated during later embryogenesis.

This chapter describes the isolation and characterisation of *Psc1* cDNA clones likely to represent either the 3.5 kb or 3.7 kb *Psc1* transcripts, and the further characterisation of *Psc1*

expression during embryogenesis. The results presented in this chapter indicate that the *Psc1* protein could be localised to subdomains within the nucleus, and that *Psc1* activity could be a component of a recurring developmental function.

5.2 Isolation and Sequence Analysis of *Psc1* cDNA Clones

5.2.1 Isolation of *Psc1* cDNA Clones

A λ Zap II cDNA library was generated from D3 ES cells (Doetschman *et al.* 1985). Separate reverse transcription reactions of D3 poly(A)_n⁺ RNA were primed with either oligo d(T) or random primers. The first strand cDNA was pooled for second strand synthesis, *EcoRI* linker addition, *EcoRI* digestion and ligation into λ Zap II (Clontech). A total of 5 x 10⁵ plaques were screened using the 458 bp *Psc1* (B04/400) DDPCR clone (3.4.2), generating two third round positive plaques. cDNA inserts were "Zapped" into pBluescript SK by helper phage super infection (2.3.15), followed by restriction mapping and sequencing of the termini with T3 and T7 primers. The sizes of the *Psc1* clones were 1084 bp and 2611 bp, with the smaller clone residing within the span of the 2611 bp *Psc1* cDNA clone (not shown). This indicated that full length *Psc1* clones were not isolated (Fig. 5.1), as *Psc1* transcript sizes were determined to be 5.5, 3.7 and 3.5 kb by northern blot (Figs. 3.9A, 3.10C).

The 2611 bp *Psc1* cDNA clone was sequenced in both directions. The 2611 bp insert did not contain an *EcoRI* site at the 3' end, as expected from library construction and was therefore isolated by digestion with *Bam*HI and *Hind*III. The *Bam*HI/*Hind*III *Psc1* insert was digested with *Alu*I, *Hae*III or *Rsa*I and subcloned into pBluescript KS to generate a series of shorter overlapping fragments (2.3.12). Subcloned fragments were sequenced in both directions with T3 and T7 primers and an overlapping contiguous sequence was aligned for the 2611 bp *Psc1* insert (2.3.36). Regions of the contiguous sequence that contained compressions or were spanned in only a single direction, were sequenced with 7-deaza-dGTP reactions (2.3.13iii), by the generation of additional subclones (2.3.12), or using custom oligonucleotides designed against known sequence. Sequence analysis demonstrated that the

Figure 5.1 *Psc1* Transcript Cloning and Sequencing Strategy.

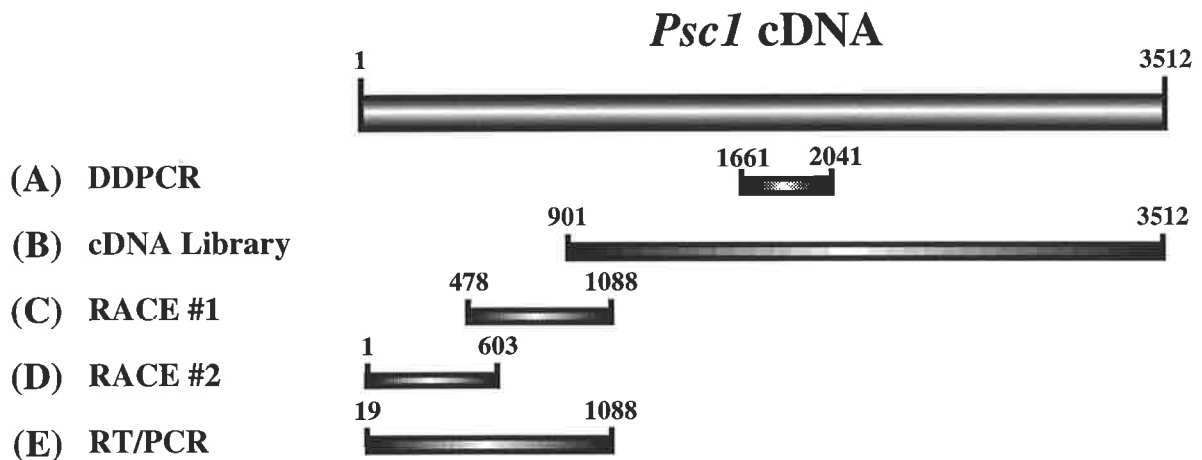
Schematic representation of the sizes, relative positions and generation of *Psc1* cDNA clones, compared to the compiled 3512 bp *Psc1* transcript.

(A) The *Psc1* 3'-hp DDPCR clone (3.4, 3.4.2) spanned positions 1661-2041, disregarding primer sequences.

(B) A 2611 bp *Psc1* cDNA (spanning positions 901-3512) was isolated from a D3 ES cell library by hybridisation to the *Psc1* DDPCR clone. The *Psc1* library clone was sequenced in both directions by assembling an overlapping contiguous sequence of short fragments (2.3.12), generated by *AluI*, *HaeIII* or *RsaI* digestion, subcloning and sequencing with T3 and T7 primers. Regions of the contiguous sequence that contained compressions or were spanned in only a single direction, were sequenced with 7-deaza-dGTP reactions (2.3.13iii), by the generation of additional subclones (2.3.12), or with custom oligonucleotides designed against known sequences. The 1084 bp *Psc1* cDNA library clone (5.2.1) spanned positions 1175-2259 (not shown).

(C) Sequence derived from the *Psc1* DDPCR and library clones was used to generate primers to amplify *Psc1* 5' regions by amplifINDER RACE/PCR (2.3.19). D3 ES cell (A)_n⁺ RNA was reverse transcribed (2.3.16i) with the 3' *Psc1* primer 1736 (2.2.8), corresponding to nucleotide positions 1736-1755. Anchor ligated cDNA (2.3.19) was amplified using the Anchor primer (2.2.8) and the *Psc1* 3' oligonucleotide 1064 (2.2.8), that spanned positions 1064-1088 and contained an *EcoRI* site to facilitate cloning. The *Psc1* cDNA clone RACE #1 was isolated by *EcoRI* digestion and ligation to pBluescript KS *EcoRI*. The RACE #1 clone was sequenced with custom oligonucleotides designed against known sequences, or by subcloning and sequencing with T3 and T7 primers. The RACE #1 clone was sequenced in both directions and spanned positions 478-1088.

(D) Sequence derived from the *Psc1* RACE #1 clone was used to generate primers to amplify additional 5' regions by amplifINDER RACE/PCR (2.3.19). D3 ES cell (A)_n⁺ RNA was reverse transcribed (2.3.16i) with the 3' *Psc1* primer 1275 (2.2.8), corresponding to nucleotide positions 1275-1294. Anchor ligated cDNA (2.3.19) was amplified using the Anchor primer (2.2.8) and the *Psc1* 3' oligonucleotide 582 (2.2.8), that spanned positions 582-603 and contained an *EcoRI* site to facilitate cloning. The *Psc1* cDNA clone RACE #2 was isolated by *EcoRI* digestion and ligation to pBluescript KS *EcoRI*. The RACE #2 clone was sequenced in both directions by T3 and T7 primers or custom oligonucleotides designed against known sequences and spanned positions 1-603.



(E) *Psc1* cDNA clones RACE #1 and RACE #2 contained several sequence ambiguities in their overlapping region that were likely to represent Taq polymerase errors. To confirm the sequence of the *Psc1* 5' region, RT/PCR clones spanning positions 19-1088 were generated. D3 ES cell (A)_n⁺ RNA was reverse transcribed (2.3.16ii) with oligo 1275 (2.2.8 and (C)). cDNA was amplified with the 1064 primer (2.2.8 and (C)) and the *Psc1* 5' primer (2.2.8), that spanned positions 19-36 and contained an *EcoRI* site to facilitate cloning. PCR products were digested with *EcoRI* and ligated to pBluescript KS *EcoRI*. Three clones were isolated from independent amplification reactions and were sequenced in both directions using T3 and T7 or custom oligonucleotides.

Psc1 open reading frame (ORF) extended to the 5' end of the 2611 bp cDNA, revealing that this clone was incomplete at the 5' end (Fig. 5.1, 5.2).

Sequence derived from the 2611 bp *Psc1* clone was used to design primers for the isolation of additional 5' regions by RACE/PCR (2.3.19) and RT/PCR (2.3.17) as described in figure 5.1. cDNA clones amplified from D3 ES cell poly(A)_n⁺ RNA by this approach contained a 5' untranslated region and an ATG codon at the start of the *Psc1* ORF (Figs. 5.1, 5.2).

5.2.2 Sequence Analysis of *Psc1* cDNA Clones

The compiled sequence of the *Psc1* cDNA and the deduced amino acid sequence of the *Psc1* protein are presented in figure 5.2. The *Psc1* cDNA was 3512 bp with an ORF that consisted of 1005 amino acids, from position 157 bp to 3171 bp. The presence of two in-frame stop codons (light underline, Fig. 5.2) at positions 28 bp (tag) and 85 bp (tga) confirmed the identification of the ATG initiation codon. The *Psc1* 3' untranslated region (3172-3512 bp) did not contain a poly(A) tail or a AATAAA poly(A) addition signal, although a non-consensus sequence GATAAA was identified at position 3501 bp to 3506 bp (double underline, Fig. 5.2). This could indicate that the *Psc1* cDNA is not quite full length at the 3' end.

Computer analysis of the translation product of *Psc1* did not reveal the presence of known protein motifs, although several features of interest were observed. The *Psc1* ORF was predicted to be highly basic with a [pI] of 9.5, with arginine (6.7%), histidine (2.8%) and lysine (8.2%) contributing 17.7% of all residues. Proline contributed 9.6% of residues and *Psc1* contained a proline rich region (30% of residues) from position 272 to 417 (proline residues of this region in bold, Fig. 5.2). Proline rich regions have been implicated as protein-protein interaction interfaces and transcription activation domains (Williams and Tjian 1991). *Psc1* was also relatively serine rich at 10.2% of all residues.

The *Psc1* protein contained three regions of different dipeptide repeat sequences (dashed underline, Fig. 5.2). A serine-arginine rich sequence at positions 164-187 contained 6 serine-arginine repeats; 11 proline-glycine repeats were located at positions 337-358; and 6

Figure 5.2 Compiled Nucleotide Sequence of *Psc1*.

The 1005 amino acid *Psc1* open reading frame spanned positions 157 to 3171. In-frame termination codons (light underline) upstream of the initiation codon were located at nucleotide positions 28 (tag) and 85 (tga). The non-consensus GATAAA poly(A) addition signal was located at positions 3501 to 3506 (double underline). Boxed regions exhibited similarity to *C. elegans* CLEB0336.3 10 (residues 1 to 79, 277 to 330 and 535 to 622) and human expressed sequenced tag HFBDS04 (residues 835 to 885), as indicated in figure 5.3. The proline rich region spanned amino acid positions 272 to 417 (prolines of this region in bold). Potential bipartite nuclear localisation signals were located at amino acid positions 114 to 138 and 759 to 776 (underlined). Regions containing dipeptide repeat sequences (dashed underline) were located at residues 164 to 187 (serine-arginine), 337 to 358 (proline-glycine) and 897 to 910 (glycine-arginine). The termination codon was at position 3172 (asterisk). The compiled *Psc1* sequence was generated as described in figure 5.1: nucleotide positions 1 to 36 (RACE #2 clone), positions 36 to 900 (RT/PCR clones) and positions 901 to 3512 (2611 bp cDNA library clone).

gttctgttaggccccggccagggagtaggtgaagtctataagatgcccggtgggtcagggcactcggagctgtga 77

gggaacgtgaggcgccggaacgtagaccgcggcgcccttaggaggagcgccgagccgccttcctgcacc 156

ATG CTC ATA GAA GAT GTG GAT GCC CTC AAG TCC TGG CTG GCC AAG TTA CTG GAG CCG ATA 216
M L I E D V D A L K S W L A K L L E P I 20

TGT GAT GCT GAC CCT TCA GCC CTA GCC AAC TAT GTT GTA GCA CTG GTC AAG AAG GAC AAA 276
C D A D P S A L A N Y V V A L V K K D K 40

CCT GAG AAG GAA TTG AAA GCC TTT TGT GCT GAC CAA CTT GAT GTC TTT TTA CAA AAA GAG 336
P E K E L K A F C A D Q L D V F L Q K E 60

ACT TCA GGT TTT GTG GAC AAA CTA TTT GAA AGT CTT TAT ACT AAG AAC TAC CTT CCA CCA 396
T S G F V D K L F E S L Y T K N Y L P P 80

TTG GAA CCA GTT AAA CCT GAG CCA AAA CCA CTA GTC CAA GAA AAA GAA GAA ATT AAA GAA 456
L E P V K P E P K P L V Q E K E E I K E 100

GAG GTA TTT CAA GAG CCA GCA GAA GAA GAA CGA GAT ACT AGA AAG AAG AAA TAT CCT AGT 516
E V F Q E P A E E E R D T R K K K Y P S 120

CCT CAG AAG AGT CGT TCA GAA TCT AGT GAA CGA AGA ACA CGT GAG AAG AAA AGA GAA GAT 576
P Q K S R S E S S E R R T R E K K R E D 140

GGC AAG TGG AGA GAC TAT GAG CGG TAC TAC GAG CGG AAT GAG CTG TAT CGG GAG AAG TAT 636
G K W R D Y E R Y Y E R N E L Y R E K Y 160

GAC TGG AGA AGG GGC AGG AGC AAG AGC AGG AGT AAG AGC CGA GGC TTG AGT CGG AGT AGA 696
D W R R G R S K S R S K S R G L S R S R 180

AGC CGG AGC AGG GGC CGC AGC AAA GAC CGG GAT CCA AAC AGG AAC GTT GAG CAC AGG GAA 756
S R S R G R S K D R D P N R N V E H R E 200

AGA TCA AAG TTT AAG AGC GAA AGA AAT GAC CTG GAG AGC TCT TAT GTG CCT GTA TCT GCT 816
R S K F K S E R N D L E S S Y V P V S A 220

CCA CCT CCA AGC TCT TCT GAG CAG TAT TCC TCT GGG GCA CAG TCT ATT CCC AGC ACT GTC 876
P P P S S S E Q Y S S G A Q S I P S T V 240

ACT GTG ATT GCA CCT GCG CAC CAT TCT GAG AAC ACA ACT GAG AGC TGG TCT AAT TAC TAT 936
T V I A P A H H S E N T T E S W S N Y Y 260

AAC AAT CAT AGC TCT TCC AAT TCT TTT GGT CGA AAC CCA CCA CCA AAG AGG CGC TGC AGA 996
N N H S S S N S F G R N P P P K R R C R 280

GAT TAC GAT GAA AGA GGA TTT TGT GTA CTT GGT GAC CTT TGT CAG TTT GAT CAT GGA AAT 1056
D Y D E R G F C V L G D L C Q F D H G N 300

GAT CCT CTA GTT GTT GAT GAA GTT GCT CTG CCA AGT ATG ATT CCT TTC CCA CCC CCT CCT 1116
D P L V V D E V A L P S M I P F P P P P 320

CCT GGG CTT CCT CCT CCA CCA CCT CCC GGA ATG TTA ATG CCT CCA ATG CCA GGT CCA GGC 1176
P G L P P P P P P G M L M P P M P G P G 340

CCA GGC CCT GGT CCT GGC CCT GGC CCT GGT CCC GGC CCA GGC CCA GGC CCT GGT CAT AGT 1236
P G P G P G P G P G P G P G H S 360

ATG AGA CTT CCT GTT CCT CAA GGA CAT GGT CAG CCT CCA CCT TCT GTT GTG CTT CCC ATA 1296
M R L P V P Q G H G Q P P P S V V L P I 380

CCA AGA CCT CCC ATT TCA CAG TCA AGT TTA ATA AAC AGC CGT GAC CAG CCT GGG ACA AGT 1356
 P R P P I S Q S S L I N S R D Q P G T S 400

GCA GTG CCC AAT CTT GCA CCA GTG GGA GCA AGA CTA CCT CCT CCT TTA CCG CAG AAC CTC 1416
 A V P N L A P V G A R L P P P L P Q N L 420

CTT TAT ACA GTA TCA GAA AAT ACA TAT GAA CCA GAT GGC TAT AAT CCT GAA GCA CCT AGC 1476
 L Y T V S E N T Y E P D G Y N P E A P S 440

ATC ACC AGT TCT GGT AGA TCA CAG TAC AGA CAG TTC TTC TCA AGA GCA CAG ACA CAG CGT 1536
 I T S S G R S Q Y R Q F F S R A Q T Q R 460

CCC AAC CTG ATT GGC CTC ACA TCT GGA GAT ATG GAT GCA AAT CCA AGA GCT GCT AAT ATT 1596
 P N L I G L T S G D M D A N P R A A N I 480

GTC ATC CAG ACT GAA CCA CCA GTT CCT GTT TCA GTT AAT AGC AAC GTC ACC AGA GTA GTC 1656
 V I Q T E P P V P V S V N S N V T R V V 500

CTG GAA CCA GAG AGC CGA AAG AGA GCT ATT AGT GGT TTG GAG GGG CCA CTT ACA AAG AAG 1716
 L E P E S R K R A I S G L E G P L T K K 520

CCT TGG CTG GGA AAA CAA GGG AAC AAC AAT CAA AGT AAA CCA GGC TTC CTG AGG AAG AAT 1776
 P W L G K Q G N N N Q S K P G F L R K N 540

CAT TAT ACA AAC ACC AAA CTA GAG GTG AAA AAA ATC CCC CAG GAA TTG AAC AAC ATT ACC 1836
H Y T N T K L E V K K I P Q E L N N I T 560

AAG CTC AAC GAA CAC TTC AGC AAA TTT GGA ACC ATA GTT AAT ATC CAG GTT GCT TTT AAG 1896
K L N E H F S K F G T I V N I Q V A F K 580

GGT GAT CCA GAA GCA GCA CTC ATC CAA TAC CTT ACC AAT GAA GAG GCC AGG AAA GCC ATT 1956
G D P E A A L I Q Y L T N E E A R K A I 600

TCT AGC ACA GAA GCA GTT TTG AAC AAT CGG TTC ATT CGA GTC CTG TGG CAT AGA GAG AAC 2016
S S T E A V L N N R F I R V L W H R E N 620

AAT GAG CAA CCA GCA CTC CAG TCC TCA GCA CAG ATT CTC CTG CAG CAG CAA CAC ACC CTG 2076
N E Q P A L Q S S A Q I L L Q Q Q H T L 640

AGT CAC CTT TCA CAG CAA CAC CAT AGC CTC CCA CAG CAT CTT CAT CCG CAG CAG GTG ATG 2136
 S H L S Q Q H H S L P Q H L H P Q Q V M 660

GTA ACC CAG TCT TCC CCC TCA TCG GTC CAT GGA GGT ATC CAA AAG ATG ATG GGC AAA CCA 2196
 V T Q S S P S S V H G G I Q K M M G K P 680

CAG ACC TCT GGT GCC TAT GTT CTT AAC AAA GTC CCT GTT AAA CAC CGT CTT GGA CAT GCA 2256
 Q T S G A Y V L N K V P V K H R L G H A 700

AGT ACT AAC CAG AGT GAC ACG TCT CAC TTG CTA AAT CAG ACT GGT GGT TCT TCT GGG GAA 2316
 S T N Q S D T S H L L N Q T G G S S G E 720

GAT TGT CCG GTA TTT TCT ACT CCA GGC CAT CCA AAA ACA ATT TAC AGC TCT TCA AAC TTA 2376
 D C P V F S T P G H P K T I Y S S S N L 740

AAG GCA CCT TCA AAA CTT TGT TCA GGG TCT AAA TCT CAT GAT GTT CAA GAA GTT CTT AAG 2436
 K A P S K L C S G S K S H D V Q E V L K 760

AAG AAA CAG GAA GCA ATG AAG TTA CAG CAA GAC ATG AGG AAA AAG AAA CAA GAA ATG TTG 2496
K K Q E A M K L Q Q D M R K K K Q E M L 780

GAA AAG CAG ATA GAA TGC CAG AAG ATG CTG ATT TCT AAA CTA GAA AAG AAC AAA AAC ATG 2556
 E K Q I E C Q K M L I S K L E K N K N M 800

AAA CCA GAA GAA AGA GCA AAT ATA ATG AAG ACC TTG AAA GAA CTG GGA GAG AAG ATC TCA 2616
 K P E E R A N I M K T L K E L G E K I S 820

CAA TTG AAA GAT GAA TTA AAA ACA TCT TCT ACA GTC TCC ACA CCG TCC AAA GTA AAG ACA 2676
 Q L K D E L K T S S T V S T P S K V K T 840

AAA ACA GAG GCC CAG AAG GAG TTG TTG GAC ACA GAG TTG GAC CTC CAC AAG AGG CTC TCT 2736
K T E A Q K E L L D T E L D L H K R L S 860

TCA GGA GAA GAT ACA ACA GAA TTA CGT AAA AAG CTC AGT CAG TTA CAA GTT GAG GCT GCA 2796
S G E D T T E L R K K L S Q L Q V E A A 880

AGA TTA GGT ATT TTA CCT GTG GGG AGA GGA AAG ACC ATA TCC TCT CAA GGC CGA GGA AGA 2856
R L G I L P V G R G K T I S S Q G R G R 900

GGC AGG GGT CGA GGA AGA GGG AGG GGC TCA CTG AAT CAC ATG GTG GTG GAT CAC CGC CCC 2916
G R G R G R G R G S L N H M V V D H R P 920

AAA GCT CTC CCA GGT GGA GGA TTC ATT GAG GAG GAA AAA GAT GAG TTG CTT CAG CAT TTC 2976
 K A L P G G G F I E E E K D E L L Q H F 940

TCA GCT ACA AAC CAA GCA TCA AAA TTC AAA GAC CGT CGG CTA CAG ATC TCA TGG CAT AAG 3036
 S A T N Q A S K F K D R R L Q I S W H K 960

CCC AAG GTA CCA TCT ATA TCT ACT GAG ACC GAG GAA GAG GAA GTC AAA GAA GAG GAA ACA 3096
 P K V P S I S T E T E E E E V K E E E T 980

GAA ACC TCA GAT TTG TTT TTG CAT GAT GAT GAT GAT GAA GAT GAA GAT GAA TAT GAG TCT 3156
 E T S D L F L H D D D D E D E D E Y E S 1000

CGT TCG TGG CGA AGA TGA aatttgatgagtggtgtgtaatttttagaaatattgtttagaagaacaacttttt 3229
 R S W R R * 1005

agaattatttaaagaagtcaatgagccaaaaaattttttctcttcaacacagtaggtttaaggacagcaagttgctat 3308

tcaaacacatctcataattgtctatgatatcaaatcaccaaaggccgtgactttacacagttgtgaagatccacagaag 3387

tgactggattctctagaaccttaccttccacaccatctgtttctgctcttggtgctggattacaatgtaaagacaggg 3466

gttcagaaattttattttggattataaatctgctgataaaaatttc 3512

glycine-arginine repeats, with a glycine-serine interruption, were located at positions 897-910.

Two sequences similar to bipartite nuclear localisation signals (NLS; reviewed in Boulikas 1993) were detected in the Psc1 ORF (underlined, Fig. 5.2), at positions 114 to 138 (RKKK_{x16}REKKR) and 759 to 776 (KKK_{x10}RKKK). The potential NLS at positions 759 to 776 was similar to the bipartite NLS (KRK_{x11}KKKS_{KK}) of the human poly (ADP-ribose) polymerase, which is functionally required for the nuclear localisation of this protein (Schreiber *et al.* 1992).

5.2.3 Similarity of Psc1 with CLEB0336.3 10 and HFBDS04

Database comparison indicated that the Psc1 protein contained four regions of similarity to previously identified ORFs (boxed areas, Fig. 5.2); three to the predicted *Caenorhabditis elegans* gene CLEB0336.3 10 (Wilson *et al.* 1994; Accession No. U32305) at positions 1-80 (region I), 276-330 (region II) and 535-622 (region III); and one to the human expressed sequence tag HFBDS04 (Adams *et al.* 1993; Accession No. T06332) at positions 835-886 (region IV). *C. elegans* CLEB0336.3 10 represents a putative protein coding sequence, predicted from the sequence of the genomic cosmid clone B0336.3 and thus requires confirming experimental evidence. HFBDS04 is an incomplete 472 bp sequence of a human cDNA and represents an open reading frame (174 bp) with a 3' untranslated region (298 bp).

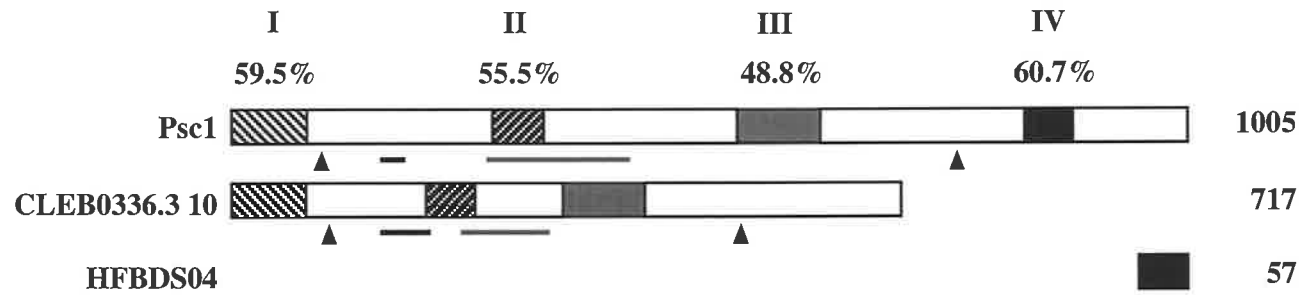
The structural arrangement and alignments of Psc1 sequence similarity with CLEB0336.3 10 and HFBDS04 are demonstrated in figure 5.3. Psc1 amino acid identity with CLEB0336.3 10 was 59.5% (region I), 55.5% (region II) and 48.8% (region III), while the intervening regions and C-termini were not of similar size (Fig. 5.3A) or sequence (data not shown). Similarity to Region IV was not observed in CLEB0336.3 10. Interestingly, the order of these motifs was conserved in Psc1 and CLEB0336.3 10, as were additional shared features, such as a proline rich region (grey underline), a serine-arginine dipeptide repeat region (black underline) and potential bipartite nuclear localisation signals (arrowheads). Specifically, CLEB0336.3 10 contained a proline rich region (30% of residues) from positions 248 to 323 (grey underline) and also contained 11 serine-arginine dipeptides dispersed within a 62 residue region (black underline, positions 149-211). Two potential

Figure 5.3 Comparisons and Alignment of Psc1 to CLEB0336.3 10 and HFBDS04 Open Reading Frames.

(A) Schematic alignment of Psc1, CLEB0336.3 10 and HFBDS04 open reading frames. Regions of similarity are numbered I to IV and specifically shaded. Amino acid identity between each region is indicated as percentage above the Psc1 protein. Psc1 and CLEB0336.3 10 contained a proline rich region (grey underline), a serine-arginine dipeptide repeat region (black underline) and potential bipartite nuclear localisation signals (arrowheads). Psc1 and CLEB0336.3 10 were predicted to be basic proteins ([pI] of 9.5 and 9.73 respectively). The size of each ORF is indicated to the right.

(B) Alignment of Psc1 regions I, II, III to CLEB0336.3 10 sequences and Psc1 region IV to the HFBDS04 sequence. Open reading frame amino acid positions are indicated to the left and right. Positions of identity are indicated by vertical lines, percentage amino acid identity for each region is indicated in (A). A sequencing ambiguity existed in HFBDS04 at position 29. Comparisons were carried out with the BLAST program (2.3.36).

A



B

I

| | | | |
|----------------------|---|---|----|
| Psc1 | 1 | MLIEDVDALKSWLAKLLEPICDADPSALANYVVALVKK-DKPEKELKAFCADQLDVFLQKETS | 79 |
| | | | |
| CLEB0336.3 10 | 1 | MHIDNEDALFDWLSDELSPI TDADPNALAKYVLALVKKPKDGHDELKAFTNEQLNVFLTDHTAPFVDKVF | 80 |

II

| | | | |
|----------------------|-----|---|-----|
| Psc1 | 277 | RRCRDYDERGFCVLGDLGQFDHGN | 330 |
| | | | |
| CLEB0336.3 10 | 210 | RRCKDFEERGYCIRGDQCPYYHGRDPVVVDENALGSMVPLPATAP-NFSLPPG | 263 |

III

| | | | |
|----------------------|-----|--|-----|
| Psc1 | 535 | GFLRKNHYTNTKLEVKKIPQELNITKLN | 622 |
| | | | |
| CLEB0336.3 10 | 347 | GFTGAINRDNCTLQVAKIPPEMNTIAKLNEHFATFGTV | 443 |

IV

| | | | |
|----------------|-----|-------------------------------------|-----|
| Psc1 | 835 | PSKVTKTEAQKELLDTELDLHKRLSSGEDTTELRK | 885 |
| | | | |
| HFBDS04 | 6 | PKSIKTKTQMQLDLKXMQAGEEVTTELRRKYTELQ | 57 |

bipartite nuclear localisation signals were located at positions 132 to 147 (RKR₁₀KEK) and 537 to 553 (DKK₁₀KKIK) within the CLEB0336.3 10 sequence (arrowheads). CLEB0336.3 10 was also predicted to be relatively proline (8.78%) and serine (9.62%) rich and a basic protein ([pI]=9.73). *Psc1* amino acid identity with HFBDS04 was 60.7% over 51 residues (region IV) and this region was located at the C-terminal regions of both these proteins (Fig. 5.3). The N-terminal 8 residues of HFBDS04 did not demonstrate any similarity to the corresponding *Psc1* sequence (data not shown). The sequence of other regions of HFBDS04 was not available for comparison.

5.3 *Psc1* Expression in Foetal and Adult Murine Tissues

To investigate the expression of *Psc1* at later stages of murine development and in adult tissues, RNase protection analysis (2.3.28) was carried out using an antisense riboprobe derived from the *Psc1* (B04/400) DDPCR clone (2.2.7ii). RNA samples from 10.5 d.p.c. and 16 d.p.c. murine embryos and RNA isolated from specific tissues of 16 d.p.c. murine embryos and adult mice, were screened by RNase protection (Fig. 5.4). The expression level of *Psc1* was quantitated by volume integration (2.3.35), normalised by comparison to the *GAPDH* loading control probe and defined as a percentage of D3 ES cell expression. *Psc1* was detected at 10.5 d.p.c and 16.5 d.p.c., indicating that *Psc1* was expressed during later embryogenesis. *Psc1* was differentially regulated between tissues at 16.5 d.p.c. and was expressed at a higher level in lung, limbs and brain, as compared to liver, kidney, heart, intestine and skin. *Psc1* was also differentially regulated between tissues in adult mice and was expressed highly in lung and placenta, and at a reduced level in liver, kidney, heart and muscle.

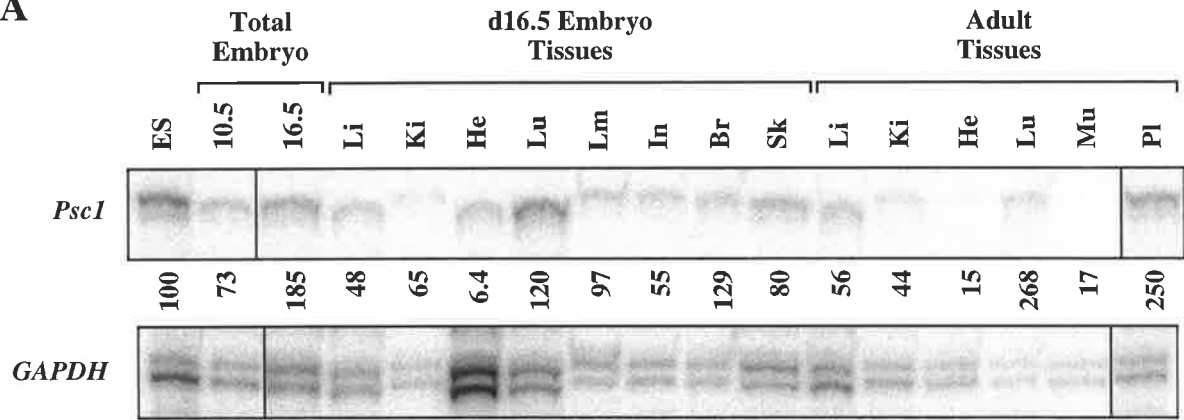
Analysis of *Psc1* during later development therefore indicated that expression was not restricted to peri-implantation pluripotent cells and the ectoplacental cone during early embryogenesis. Importantly, *Psc1* expression was differentially regulated between individual tissues at 16.5 d.p.c. and in the adult. For example, *Psc1* was expressed at a high level in the lung at 16.5 d.p.c. and in the adult, but was not expressed at a high level in the heart during these stages. The differential regulation of *Psc1* expression between individual tissues

Figure 5.4 Expression of *Psc1* During Later Development.

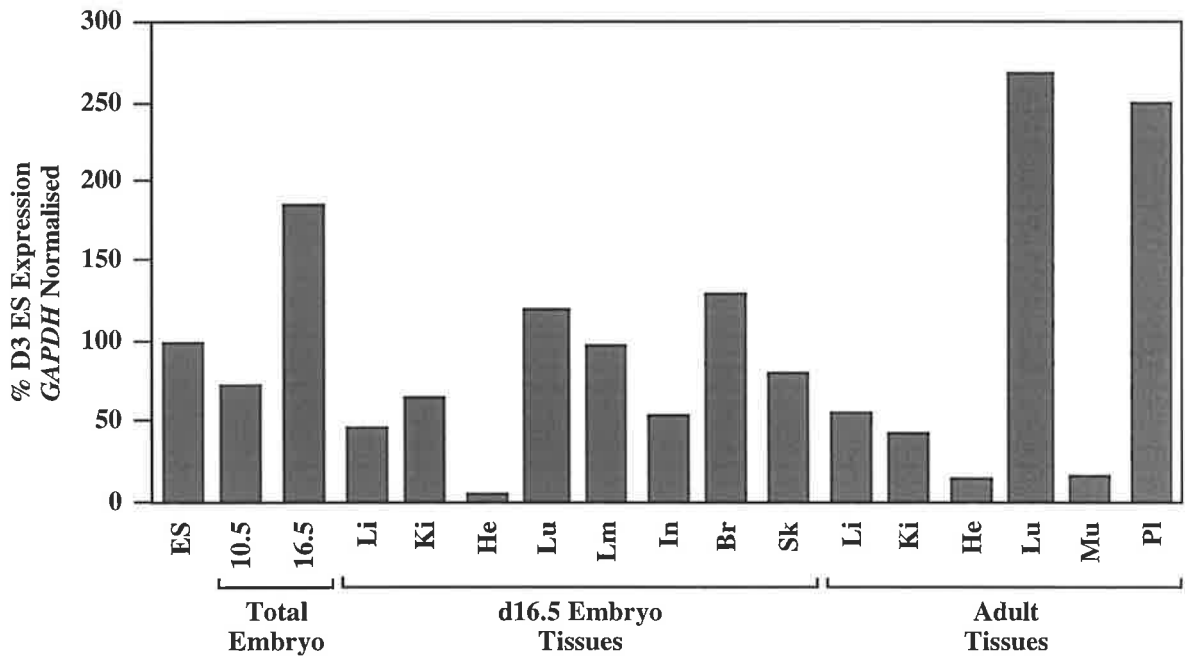
(A) *Psc1* expression was detected using an antisense riboprobe of the 458 bp B04/400 clone (2.2.7ii). RNase protection (2.3.28) was performed on 10 µg of total RNA from 10.5 and 16.5 d.p.c. CBA embryos, tissue specific RNA from 16.5 d.p.c CBA embryos and tissue specific RNA from adult CBA mice. Li, Liver; Ki, Kidney; He, Heart; Lu, Lung; Lm, Limbs; In, Intestine; Br, Brain; Sk, Skin; Mu, Muscle; Pl, Placenta. Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) was used as a loading control (2.2.7iv). *Psc1* signal was quantitated by volume integration (2.3.35), normalised to *GAPDH* expression and presented as a percentage of the D3 ES cell expression level.

(B) Graphic representation of *Psc1* expression levels.

A



B



indicates that *Psc1* is likely to have a developmental function during later stages of gestation and in the adult, particularly in the lung, brain and placenta.

The placenta has important functions during the growth and development of the embryo. These include the exchange of metabolites and nutrients between the mother and the foetus and the synthesis of hormones involved in coordinating maternal and embryonic physiology's (Hogan *et al.* 1994). Whole mount *in situ* hybridisation previously detected *Psc1* expression in the ectoplacental cone (EPC) at approximately 5.5-5.75 d.p.c. (Fig. 4.10). The EPC consists of trophoblast derived cell types that contribute to a large proportion of the placenta during gestation (Hogan *et al.* 1994). The expression of *Psc1* in the in the placenta at 16.5 d.p.c. confirms that *Psc1* could function during the generation and development of this specific tissue.

5.4 Discussion

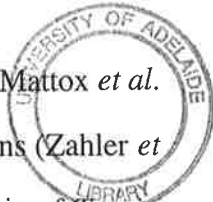
5.4.1 Analysis of *Psc1* Transcripts

cDNA isolation and analysis generated a 3512 bp *Psc1* sequence, encoding a 1005 residue ORF and representing either the 3.5 kb or 3.7 kb *Psc1* transcript. Generation of transcript specific probes from *Psc1* cDNA will be required to define which mRNA is represented by the isolated sequence. The *Psc1* transcript 3' terminus or a poly(A) addition signal were not identified in this analysis, indicating that the sequence was not full length. However, the presence of an non-consensus poly(A) signal (position 3501-3506) may indicate that the poly(A) tail lies adjacent to the isolated 3' UTR.

The detailed nature of the three *Psc1* transcript variants was not addressed in this analysis. Future investigation utilising RACE/PCR, RT/PCR and genomic analysis may be required to define *Psc1* mRNA differences resulting from alternate transcriptional or processing events.

5.4.2 Cellular Functions of *Psc1* and Nuclear Speckle Domains

Several features of the deduced *Psc1* protein were potentially indicative of protein function. Serine-arginine repeat sequences, termed SR or RS domains, are found in



regulators of pre-mRNA splicing such as *Drosophila Tra* and *Tra2* (reviewed in Mattox *et al.* 1992) and in general splicing factors such as the multigene family of SR proteins (Zahler *et al.* 1992) or the U1 snRNP protein, U1-70K (Spritz *et al.* 1987). The SR domain of *Tra*, in combination with a bipartite NLS, is capable of directing subnuclear localisation, to regions referred to as nuclear “speckles” (Hedley *et al.* 1995). Nuclear speckles are interchromatin granule regions predicted to be sites of spliceosome assembly or storage (Zhang *et al.* 1994). The minimal sequence required to direct *Tra* to nuclear speckle domains was predicted to be a bipartite NLS and a stretch of basic residues adjacent to at least three SR dipeptide repeats (Hedley *et al.* 1995). This minimal sequence was identified within the SR domains of several splicing related proteins that localise to nuclear speckles (Hedley *et al.* 1995). Analysis of the serine-arginine dipeptide repeats of Psc1 and CLEB0336.3 10 (5.2.2, 5.2.3) revealed regions that were similar to this minimal sequence and may therefore function as subnuclear localisation signals, in combination with their bipartite NLS. These were residues 163-187 of Psc1 (RRGRSKSRSKSRGLSRRSRSRGRS; basic residues underlined, SR dipeptides in bold) and residues 182-194 of CLEB0336.3 10 (RHSKRRSRSRSRS; basic residues underlined, SR dipeptides in bold).

The presence of this domain suggests that Psc1 is a nuclear localised protein and may be sublocalised to nuclear speckle regions. Nuclear speckles harbour spliceosomes and other factors responsible for pre-mRNA splicing (reviewed in Moore *et al.* 1993). This suggests that Psc1 may co-localise with other SR domain proteins to spliceosome containing regions of the nucleus. Psc1 did not contain characterised RNA binding domains, although close association with RNA may be possible, given the highly basic nature of the Psc1 protein. The proline rich region of Psc1 could enable protein-protein interactions within the nuclear speckle regions.

Psc1 contained four regions of cross-species similarity, three to the predicted *C. elegans* gene CLEB0336.3 10 (regions I, II and III) and one to the incomplete ORF of the human EST HFBDS04 (region IV). The similarity of regions I, II and III between Psc1 and CLEB0336.3 10, compared to the dissimilarity of intervening sequence, may define novel domains that are functionally required for the activity of these proteins. The isolation of additional genes with similar regions may reveal consensus sequences for these motifs and

demonstrate if they function independently or in combination. Interestingly, *Psc1* and CLEB0336.3 10 shared several features, such as the similarity and arrangement of regions I, II and III, proline rich regions, bipartite NLS and SR subnuclear localisation signals and their basic nature. *Psc1* and CLEB0336.3 10 may therefore define the founding members of a novel gene class and could perform similar cellular functions. This indicates that CLEB0336.3 10 function may also involve localisation to nuclear speckle regions.

Database comparison identified similar proline-glycine dipeptide repeats to those found in *Psc1*; in the human DNA-dependent ATPase transcriptional activator, BRG1 (Accession no. 440970); in the *Drosophila virilis* zinc finger protein, suppressor of sable (Su(s)) (Accession no. 671708); and in the human sperm membrane protein BS-84 (Accession no. 532731). Specific functions of proline-glycine repeat sequences have not been described.

5.4.4 Function of *Psc1* During Development

The prediction that *Psc1* is localised to nuclear speckle regions, which are rich in spliceosomes and splicing factors, may give some indication as to the developmental role of this gene. *Psc1* expression correlated with specific events during peri-implantation development such as pluripotent cell proliferation, differentiation and apoptosis (4.4.3). *Psc1* function may therefore regulate the pre-mRNA splicing of cellular factors that are critical for these processes. Regulation of pre-mRNA splicing is a common feature in the control of eukaryotic gene expression. For example, sex determination in *Drosophila melanogaster* requires the Transformer (*Tra*) and Transformer 2 (*Tra2*) splicing proteins which activate the female-specific splicing of doublesex (*dsx*) transcripts, leading to female sex determination (Ryner and Baker 1991). In a similar manner, *Psc1* may be an activator or repressor of splicing of a specific transcript or series of pre-mRNAs that control peri-implantation stem cell development.

Expression analysis indicated that *Psc1* was not restricted to the pluripotent cells of the peri-implantation embryo or the ectoplacental cone, but was differentially expressed in a range of tissues during organogenesis and in the adult. This indicated that *Psc1* was likely to be functional at multiple stages of embryonic development and in the adult. This was confirmed by the tissue specific regulation of *Psc1* expression; by elevated levels in 16.5 d.p.c. lung and

brain, adult lung and placenta; and reduced levels in 16.5 d.p.c. liver, kidney and heart, adult liver, kidney, heart and muscle. *Psc1* expression during peri-implantation development may therefore indicate a critical cellular decision or process, such as proliferation or apoptosis, rather than functions unique to implantation. Definition of the processes that are highlighted by *Psc1* expression will require detailed analysis of transcript and protein expression at the cellular level in multiple tissues during development.

5.4.5 Future Analysis of the Cellular Function of Psc1

The isolated *Psc1* cDNA clones will enable the design of approaches directed toward investigating the functional role of this protein. These are likely to include production of Psc1 protein and generation of specific antibodies. Anti-Psc1 antibodies will be crucial to functional analysis, allowing investigation of the predicted subnuclear localisation of Psc1 and potential interaction with components of the splicing machinery. Of particular interest would be to define whether Psc1 functions as a specific activator or repressor of splicing, and subsequently, which transcripts are regulated. These approaches are likely to be complimented by the isolation of *Psc1* genomic clones and allele disruption experiments *in vitro* and *in vivo* (Chapter 6).

CHAPTER 6

Final Discussion

CHAPTER 6

FINAL DISCUSSION

6.1 A System for the Isolation of Markers for Subpopulations of Murine Pluripotent Cells

The work in this thesis described the development of an integrated system for the generation of markers for early embryonic pluripotent cell subpopulations, based on gene expression profiles during the differentiation and development of embryonic stem cells *in vitro*. The system combined the identification and isolation of markers by differential display PCR with verification procedures and embryonic expression, as demonstrated by the generation and analysis of *Icm1* and *Psc1*. In this way, molecular heterogeneity between pluripotent cell types was identified *in vitro*, enabling subsequent identification of heterogeneity within the pluripotent stem cell pool *in vivo*.

6.2 Pluripotent Cell Subpopulations or “States” During Early Embryogenesis

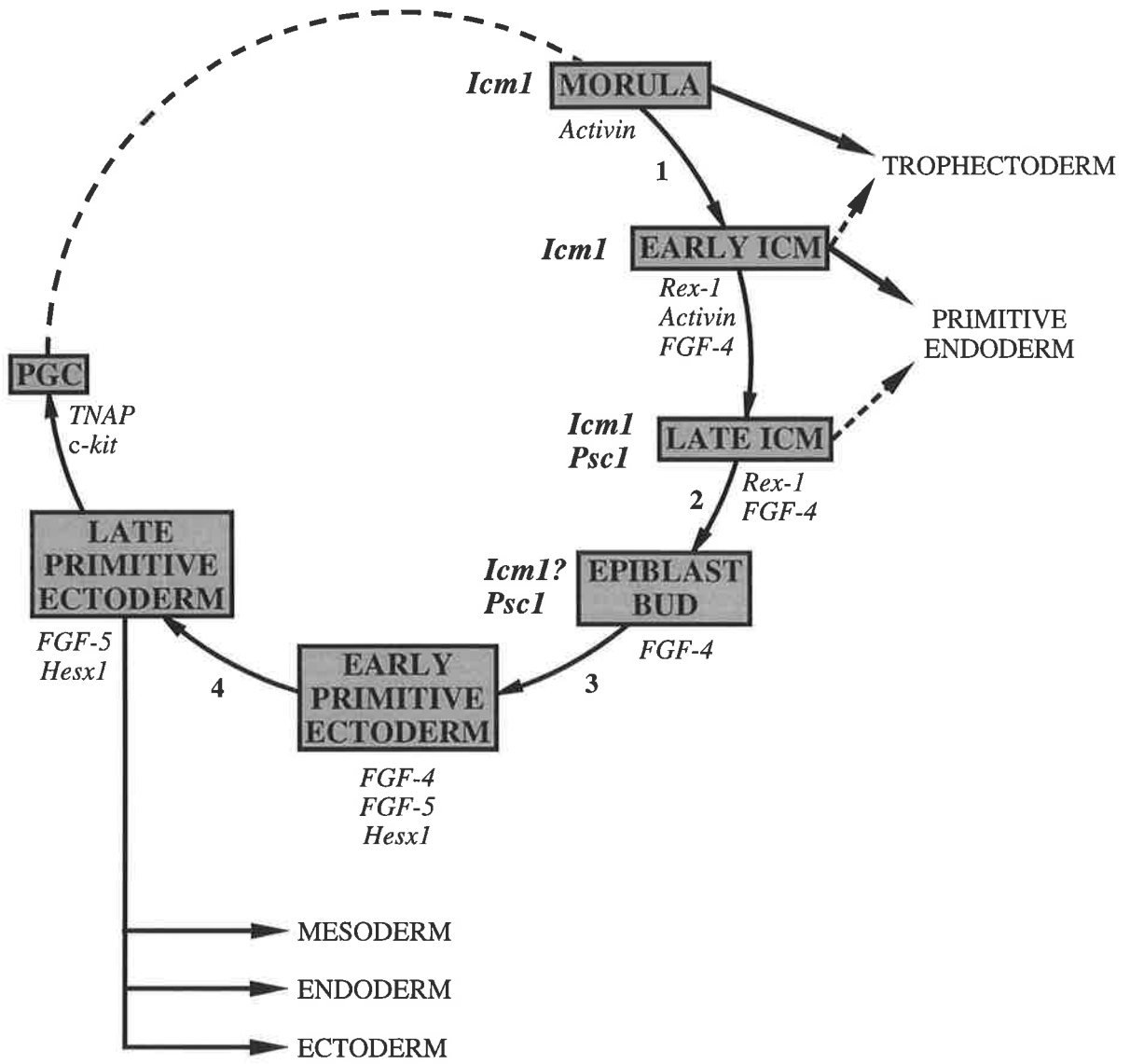
The distribution of *Icm1* and *Psc1* during early embryogenesis revealed the presence of pluripotent cell subpopulations within the *Oct-4*⁺ pool (4.4.2). Together with previous descriptions of molecular, developmental and structural heterogeneity, a model of specific stem cell subpopulations or “states” and their developmental transitions can be formulated for cell types between the morula and primordial germ cells (Fig. 6.1).

The prediction of distinct pluripotent cell subpopulations highlights the complexity within the developing stem cell pool and indicates that pluripotent cells develop through a series of highly regulated stages, all critical to murine embryogenesis. These subpopulations could be distinct cell types such as the ICM and primitive ectoderm, or could be subsets of cells within a homogeneous population that respond to different developmental signals. For example, specific regions of primitive ectoderm are induced to form particular tissues following gastrulation (Fig. 1.6). However, primitive ectodermal cell types are not committed and can reprogram developmental fate when transplanted to a different region of the primitive ectoderm (1.5.1). This indicates that different pluripotent cell “states” within the primitive

Figure 6.1 Heterogeneity and Developmental Progressions of Pluripotent Cell Subpopulations.

Subdivisions of early embryonic pluripotent cells between the morula and PGC, on the basis of gene expression (1.5.2), differentiation potential (1.5.1) and structural morphology (1.2-1.4.3). Pluripotent cell populations express *Oct-4*, SSEA-1 antigen and alkaline phosphatase (shaded boxes). Gene expression markers that are restricted during pluripotent cell development (1.5.2) are indicated by italics. Developmental or differentiation potential is shown by arrows and restriction in differentiation potential is shown by dashed arrows. Pluripotent cell types between PGCs and morula are indicated by the dashed line. Morphological distinctions between pluripotent cell subpopulations include blastocoel formation (1), ICM proliferation (2), proamniotic cavitation (3), and primitive streak formation (4).

The early ICM is distinguished from the morula by the induction of *FGF-4* expression, gain of the developmental capacity to generate primitive endoderm and gradual restriction in the capacity to differentiate to trophectoderm. The early ICM progresses the late ICM, inducing *Psc1* expression, downregulating the expression of Activin peptides and becoming restricted in the capacity to differentiate to primitive endoderm. The late stage ICM proliferates to form the *Psc1*⁺/*FGF-4*⁺ embryonic ectoderm of the epiblast bud stage embryo. Downregulation of *Icm1*, *Psc1* and *Rex-1* expression and induction of *FGF-5* expression distinguishes the early primitive ectoderm. Expression of *Hesx1* identifies an anterior subpopulation within the pluripotent primitive ectoderm. Additional "states" within the primitive ectoderm may be indicated by developmental fate (Fig. 1.6). Late stage primitive ectoderm undergoes mesoderm induction and the processes of gastrulation to form the three primary germ layers. *FGF-4* expression may delineate a posterior pluripotent cell subpopulation during these stages. Primordial germ cells are identified by the expression of tissue non-specific alkaline phosphatase (*TNAP*) and *c-kit*. Expression of *Rex-1* in the morula or epiblast bud has not been determined.



ectoderm do not represent distinct cell types but could be due to different responses to heterogeneous developmental signals or inductive events. A series of markers that define distinct pluripotent cell sub-types, or identify cell populations responding to temporally or spatially restricted inductive signals will therefore be a powerful accessory to deciphering the developmental progression of the pluripotent cell pool.

6.2.1 Identification of Pluripotent Cell Subpopulations or “States” During Early Embryogenesis

Expanded utilisation of the approach detailed in this work is likely to define a series of markers that delineate the complexity of the ES to X cell transition *in vitro*, and identify a progression of specific pluripotent cell subpopulations or “states” *in vivo*. Promising applications of this system have included the identification of a marker that is rapidly induced in Xp0 cells and is downregulated in Xp3 cells. Expression of this marker further subdivides “early” pluripotent cell populations *in vitro* into ES and “early” X cell components (S. Sharma unpublished data, 3.3.3). This marker therefore has the potential to distinguish pluripotent cell subpopulations during peri-implantation development (4.4.4).

“Late” X cells express the primitive ectoderm marker *FGF-5* and have downregulated expression of *Icm1*, *Psc1* and *Rex-1*, indicating that they represent pluripotent cell types of the epiblast after proamniotic cavitation. “Late” X cell types also express *FGF-4* (Bettess 1993), which is expressed in the early primitive ectoderm, but is restricted to the posterior region of primitive ectoderm at 6.25 d.p.c. (1.5.2). Thus the identification of markers that are specifically induced in “late” X cell populations (S. Sharma unpublished data, 3.3.3) could enable the definition of pluripotent cell “states” within the early primitive ectoderm or in posterior regions of the late primitive ectoderm (4.4.4).

6.3 Future Work and Application of *Icm1* and *Psc1* to the Analysis of Pluripotent Cell Development

The identification of *Icm1*, *Psc1* and additional markers of stem cell “states” is expected to provide valuable tools for the functional characterisation of pluripotent cell development (1.7). *Icm1* and *Psc1* could enable the analysis of pluripotent cell development through

several approaches, including investigation of mutations that affect pluripotent cells, following inductive events during embryogenesis, mutational analysis of these genes *in vitro* and *in vivo*, and the isolation and application of *Icm1* and *Psc1* promoters. These are described in more detail in the following sections.

6.3.1 Analysis of Mutations that Affect Pluripotent Cell Development

Unlike *Oct-4*, *Icm1* and *Psc1* divide the pluripotent cell pool into distinct subpopulations. These markers, in combination with *Rex-1* and *FGF-5* (1.5.2), could therefore be used to map specific stages of pluripotent cell development in mouse mutations that lead to peri-implantation pluripotent cell defects. This could include the analysis of naturally occurring or induced mutations, or in lines generated by gene targeting. Examples of this include the induced mutation *Hba^{th-J}*, which is lethal during early postimplantation development and exhibits ICM defects in blastocyst outgrowths (Hendrey *et al.* 1995), and the targeted mutation of *FGF-4* (Feldman *et al.* 1995), which exhibits early postimplantation lethality. The analysis of *Icm1* and *Psc1* expression during pluripotent cell development in these mutant embryos could indicate whether important developmental progressions occur, such as the generation of the *Psc1*⁺ late stage ICM or the formation of primitive ectoderm. In this manner it could be possible to refine the stage at which defects are observed and enable detailed characterisation of the developmental effects of these mutations.

6.3.2 Identifying Developmental Signals During Embryogenesis

The pluripotent cell subpopulation identified by *Psc1* could be different from the earlier and latter populations because it is responding to a developmental signal that is temporally restricted to this stage. Therefore the expression of *Psc1* could be used as a reporter to follow the origin and nature of this signal, enabling its identification and characterisation. For example, the expression of *Psc1* in the late stage ICM correlated with the differentiation of primitive endoderm (4.4.3). Therefore, the differentiated cells of the primitive endoderm could be emitting a developmental signal that induces *Psc1* expression in the late stage ICM. In this way *Psc1* expression could be used as a marker to investigate developmental processes such as cell-cell signalling, lineage restriction or cellular differentiation. Markers of other

pluripotent cell subpopulations or “states” could be used to follow inductive and developmental signals during embryogenesis. For example, markers of individual primitive ectoderm “states” could be used as reporters to follow and identify signals that specify cellular fate during gastrulation (1.5.1).

6.3.3 Analysis of Pluripotent Cell Development

Functional analysis of cell and developmental biology requires the construction of specific genetic alterations. It is expected that mutation of *Icm1* and *Psc1* by gene targeting will enable sophisticated analysis of pluripotent cell development, both *in vitro* and *in vivo*. This would initially require the isolation and characterisation of genomic clones for these genes and establishment of single allele modifications of *Icm1* or *Psc1* by homologous recombination in ES cells.

An appropriate approach could be the inactivation of *Icm1* or *Psc1* loci by insertion of an IRES- β geo cassette (Friedrich and Soriano 1991, Mountford *et al.* 1994). Firstly, this approach would enable Neo^R selection of recombinants driven from *de novo* promoter elements in ES cells and could lead to a high frequency of correctly targeted lines. Secondly, the utilisation of an IRES- β geo cassette would enable detailed mapping of *Icm1* or *Psc1* expression in heterozygous mice by β gal staining during later development. In particular, this may identify specific cell types that express *Psc1* in the lung, brain and placenta at 16.5 d.p.c., or highlight a common cellular process at the sites of *Psc1* expression (5.4.4).

6.3.3i Functional Analysis of Pluripotent Cell Development *in vitro*

The development of pluripotent cells during the ES to X cell transition and during ES cell differentiation *in vitro* provides a model for the decisions and development of the pluripotent cell pool *in vivo* (1.6.1, 1.6.9). Mutational approaches in ES cells may enable the functional analysis of *Icm1* or *Psc1* in these cell populations. *Icm1* and *Psc1* were expressed in ES cells and “early” X cell types, but expression of these genes was downregulated in “late” X cells and during cellular differentiation. This indicates that *Icm1* and *Psc1* could have functional roles during the development of early pluripotent cell types *in vitro*, such as maintenance of the pluripotent cell state or inhibition of cellular progression or differentiation.

The inactivation of both alleles of *Icm1* or *Psc1* in ES cells could therefore lead to disruption of pluripotent cell development *in vitro*. These approaches could be complimented by overexpression of *Icm1* or *Psc1* *in vitro*, which could also lead to alterations in pluripotent cell development. Disrupted pluripotent cell development could be monitored *in vitro*, by assaying gene expression, or alteration to the self-renewal or differentiation capabilities of pluripotent cells.

These analyses could enable the investigation of the functional roles of *Icm1* and *Psc1* during the progression and differentiation of pluripotent cells *in vitro* and therefore provide insight into the processes that regulate the development of the pluripotent cell pool *in vivo*.

6.3.3ii Mutational Analysis of *Icm1* and *Psc1* *in vivo* and Gene Function During Embryogenesis

ES cell lines with a single modified *Icm1* or *Psc1* allele could be reintroduced into the mouse by blastocyst injection. Chimeric mice could be onbred to obtain offspring homozygous for the mutant allele, to analyse *Icm1* or *Psc1* function during the development of pluripotent cells during early embryogenesis. Analysis of phenotypes in embryos lacking *Icm1* or *Psc1* could reveal critical events that require gene function and therefore enable analysis of pivotal stages of pluripotent cell development. These approaches could be complimented by ectopic expression of *Icm1* or *Psc1* during the equivalent stages of development. In this way, constructs designed to overexpress *Icm1* or *Psc1* could be introduced into the fertilised egg by microinjection and the corresponding developmental phenotypes analysed. Pluripotent cell or developmental defects that arise during the mutational analysis of *Icm1* or *Psc1* *in vivo* could be characterised by morphological and gene expression analysis.

Icm1 was expressed in pluripotent cells during preimplantation development and expression was downregulated during cellular differentiation to trophectoderm and primitive endoderm and prior to the formation of the primitive ectoderm. The function of *Icm1* during early embryogenesis could therefore involve regulation of pluripotency in the ICM and in ICM primordia, or could be associated with lineage restriction during trophectoderm or primitive endoderm differentiation (4.4.3). Embryos with altered *Icm1* expression could

exhibit dysfunction in the generation or maintenance of pluripotent cell types during preimplantation development, such as during the formation of the morula or ICM. Alternatively, morula or ICM cells could exhibit an altered propensity to differentiate, particularly to trophoctodermal or primitive endodermal lineages respectively. Peri-implantation defects could also be observed, prior to the formation of primitive ectoderm.

Psc1 was expressed in the late stage ICM and in the embryonic ectoderm during peri-implantation development and expression was downregulated at the onset of proamniotic cavitation. The function of *Psc1* during peri-implantation development could therefore involve primitive endoderm differentiation, stem cell proliferation, or proamniotic cavitation (4.4.3). Pluripotent cell defects expected in embryos with altered *Psc1* expression, range from effects on primitive endoderm differentiation, to peri-implantation dysfunctions such as altered proliferation, proamniotic cavitation, and primitive ectoderm formation.

It is expected that mutational analysis of *Icm1* and *Psc1* *in vivo*, coupled with *in vitro* analysis (6.3.3i), would therefore reveal the functional importance of these genes during pluripotent cell development and in cellular events during embryogenesis.

Psc1 was also expressed in the ectoplacental cone (EPC) at 5.5 d.p.c. and in the placenta at 16.5 d.p.c., indicating potential functions during the formation and development of the placenta. Disruption to EPC and placental development could therefore be observed in pregnancies of *Psc1* null mutant embryos, possibly leading to placental failure and embryonic lethality. Embryonic lethality due to placental failure is observed at 10.0 d.p.c. in a targeted mutation of the helix-loop-helix gene *Mash-2*, which is first expressed in the ectoplacental cone at 6.5 d.p.c. (Guillemot *et al.* 1995), and is essential for the development of extraembryonic lineages and the placenta (Guillemot *et al.* 1994). Similar phenotypes in *Psc1* null mutant pregnancies could indicate a role for *Psc1* in extraembryonic lineages and contribute to the functional analysis of placental development. Other developmental defects could be observed during mutational analysis of *Psc1*. For example, the expression of *Psc1* in embryonic lung and brain, and in adult lung, could suggest a functional role of *Psc1* in these tissues and point to corresponding developmental defects during mutational analysis.

6.3.4 Isolation and Application of *Icm1* and *Psc1* Promoters

Promoter regions that regulate *Icm1* or *Psc1* expression could be isolated from genomic clones by standard methods, and could provide a valuable resource for functional analysis of pluripotent cell biology. This could enable the investigation of cellular mechanisms during development, via the disruption of pluripotent cells at specific stages of embryogenesis. For example, the molecular basis of proamniotic cavitation could be analysed by ectopic expression of inhibitors of apoptosis during peri-implantation development, under the control of a *Psc1* promoter.

Icm1 or *Psc1* promoters could enable analysis of the molecular controls of pluripotent cell development. The generation of a series of deletion constructs of *Icm1* or *Psc1* promoter regions, combined with expression analysis during the ES to X cell transition *in vitro*, could define sequences responsible for gene expression and regulation. This could identify potential transcription factor binding sites in *Icm1* or *Psc1* promoters and enable the investigation of regulatory hierarchies that control pluripotent cell development *in vitro* and during early embryogenesis.

APPENDIX

An Unusual Arrangement of 13 Zinc Fingers in the Vertebrate Gene Z13

FOREWORD TO THE APPENDIX

The material in this thesis is documented as two sections, representing two distinct projects.

The appendix presents a study conducted by the candidate in the laboratory of Dr. J. R. E. Wells commencing in March 1992. The project was discontinued in December 1993 after the untimely death of Dr. Wells in November of that year. All the work presented in the appendix was conducted by the candidate. This material has been published:

Schulz, T. C., Hopwood, B., Rathjen, P. D. and Wells, J. R. E. (1995). An unusual arrangement of 13 zinc fingers in the vertebrate gene Z13. *Biochem. J.* **311**, 219-224.

Schulz, T.C., Hopwood, B., Rathjen, P.D., and Wells, J.R., (1995) An unusual arrangement of 13 zinc fingers in the vertebrate gene Z13.
Biochemical Journal, v. 311 (1), pp. 219-224.

NOTE:

This publication is included on pages 111-116 in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

<http://dx.doi.org/10.1042/bj3110219>

Appendix Figure 8 The Genomic Sequence of mZ13.

The 5740 bp sequence of the mZ13 genomic region is shown (see Appendix Figs. 1, 2b, 4). Exons are in upper case and bold, introns are numbered and in lower case. The POZ domain is boxed, individual zinc fingers are numbered and indicated by arrows. The full sequence of intron 1 and the location of the mZ13 5' untranslated region was not determined (asterisk). *Pst*I and *Bam*HI restriction sites are underlined (Appendix Fig. 4a).

tgctcattgt tgcggttga cccgtgctca tacctagtgt gtatctgctg tgggctatct 60
gatacttctg ctgggggact gctttgtcga gtaagctttg ggaatgactg ggctgccttg 120
gcttgcahta agtgtctgga gtcactctgc ctttttctggt agaaaagcct ctcccgtgtt 180
taattctctg tctcctctgt gagtcttacc tgggcttgct aggcagttgg cagtgagttg 240
ggctggcatc ctctcgggtgc tgacggaagg ctctggcttt ctgcag**CCAT** **GGATTTCCCC** 300
PsII
CAGCACAGCC AGCGTGTCTT GGAGCAGCTG AACCAGCAGC GGCAACTGGG CCTCCTCTGT 360
GACTGCACCT TTGTGGTGGG CGGCGTTGAC TTCAAGGCTC ACAAGGCGGT GCTGGCAGCT 420
TGCAGCGAGT ACTTCAAGAT GCTCTTTGTG GACCAGAAGG ACGTGGTGCA CCTAGACATC 480
AGTAATGCGG CAGgtaccat gctggggcct catcacgcat ggccttcta gggcaatgcc 540
agcaggccac gtcccacctg ggcctgagct ctgacctgt agataccaag gcgggagagg 600
gacagattga cagtatgctt agggcatgga gagaggagga ccagccttg tagcagggac 660
agatcgaggc aggcaggcct agaccttctt gggctctctc tgtggccaca tcagtccctgg 720
agctgacagc taggctgtgt aagatgagga gttagctgga cacagcttga gatgggctgg 780
attgaggaag gatggaccag tcttccatct gttttgagat agtcaagaca cggctactctc 840
ccattttata gatgtgcaaa cagacacaca gaggtcagat gactgagagt ttggtagagc 900
aggtgatgag gccttgaaga ccacattgtc agaagggagg gatgtccagt cccttctgtt 960
atctttttag aagctgtccc cagcccagca gacaggtggt gaggcttttg aagtgggtgt 1020
gacagccatc ctgtcctcga cag**GCCTGGG GCAGGTGCTG GAGTTCATGT ACACTGCTAA** 1080
GCTGAGCCTT AGCCCTGAGA ACGTGGACGA TGTCTGGCC GTGGCCAGCT TCCTCCAGAT 1140
GCAGGATATT **GTCACGGCCT GCCACAGCT CAAGTCCCTC GCTGAGCCGA GCAGCACCAC** 1200
TGGGGAGAGC GCAGACGCCT CAGCTGTGGA AGgtgaggtc actggccggc cactcctgaa 1260
gagtggcttc tggagctcct gcctgcctct gggctctctc actgggcctc aggttccccca 1320
tttaccttgc ccttgctctc attgtctggg gtgggggtgg gcggcgctgc taaggcgtgg 1380
ctagcacttc tggaaagggc cccatgctt accccttctt acctctgggc agcagaggca 1440
ggagccacac ctccggccaca cctgagctgt taactgcctt ttcagctctt ttggtatcct 1500
cttagggcgt cctcagaagc catagactcc tatcaggggt caggtctctc ctgtagcac 1560
tggctcacac ctagtctgca ctgcccgtgt aagatcatct cacttctgat gtggaggctg 1620
cagggctgtg aatcagtgtc ctagctgag acggctcctg aacacactag agaagtacct 1680
PsII
aggggtcctg aggggcctcg gagcctcctc attcctctgg aggatgaaaa acacttgatg 1740
aaaaacgaca agtttgctgt gttctgtctc cag**GAGGAGA CAAGAGAGCC AAAGACGAGA** 1800

AGGCCGCTGC CACCATGTTG AGCAGGCTGG GCCAAGCAAAG AGGCAGCTCC TCCACAGGCC 1860
CGGGCAGAGA ACTCAAGGAG GAGCGGGGTG GCCAGGCAGA AAGTGCATCC AGTGGTGAGT 1920
 tgtgattggg tctcatgggc tggctggcta cctggaagga ggtgacagac agttggccgg 1980
 tggccctgcc tgtcacctgt gcagcctggg cccagagaca ggaaaggcct gtgtctgaga 2040
 ggttggtggt atgcaactccc ataagtttct ggctcatgtc tagctcagac agtttgttct 2100
 gcagggcccg ttcgggcat ggaagagagc ctttggcct ctggacagag tccaatccag 2160
PstI
 cctcccgtcc ttcctcagGC **GCAGAGCAGA CGGAGAAGGC AGACGCTCCC CGGGAGCCTC** 2220
CACCTGTGGA GCTCAAGCCG GACCCACAA GCAGCATGGC GGCTGCAGAA GCCGAAGCCT 2280
PstI
TATCAGAGAG CTCAGAGCAA Ggtgcccctc cccagccag ggctcgtctc cctccgcacc 2340
 ctctggtggtc accactctct tctgtctcta gcctgtcctc tcagactcat ccttcacttc 2400
 atctctttgc ttctttcact cttgagctgt cttcattgcc agtccccatc cctgctcgt 2460
 tcctgttgtt agtgtggaag catcctgaca cagtccctgga gcttccctgct cagccctgcc 2520
 tgcgcacgac acatggtgtg ccatggtgcc cagaaacctc acaaggactt ggttggctga 2580
 aatgcccac acacatttgc ctgtggtggc atactctcta gccaaacta atgttcgatc 2640
 ctggagtcac tgtcacctaa gaagtggggg ggaacttggg ccagaccag aaagtgtctc 2700
 ggtctcagcc acgcccctcag tacctgctct tttctgtgca gAAATGGAAG **TGGAGCCAGC** 2760
CAGCAAAGGA GAAGACGGAC AAGAGGAAGA GGGTGCAGGG CCTGCCACAG TCAAAGAAGA 2820
GGGGATGCAT CTGGATAACG GGGAGCCTCC TGAGGAGAAT GAAGAGTCTG CCGGCACAGA 2880
TTCTGGGCAG GAGCTTGGCA TGGAGGGTCA GAACCTGCGC TCGGGCACCT ACGGGGATCG 2940
CACTGAGTCC AAGGCTTACG GCTCCATCAT CCACAAGTGC GAGgtgcggg gccgtggctg 3000
 gtgccagccc tgagcctcca gggctctctgc agcccctcac gcacccttc tegtgcgct 3060
 INTRON 6
PstI
 gcag**GA**CTGC **GGGAAGGAGT TCACGCACAC AGGGAACTTC AAACGGCACA TCCGCATCCA** 3120
PstI
CACGGGGGAG AAACCTTTCT CATGCAGGGA GTGCAGCAAG GCTTCTCGG ACCCTGCAGC 3180
PstI
GTGCAAGGCT CATGAGAAGA CACACAGgta cccacctgt ctctccaga ctgggcgct 3240
 INTRON 7
 gtttgggacc ctgcggcacc cttcctgtc acctagcacc ttctgacca cag**CCC**TTG 3300
AAACCCTATG GGTGTGAGGA GTGTGGCAAG AGCTACAGGC TCATCAGCCT GCTGAACCTG 3360
CATAAGAAGA GGCACCTCCG GAGGGCGCG TACCGCTGCG GGGACTGTGG CAAGCTCTTC 3420
ACCACTTACG GCAACCTCAA GCGCCACCAG CTGGTACACA GTGGCCAGAA ACCCTACCAG 3480
TGCGACTACT GTGGCCGCTC CTTCTCTGAC CCCACCTCTA AGATGCGCCA CCTGGAGACT 3540
CACGACCCG ACAAGGAACA CAAGTGTCTT CACTGCGACA AAAAGTTCAA CCAGgtgtgc 3600

accctcctga gccggcccca gatggctcct ggagctcccg cgatagcaca caacagccag 3660
 gacaggactc tcctaggatg gtccccggag ggctgtgggt gcttttctct ccagtgtctg 3720
 gacctgggct tctatcagga atagagccct gccctgtggg gaaaccagga aggccattag 3780
 tcccttgctc acggggttgg ggagaaacca gggccaagg gaacccttct gagcccctgt 3840
 ctgggtgcag **GTGGGGAACC TGAAGGCCCA CTTGAAGATC CACATTGCCG ATGGCCCTCT** 3900
TAAGTGCCGG **GAGTGCGGGA AGCAGTTCAC CACCTCAG**gt aggccagag caccctttcc 3960
 INTRON 8
 caaggcttgc ccctcgcccc cggccctca cagtgcagca cactggtctt gcctttcctc 4020
 caaag**GGAAC CTCAAGCGGC ACCTGCCGAT CCACAGTGGG GAGAAGCCGT ACGTATGCAC** 4080
CCACTGTCAG CGGCAGTTG CCGACCCAGG CGCGCTACAG CGGCACGTCC GGATCCACAC 4140
 INTRON 9
GGgtgggtga ggaggccagg tgtgggctgt ggacgggaa tggggagccg agcgagtgtc 4200
 agtgccatgt gcctctgccc cag**GTGAGAA GCCGTGCCAG TGTGTGATAT GTGGTAAGGC** 4260
CTTTACCCAA GCCAGCTCCC TCATCGCCCA TGTACGCCAA CACACGGGGG AGAAGCCCTA 4320
CGTCTGTGAA CGCTGTGGCA AGAGgttagcg acctgcccct accccagggg gagtgaggac 4380
 INTRON 10
 aaatctcagg ctgtgtccca gagtgggtggg gacctccaaa tcctctaca ctgaagctcc 4440
 catgggtcaa catgagcccc ttatttgtga cctttagat**AT TTGTCCAGTC CAGCCAGTTG** 4500
GCCAACCACA TCCGTCACCA TGACAAGATC CGACCCGACA AGTGCAGCGT GTGTAGCAAG 4560
GCCTTCGTGA ATGTGGGGGA CCTGTCCAAG CACATTATCA TCCACACCGg tgagcagccc 4620
 accatgtccc cgctgcgtg ggaccaggg gagggtggag gaagagagcc ccatgccaca 4680
 cttgctcagc agggagcctt gaggtacaca cgtggactcc acctgccttc cgggaggagg 4740
 INTRON 11
 gtgcaactcg gcttctgggt gctgccttag agaacttgtc agaaaagctg gggctgtctg 4800
 gtactgggag gcctgcaggc ggcaggagg taacttgct ccccaccag**G AGAGAAGCCT** 4860
TATTTGTGTG ACAAGTGTGG CCGTGGTTT AACCGGGTAG ACAACCTGCG TTCTCATGTA 4920
AAGACCGTGC ATCAGGGCAA GCGGGGCATC AAGATCCTGG AGCCAGAAGA GGGTGGTGAG 4980
GTCAGCGTGG TCACTGTGGA CGACATGGTC ACCCTGGCCA CTGAGGCTCT GGCAGCGACA 5040
GCGGTCACCTC AGCTCACAGg tgtgcaactgt gggcagggcg gcaggcagag ggcagctaca 5100
 INTRON 12
 gcagctggcc tcagtggcaa taatgggta gctccttctg gcttctgct cagacctgc 5160
 tgactctgcc cacag**TGGTA CCAGTGGGG CCGCAGTGAC AGCTGACGAG ACGGAAGTAC** 5220
TCAAAGCTGA GATCAGCAA GCTGTCAAGC AAGTGCAGGA AGAAGgtgag aggggcagtc 5280
 tggggcatga acgcccgtgg acccaaggtc cttcccatcc tgaccacagc cccacctcc 5340
 acag**ACCCCA ACACCCACAT CCTCTACGCT TGTGATTCTT GTGGGACAA GTTCTGGAT** 5400
 INTRON 13

13 ←

GCCAATAGCC TAGCCAGCA TGTTCGGATC CACACAGCCC AGGCACTGGT CATGTTCCAG 5460
BamHI

ACGGATGCGG ACTTCTACCA GCAGTATGGG CCAGGCAGCA CGTGGCCAGC CGGGCAGATG 5520

CTGCAGGCTG GAGAGCTCGT CTTCCGTCCT AGGGATGGGA CTGAGGGCCA ACCCACACTG 5580
PstI

GCAGAAAGTC CACCCACAGC TCCTGATTGC CTACCACCTG CCGAGTGAGC AGGTGGCCTC 5640

TGACTGTTTA TTTAAGGATG GATGGCACCC TAGGACTCAG AAGGTGCTCC TGCCCCATTC 5700

CCTAGAGAAT AAATGATTAT TTTCTAA_{tgt} tgccctgtggt 5740

REFERENCES

REFERENCES

- Adams, M. D., Kerlavage, A. R., Fields, C. and Venter, J. C.** (1993). 3,400 new expressed sequence tags identify diversity of transcripts in human brain. *Nat. Genet.* **4**, 256-267.
- 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**, 711-723.
- 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.
- Asling, B., Dushay, M. S. and Hultmark, D.** (1995). Identification of early genes in the *Drosophila* immune response by PCR-based differential display: the Attacin A gene and the evolution of attacin-like proteins. *Insect Biochem. Mol. Biol.* **25**, 511-518.
- Austbel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K.** (1994). *Current Protocols in Molecular Biology, Vol 2.*, Current Protocols, John Wiley & Sons, Inc. U.S.A.
- Barcena, A., Muench, M. O., Roncarolo, M. G. and Spits, H.** (1995). Tracing the expression of CD7 and other antigens during T- and myeloid-cell differentiation in the human fetal liver and thymus. *Leuk. Lymphoma.* **17**, 1-11.
- Becker, D. L., Leclerc-David, C. and Warner, A.** (1992). The relationship of gap junctions and compaction in the preimplantation mouse embryo. *Dev. Suppl.* 113-118.
- Beddington, R. S. P.** (1982). An autoradiographic 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. 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.** (1983). The origin of the foetal tissues during gastrulation in the rodent. In *Development in Mammals*, Vol 5, ed. Johnson, M. H. pp. 1-32. Elsevier, Amsterdam.
- Besmer, P., Manova, K., Duttlinger, R., Huang, E. J., Packer, A., Gyssler, C. and Bachvarova, R. F.** (1993). The kit-ligand (steel factor) and its receptor c-kit/W: pleiotropic roles in gametogenesis and melanogenesis. *Dev. Suppl.* 125-137.
- Bettess, M.** Gene expression in murine pluripotential stem cells. (1993). *Honours Thesis, University of Adelaide, Adelaide.*
- Biggers, J. D., Bell, J. E. and Benos, D. J.** (1988). Mammalian blastocyst: transport functions in a developing epithelium. *Am. J. Physiol.* **255**, C419-C432.
- Blok, L. J., Kumar, M. V. and Tindall, D. J.** (1995). Isolation of cDNAs that are differentially expressed between androgen-dependent and androgen-independent prostate carcinoma cells using differential display PCR. *Prostate* **26**, 213-224.
- Boulikas, T.** (1993). Nuclear localization signals (NLS). *Critical Rev. Euk. Gene Exp.* **3**, 193-227.
- 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., Kirsch, 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 hierarchy 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.

- Burglin, T. R.** (1994). A comprehensive classification of homeobox genes. In *Guidebook to the Homeobox Genes*, ed. Duboule, D. pp. 27-71. Sambrook and Tooze, Oxford University Press, Oxford.
- Callard, D., Lescure, B. and Mazzolini, L.** (1994). A method for the elimination of false positives generated by the mRNA differential display technique. *BioTechniques* **16**, 1096-1103.
- Chapman, G.** Isolation and characterisation of the murine homeobox gene *MmxB*. (1994). *Honours Thesis, University of Adelaide, Adelaide*.
- 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.
- Clark, J. M.** (1988). Novel nontemplated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. *Nucl. Acids Res.* **16**, 9677-9686.
- Copeland, N. G., Jenkins, N. A., Gilbert, D. J., Epping, J. T., Maltais, L. J., Miller, J. C., Deitrich, W. F., Weaver, A., Lincoln, S. E., Steen, R. G., Stein, L. D., Nadeau, J. H. and Lander, E. S.** (1993). A genetic linkage map of the mouse: current applications and future prospects. *Science* **262**, 57-66.
- Coucovanis, 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.
- Davis, W. Jr., De Sousa, P. A. and Schultz, R. M.** (1996). Transient expression of translation initiation factor eIF-4C during the 2-cell stage of the preimplantation mouse embryo: identification by mRNA differential display and the role of DNA replication in zygotic gene activation. *Dev. Biol.* **174**, 190-201.
- Diel, P., Walter, A., Fritzemeier, K. H., Hegele-Hartung, C. and Knauthe, R.** (1995). Identification of estrogen regulated genes in Fe33 rat hepatoma cells by differential display polymerase chain reaction and their hormonal regulation in rat liver and uterus. *J. Steroid Biochem. Mol. Biol.* **55**, 363-373.

- 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. Morph.* **87**, 27-45.
- Douglass, J., McKinzie, A. A. and Couceyro, P.** (1995). PCR differential display identifies a rat brain mRNA that is transcriptionally regulated by cocaine and amphetamine. *J. Neurosci.* **15**, 2471-2481.
- Duboule, D.** (1994). *Guidebook to the Homeobox Genes*. Sambrook and Tooze, Oxford University Press, Oxford.
- 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**, 3280-3288.
- Evans, M. J.** (1972). The isolation and properties of a clonal tissue culture strain of pluripotential mouse teratoma cells. *J. Embryol. Exp. Morphol.* **28**, 163-176.
- Evans, M. J. and Kaufman, M. H.** (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
- Feldman, B., Poueymirou, W., Papaioannou, 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.
- Finocchiaro, G., Ito, M., Ikeda, Y. and Tanaka, K.** (1988). Molecular cloning and nucleotide sequence of cDNAs encoding the alpha-subunit of human electron transfer flavoprotein. *J. Biol. Chem.* **263**, 15773-15780.
- Fleming, T. P., Javed, Q. and Hay, M.** (1992). Epithelial differentiation and intercellular junction formation in the mouse early embryo. *Dev. Suppl.* 105-112.
- Friedrich, G. and Soriano, P.** (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5**, 1513-1523.

- Gardner, R. L.** (1983). Origin and differentiation of extraembryonic tissues of the mouse. *Int. Rev. Exp. Pathol.* **24**, 63-133.
- Ginsberg, M., Snow, M. H. L. and McLaren, A.** (1990). Primordial germ cells in the mouse embryo during gastrulation. *Development* **110**, 521-528.
- Gossler, A., Joyner, A. L., Rossant, J. and Skarnes, W. C.** (1989). Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* **244**, 463-465.
- Guillemot, F., Nagy, A., Auerbach, A., Rossant, J. and Joyner, A. L.** (1994). Essential role of Mash-2 in extraembryonic development. *Nature* **371**, 333-336.
- Guillemot, F., Caspary, T., Tilghman, M., Copeland, N. G., Gilbery, D. J., Jenkins, N. A., Anderson, D. J., Joyner, A. L., Rossant, J. and Nagy, A.** (1995). Genomic imprinting of Mash-2, a mouse gene required for trophoblast development. *Nat. Genet.* **9**, 235-242.
- Guimarães, M. J., Bazan, J. F., Zlotnik, A., Wiles, M. V., Grimaldi, J. C., Lee, F. and McClanahan, T.** (1995). A new approach to the study of haematopoietic development in the yolk sac and embryoid bodies. *Development* **121**, 3335-3346.
- Harrison, S. M., Dunwoodie, S. L., Arkell, R. M., Lehrach, H. and Beddington, R. S. P.** (1995). Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo. *Development* **121**, 2479-2489.
- Haub, O. and Goldfarb, M.** (1991). Expression of the *fibroblast growth factor-5* gene in the mouse embryo. *Development* **112**, 397-406.
- Hébert, J. M., Boyle, M. and Martin, G. R.** (1991). mRNA localization studies suggest that murine FGF-5 plays a role in gastrulation. *Development* **112**, 407-415.
- Hedley, M. L., Amrein, H. and Maniatis T.** (1995). An amino acid sequence motif sufficient for subnuclear localization of an arginine / serine-rich splicing factor. *Proc. Natl. Acad. Sci. (USA)*. **92**, 11524-11528.

- Haegel, H., Larue, L., Ohsugi, M., Fedorov, L., Herrenknecht, K. and Kemler, R.** (1995). Lack of β -catenin affects mouse development at gastrulation. *Development* **121**, 3529-3537.
- Hendrey, J., Lin, D. and Dziadek, M.** (1995). Developmental analysis of the Hba^(th-J) mouse mutation: effects on mouse peri-implantation development and identification of two candidate genes. *Dev. Biol.* **172**, 253-63.
- Hermesz, E., Mackem, S. and Mahon, K. A.** (1996). *Rpx*: a novel anterior-restricted homeobox gene progressively activated in the prechordal plate, anterior neural plate and Rathke's pouch of the mouse embryo. *Development* **122**, 41-52.
- Herrmann, B. G., Labeit, S., Poutska, A., King, T. R. and Lehrach, H.** (1990). Cloning of the *T* gene required in mesoderm formation in the mouse. *Nature* **343**, 617-622.
- Hogan, B., Constantini, F. and Lacy, E.** (1986). *Manipulating the Mouse Embryo: A Laboratory Manual*. (1st. Edn.), Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.
- Hogan, B., Beddington, R. S. P., Constantini, F. and Lacy, E.** (1994). *Manipulating the Mouse Embryo: A Laboratory Manual*. (2nd. Edn.), Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.
- 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**, 5623-5629.
- Izpisua-Belmonte, J. C., Dolle, P., Renucci, A., Zappavigna, V., Falkenstein, H. and Duboule, D.** (1990). Primary structure and embryonic expression pattern of the mouse Hox-4.3 homeobox gene. *Development* **110**, 733-745.
- Jackson, B. W., Grund, C., Winter, S., Franke, W. W. and Illmensee, K.** (1981). Formation of cytoskeletal elements during mouse embryogenesis. II. Epithelial differentiation and intermediate-sized filaments in early postimplantation embryos. *Differentiation* **20**, 203-216.
- 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.

Jones, M. D. and Foulkes, N. S. (1989). Reverse transcription of mRNA by *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* **17**, 8387-8388.

Joyner, A. L. (1991). Gene targeting and gene trap screens using embryonic stem cells: new approaches to mammalian development. *Bioessays* **13**, 649-656.

Joyner, A. L., Auerbach, A. and Skarnes, W. C. (1992). The gene trap approach in embryonic stem cells: the potential for genetic screens in mice. *Ciba. Found. Symp.* **165**, 277-288.

Kaufman, M. H. (1992). *The Atlas of Mouse Development*. Academic Press, Harcourt Brace Jovanovich, London.

Khavari, P. A., Peterson, C. L., Tamkun, J. W., Mendel, D. B. and Crabtree, G. R. (1993). BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nature* **366**, 170-174.

Kidder, G. M. (1992). The genetic program for preimplantation development. *Dev. Genet.* **13**, 319-325.

Kishimoto, T., Taga, T. and Akira, S. (1994). Cytokine signal transduction. *Cell* **76**, 253-262.

Krieg, P. A. and Melton, D. A. (1987). *In vitro* RNA synthesis with SP6 RNA polymerase. *Methods Enzymol.* **155**, 397-415.

Lake, J. Differentiation of pluripotential murine embryonic stem cells. (1996). *Ph.D. Thesis, University of Adelaide, Adelaide.*

Latham, K. E., Garrels, J. I., Chang, C. and Solter, D. (1991). Quantitative analysis of protein synthesis in mouse embryos. 1. Extensive reprogramming at the one- and two-cell stage. *Development* **112**, 921-932.

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.

Leivo, I., Vaheri, A., Timpl, R. and Wartiovaara, J. (1980). Appearance and distribution of collagens and laminin in the early mouse embryo. *Dev. Biol.* **76**, 100-114.

Liang, P., Averboukh, L., Keyomarsi, K., Sager, R. and Pardee, A. B. (1992). Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelial cells. *Cancer Res.* **52**, 6966-6968.

Liang, P. and Pardee, A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967-971.

Loeb, D. D., Padgett, R. W., Hardies, S. C., Shehee, W. R., Comer, M. B., Edgell, M. H. and Hutchison, C. D. (1986). The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol. Cell. Biol.* **6**, 168-182.

Loveland, B., Wang, C. R., Yonekawa, H., Hermel, E. and Lindahl, K. F. (1990). Maternally transmitted histocompatibility antigen of mice: a hydrophobic peptide of a mitochondrially encoded protein. *Cell* **60**, 971-980.

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.

MacGregor, G. R., Zambrowicz, B. P. and Soriano, P. (1995). Tissue non-specific alkaline phosphatase is expressed in both embryonic and extraembryonic lineages during mouse embryogenesis but is not required for migration of primordial germ cells. *Development* **121**, 1487-1496.

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)* **78**, 7634-7638.

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). The formation of embryoid bodies *in vitro* by homogeneous embryonal carcinoma cell cultures derived from isolated single cells. In *Teratomas and Differentiation*, ed. Sherman, M. I. and Solter, D. pp. 169-187. Academic Press, New York.

Martin, G. R. and Evans, M. J. (1975c). Multiple differentiation of clonal teratocarcinoma stem cells following embryoid body formation *in vitro*. *Cell* **6**, 467-474.

Matsui, Y., Zsebo, K. and Hogan, B. L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841-847.

Mattox, W., Ryner, L. and Baker, B. S. (1992). Autoregulation and multifunctionality among *trans*-acting factors that regulate alternative pre-mRNA processing. *J. Biol. Chem.* **267**, 19023-19026

Moore, M. J., Qurey, C. C. and Sharp, P. A. (1993). *The RNA World*. Cold Spring Harbor Laboratory Press, New York.

Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M., Chambers, I. and Smith, A. (1994). Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression. *Proc. Natl. Acad. Sci. (USA)* **91**, 4303-4307.

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.

Niswander, L. and Martin, G. R. (1992). Fgf-4 expression during gastrulation, myogenesis, limb and tooth development in the mouse. *Development* **114**, 755-768.

Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M. and Hamada, H. (1990). A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**, 461-472.

Packer, A. I., Manova, K. and Bachvarova, R. F. (1993). A discrete LINE-1 transcript in mouse blastocysts. *Dev. Biol.* **157**, 281-283.

Papaiouannou, 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.

Pedersen, R. A. (1988). Potency, lineage, and allocation in preimplantation mouse embryos. In *Experimental Approaches to Mammalian Development*, ed. Rossant, J. and Pedersen, R. A. pp. 3-33. Cambridge University Press, Cambridge.

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.

Petters, R. M. and Markert, C. L. (1980). Production and reproductive performance of hexaparental and octaparental mice. *J. Hered.* **71**, 70-74.

Piquet-Pellorce, C., Grey, L., Mereau, A. and Heath, J. K. (1994). Are LIF and related cytokines functionally equivalent? *Exptl. Cell Res.* **213**, 340-347.

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.

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 differentiation inhibiting activity and the control of stem cell populations. *Genes Dev.* **4**, 2308-2318.

Robertson, E. J. (1987). *Embryo-derived stem cell lines. In Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, ed. Robertson, E. J. pp. 71-112. IRL press, Oxford, Washington DC.

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**, 815-824.

Rosen, B. and Beddington, R. S. P. (1993). Whole-mount in situ hybridization in the mouse embryo: gene expression in three dimensions. *Trends Genet.* **9**, 162-167.

Rosenthal, M. D., Wishnow, R. M. and Sato, G. H. (1970). *In vitro* growth and differentiation of clonal populations of multipotent mouse cells derived from a transplantable testicular teratocarcinoma. *J. Natl. Cancer. Inst.* **44**, 1001-1014.

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. (1993). Immortal Germ Cells? *Current Biology* **3**, 47-49.

Rossant, J. and Pedersen, R. A. (1988). *Experimental Approaches to Mammalian Embryonic Development*. Cambridge University Press, Cambridge.

Rothstein, J. L., Johnson, D., DeLoia, J. A., Skowronski, J., Solter, D. and Knowles, B. (1992). Gene expression during preimplantation mouse development. *Genes. Dev.* **6**, 1190-1201.

Rudnicki, M. A. and McBurney, M. W. (1987). Cell culture and induction of differentiation of embryonal carcinoma cell lines. In *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, ed. Robertson, E. J. pp. 19-49. IRL Press, Oxford, Washington DC.

Ryner L. C. and Baker B. S. (1991). Regulation of *doublesex* pre-mRNA processing occurs by 3'-splice site activation. *Genes Dev.* **5**, 2071-2085.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. (2nd edn.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schichman, S. A., Severynse, D. M., Edgell, M. H. and Hutchison, C. A. (1992). Strand-specific LINE-1 transcription in mouse F9 cells originates from the youngest phylogenetic subgroup of LINE-1 elements. *J. Mol. Biol.* **224**, 559-574.

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**, 2185-2195.

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**, 435-439.

Schreiber, V., Molinete, M., Boeuf, H., de Murcia, G. and Menissier-de Murcia, J. (1992). The human poly (ADP-ribose) polymerase nuclear localization signal is a bipartite element functionally separate from DNA binding and catalytic activity. *EMBO. J.* **11**, 3263-3269.

Scott, M. P., Tamkun, J. W. and Hartzell III., G. W. (1989). The structure and function of the homeodomain *Biochim. Biophys. Acta.* **989**, 25-48.

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.

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. *Semin. Cell Biol.* **3**, 385-399.

Smith, A. G. and Hooper, M. L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev. Biol.* **121**, 1-9.

Snow, M. H. L. (1977). Gastrulation in the mouse: growth and regionalization of the epiblast. *J. Embryol. Exp. Morphol.* **42**, 293-303.

Snow, M. H. L. and Tam, P. P. L. (1979). Is compensatory growth a complicating factor in mouse teratology? *Nature* **279**, 554-557.

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.

Spritz, R. A., Strunk, K., Surowy, C. S., Hoch, S. O., Barton, D. E. and Franchke, U. (1987). The human U1-70K snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing and RNA binding. *Nuc. Acids Res.* **16**, 10373-10391.

Spyropoulos, D. D. and Capecchi, M. R. (1994). Targeted disruption of the even-skipped gene, *evx-1*, causes early postimplantation lethality of the mouse conceptus. *Genes Dev.* **8**, 1949-1961.

- Stone , B. and Wharton, W.** (1994). Targeted RNA fingerprinting: the cloning of differentially-expressed cDNA fragments enriched for members of the zinc finger gene family. *Nucl. Acids Res.* **22**, 2612-1618.
- 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**.
- Subramaniam, M., Harris, S. A., Oursler, M. J., Rasmussen, K., Riggs, B. L. and Spelsberg, T. C.** (1995). Identification of a novel TGF-beta-regulated gene encoding a putative zinc finger protein in human osteoblasts. *Nucleic Acids Res.* **23**, 4907-4912.
- Suzuki, N., Rohdewold, H., Neuman, T., Gruss, P. and Schöler, H. R.** (1990). Oct-6: a POU transcription factor expressed in embryonal stem cells and in the developing brain. *EMBO J.* **9**, 3723-3732.
- Sweeney, C., Murphy, M., Kubelka, M., Ravnik, S. E., Hawkins, C. F., Wolgemuth, D. J. and Carrington, M.** (1996). A distinct cyclin A is expressed in germ cells in the mouse. *Development* **122**, 53-64.
- Tam, P. P. L. and Beddington, R. S. P.** (1992). Establishment and organization of germ layers in the gastrulating mouse embryo. *Ciba. Found. Symp.* **165**, 27-41.
- 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 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.

- Temeles, G. L., Ram, P. T., Rothstein, J. L. and Schultz, R. M.** (1994). Expression patterns of novel genes during mouse preimplantation embryogenesis. *Mol. Reprod. Dev.* **37**, 121-129.
- Thomas, P. Q.** (1995). Homeobox gene expression in murine embryonic stem cells. *Ph.D. Thesis, University of Adelaide, Adelaide.*
- Thomas, P. Q., Johnson, B. V., Rathjen, J. and Rathjen, P. D.** (1995). Sequence, genomic organization, and expression of the novel homeobox gene *Hesx1*. *J. Biol. Chem.* **270**, 3869-3875.
- Thomas, P. Q. and Rathjen, P. D.** (1992). HES-1, a novel homeobox gene expressed by murine embryonic stem cells, identifies a new class of homeobox genes. *Nucleic Acids Res.* **20**, 5840.
- Williams, T. and Tjian, R.** (1991). Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. *Genes Dev.* **5**, 670-682.
- Wilson, R. et al.** (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**, 32-38.
- 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**, 171-179.
- Zahler, A. M., Lane, W. S., Stolk, J. A. and Roth, M. B.** (1992). SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6**, 837-847.
- Zimmermann, J. W. and Schultz, R. M.** (1994). Analysis of gene expression in the preimplantation mouse embryo: use of mRNA differential display. *Proc. Natl. Acad. Sci. (USA)* **91**, 5456-5460.
- Zhang, G., Taneja, K. L., Singer, R. H. and Green, M. R.** (1994). Localization of pre-mRNA splicing in mammalian nuclei. *Nature* **372**, 809-812.