Declaration

I declare that the work presented in this thesis is my own, unless otherwise stated, and has not be submitted for any other degree or professional qualification.

William Hamilton
2011
Acknowledgements

I would like to thank my supervisor Tilo for his support and guidance, for allowing me the freedom to peruse lines of investigations that I found most interesting, and for always maintaining a positive attitude towards this project, even when things weren’t going particularly well. Also to Mike for exposing me to novel ideas and for financial support towards the end of my PhD. I also owe a big thanks to my committee members, Thierry for always putting things into perspective, and Val for her critical and strategic input. I am indebted to several people in the ISCR who were an invaluable source of both technical and conceptual advice. Kei Kaji for always taking the time to answer what I’m sure were generally tedious questions. Nick Mullin for help with all things protein. Ian Chambers for providing reagents, advice on targeting, and showing me that sometimes going back-to-bluescript is the best strategy. To the various smokers that have passed through the ISCR that have had to listen to my complaining when all they wanted was a 10 minute break, as well as to the other, healthier friends I’ve made in the ISCR, I thank you all for making the past four years a largely enjoyable experience. To my Mum, Heather and Andy, for your support and encouragement. And finally to Noel, for all the red-eye flights to and from Edinburgh, for support and encouragement, and for pretending there was such a thing as ‘GCNF4’.

This work was supported by a University of Edinburgh SBS studentship and by funding from Parkinson's UK.
Abstract

The ERK-MAPK pathway is a dynamic signaling module, conserved across Eukarya, and capable of processing a myriad of environmental and cellular signals. It has been implicated in controlling important cell fate decisions in many cell types and species. In mES cells, growth factor activation of the ERK-MAPK pathway is involved in the earliest stages of lineage segregation, however very little is currently known about the mechanism by which this is accomplished.

Taking a loss-of-function gene targeting approach I have reexamined the relative contribution of ERK2 activity to FGF-ERK signaling. Although ERK2 depletion results in an attenuation of the combined ERK1/2 activity, this is compensated for by the hyperactivation of the remaining ERK1 isozyme. Normal ERK1/2 function can be restored to ERK2 deficient cells by transgenic expression of either ERK1 or ERK2, indicating a degree of functional redundancy between both isoforms.

When subjected to the appropriate cues, lineage commitment proceeded normally in ERK2 deficient cells, however increased self-renewal was observed under standard culture conditions. Several attempts were made to further probe ERK1/2 function by siRNA depletion, and dominant negative inhibition of ERK1 in Erk2 knockout cells, however both approaches failed to provide further insight. Furthermore, taking a candidate approach, the role of Srf, a canonical target of ERK1/2 signaling, was examined. Initial experiments indicated a role for SRK in neural differentiation, however due to issues of culture adaptation and instability in several cell lines it was not possible to conclude this line of research within the time frame of this thesis.

IP-MS/MS analysis identified several proteins known to interact with ERK2 and indicated an involvement in nuclear pore function through TPR as well as transcriptional and translational regulation through RSK proteins. Moreover, this study identified DUSP6 and DUSP9 as the primary induced dual specificity phosphatases that regulate ERK2 activity in mES cells.
To further probe the functional significance of the ERK:p90RSK interaction I examined a mES cell line genetically depleted for PDK1, a crucial regulator of p90RSK function. This cell line exhibits no detectable p90RSK activity, however in contrast to studies in other cell lines, p90RSK activity is dispensable for mitogen-induced $cFos$ expression in mES cells. Subsequent experiments demonstrated a requirement for PDK1 activity in either the specification or maintenance of mES cell derived neurons. Further analysis indicated that p90RSK may be involved in a negative feedback loop regulating ERK1/2 activity, and if so may represent a point whereby ERK1/2 activity can be manipulated. To examine this I determined the effect pharmacological inhibition of p90RSK has on ERK1/2 activity and self-renewal using a novel p90Rsk inhibitor, BI-D1870. Although treatment with BI-D1870 correlated with enhanced ERK1/2 phosphorylation, the off-target effects this molecule exhibits made it impossible to draw any firm conclusions from these experiments.

Overall this study has demonstrated a degree of redundancy between ERK1/2 isoymes in mES cells. It has highlighted the complex nature of ERK1/2 regulation as well as the robustness of this pathway to perturbations in ERK dose. Furthermore, it has underscored some of the common pitfalls encountered when studying differentiation phenotypes in mES cells. Although this study failed to highlight anything more than a coincidental relationship between ERK1/2 activity and self-renewal capacity of mES cells, it has helped to highlight some important behavioral characteristics of the FGF-MAPK pathway in mES cells and provide a platform for further study.
## Table of Contents

**Declaration** ii

**Acknowledgements** iii

**Abstract** iv

**List of Figures** xiii

**List of Tables** xvi

**Chapter 1 - Introduction** 1

1.1 Signal transduction 1
   1.1.1 Phosphorylation 1
   1.1.2 Kinases 2
      1.1.2.1 The kinase domain 3
      1.1.2.2 Docking motifs 5
   1.1.3 Protein Phosphatases 7
   1.1.4 Signal transduction pathways 8
   1.1.5 The MAPK Pathway 9
      1.1.5.1 MKKKs 10
      1.1.5.2 MKKs 12
         1.1.5.2.1 Mek1/2 12
         1.1.5.2.2 Mek3/6 13
         1.1.5.2.3 Mek4/7 13
         1.1.5.2.4 Mek5 14
      1.1.5.3 MAPKs 15
         1.1.5.3.1 Erk1/2 16
         1.1.5.3.2 p38 \(\alpha/\beta/\gamma/\delta\) 17
         1.1.5.3.3 Jnk1/2/3 18
         1.1.5.3.4 Erk5 19
         1.1.5.3.5 Atypical MAPKs 20
      1.1.5.4 MAPK-activated protein kinases 20
      1.1.5.5 The use of small molecule inhibitors to probe MAPK function 24

1.2 The Raf-Mek-Erk pathway 26
   1.2.1 Receptor activation 26
   1.2.2 Scaffolds 27
1.2.3 Interactions within the MAPK core 29
1.2.4 The Immediate Early Gene paradigm 31
    1.2.4.1 IEG expression through Serum Response Factor 32
1.3 Erk activity during early embryogenesis 35
    1.3.1 Early murine development 35
    1.3.2 The role of Fgf-Mek-Erk signaling in the developing embryo 39
    1.3.3 Functional redundancy of Erk1 and Erk2 43
1.4 mES cells 47
    1.4.1 Extrinsic factors that regulate pluripotency 48
        1.4.1.1 Leukemia inhibitory factor 48
        1.4.1.2 Bone morphogenic protein 4 50
    1.4.2 Transcriptional Regulation Of mES Cell Identity 50
        1.4.2.1 Oct4 51
        1.4.2.2 Sox2 52
        1.4.2.3 Nanog 53
    1.4.3 Epigenetic regulation of mES cell identity 54
    1.4.4 miRNA 56
    1.4.5 mES cell heterogeneity 57
1.5 The role of Fgf4 signaling in mES cell biology 58
    1.5.1 PI3K signaling 59
    1.5.2 ERK MAPK signaling 60
    1.5.3 3i culture and the ‘Ground State' hypothesis 62
1.6 Technical notes 64
    1.6.1 Genome engineering 63
    1.6.2 Mass spectrometry based approaches to 66
determine kinase interactions
        1.6.2.1 IMAC and phosphopeptide enrichment 68
        1.6.2.2 IP-LC-MS/MS 69
1.7 Aims of this thesis 73

Chapter 2: Materials and Methods 75

2.1: Materials 75

    2.1.1: Kits and Reagents 75
        2.1.2.1 Kits 75
2.1.2.2 Reagents
2.1.2: Solutions
2.1.3: Plasmids
2.1.4: Primers/Synthetic oligonucleotides
2.1.5: Antibodies
2.1.6: Cell Lines
2.1.7: Culture Medium
  2.1.7.1: Bacterial culture medium
  2.1.7.2: Mammalian cell culture medium
2.1.8: In-silico resources

2.2: Methods

2.2.1: DNA Techniques
  2.2.1.1: DNA isolation from bacteria
    2.2.1.1.1: Plasmid isolation from DH5α
    2.2.1.1.2: BAC isolation from DH10β
  2.2.1.2: DNA isolation from mammalian cells
    2.2.1.2.1: DNA isolation from 96 well plates
    2.2.1.2.2: DNA isolation from 6 well plates
  2.2.1.3: DNA isolation from mouse ear clippings
  2.2.1.4: DNA cloning
    2.2.1.4.1: Restriction enzyme digestion
    2.2.1.4.2: Dephosphorylation of DNA fragment ends
    2.2.1.4.3: Blunting DNA fragment overhangs
    2.2.1.4.4: DNA fragment ligation
    2.2.1.4.5: DNA electrophoresis
    2.2.1.4.6: DNA extraction from agarose gels
    2.2.1.4.7: DNA extraction from solution
    2.2.1.4.8: DNA precipitation
    2.2.1.4.9: DNA quantification
    2.2.1.4.10: Bacterial transformation
      2.2.1.4.10.1 Chemically competent DH5α
      2.2.1.4.10.2: Bacterial transformation
  2.2.1.5: Polymerase chain reaction
    2.2.1.5.1: Generic PCR
2.2.1.5.1.1: Site-directed DNA mutagenesis 98
2.2.1.5.1.2: PCR fusion by overlap extension 98
2.2.1.5.2: Blunt cloning of PCR products 99
2.2.1.5.3: Quantitative genomic PCR 99
2.2.1.5.4: Quantitative RT-PCR 99
2.2.1.6: DNA sequencing 99
2.2.1.7: Southern blotting 100
2.2.1.7.1: Sample preparation 100
2.2.1.7.2: DNA electrophoresis and denaturation 100
2.2.1.7.3: Sample transfer 100
2.2.1.7.4: Probe preparation and hybridization 101
2.2.1.7.5: Membrane washing and exposure 101
2.2.2: RNA Methods 102
2.2.2.1: Total RNA extraction 102
2.2.2.2: RNA quantification 103
2.2.2.3: First strand cDNA synthesis using random hexamers 103
2.2.3: Protein Methods 103
2.2.3.1: Total protein isolation from mammalian cells 103
2.2.3.2: Protein quantification 104
2.2.3.3: Protein separation by SDS-PAGE 104
2.2.3.4: Protein detection by coomassie staining 104
2.2.3.5: Protein detection by silver staining 105
2.2.3.6: Western blotting 105
2.2.3.6.1: Sample preparation 105
2.2.3.6.2: Protein transfer to nitrocellulose membrane 105
2.2.3.6.3: Membrane blocking, probing and washing 106
2.2.3.6.4: Signal detection 107
2.2.3.7: Immunopercipitation of tagged proteins 107
2.2.3.7.1: Sample preparation 107
2.2.3.7.2: Protein capture 108
2.2.3.8: Sample preparation for Mass Spectrometry 108
2.2.3.8.1: Gel-free digestion 108
2.2.3.8.2: In-gel digestion 109
2.2.3.9: Size exclusion chromatography 110
2.2.3.10: TCA precipitation of proteins 110
2.2.4: Tissue Culture 110
2.2.4.1: Routing culture of mES cells
   2.2.4.1.1: Culture of 129/OLA derived mES cells
   2.2.4.1.2: Culture of C57/BL6-J derived mES cells
2.2.4.2: Transfection of mES cells
   2.2.4.2.1: Transfection by liposomes
   2.2.4.2.2: Transfection by electroporation
2.2.4.3: Picking mES cell colonies
2.2.4.4: Freezing mES cells
2.2.4.5: Thawing mES cells
2.2.4.6: Karyotyping
2.2.4.7: hLIF05 expression in COS7 cells
2.2.4.8: Acute stimulation of mES cells
2.2.4.9: Differentiation of mES cells
   2.2.4.9.1: Neural monolayer differentiation
   2.2.4.9.2: Non-neural surface ectoderm differentiation
   2.2.4.9.3: Multilineage differentiation in hanging drops
   2.2.4.9.4: Mesoderm monolayer differentiation
   2.2.4.9.5: Clonal assay
2.2.5: Histological techniques
   2.2.5.1: Alkaline phosphatase staining
   2.2.5.2: Immunocytochemistry
   2.2.5.3 Flow Cytometry
2.2.6: Mouse and embryo techniques
   2.2.6.1: Animal housing
   2.2.6.2: mES cell derivation

Chapter 3: Analysis of ERK2 regulation in mES cells

3.1 Introduction
   3.1.2 Hypothesis
   3.1.3 Aims of this chapter
3.2 Results
   3.2.1: Transgenic rescue of ERK2 deficiency
   3.2.2: Genetic depletion of Erk2 in 129/Ola mES cells
   3.2.3: Analysis of Erk2 depletion on FGF/ERK signal transduction
3.3 Discussion

3.3.1 Analysis of embryo-derived Erk2 mutant cells 143
3.3.2 Generation of Erk2 knockout lines 145
3.3.3 Robustness of the MAPK pathway in mES cells 146
3.3.4 Redundancy of ERK isozymes in mES cells 149

3.4 Future directions 149

Chapter 4: Functional characterization of ERK2 in mES cells 152

4.1 Introduction 152

4.1.1 Aims of this chapter 152

4.2 Results 153

4.2.1 ERK2 is dispensable for multi-lineage differentiation 153
4.2.2 Attenuation of ERK signaling by overexpressing ERK2 mutant proteins 163
4.2.3 The role of Srf in mES lineage commitment 169
4.2.4 Derivation and analysis of Srf+/fl and Srf/fl/fl lines 177

4.3 Discussion 181

4.3.1 ERK2 deficiency enhances self-renewal standard culture conditions, but is dispensable for lineage commitment 181
4.3.2 ERK activity cannot be attenuated by overexpressing dominant negative ERK2 mutants 183
4.3.2 The role of Srf in mES cell biology 184

4.4 Future directions 187

Chapter 5: Identification and analysis of ERK2 interacting partners 189

5.1 Introduction 189

5.1.1 Aims of this chapter 190

5.2 Results 191

5.2.1 Optimizing the IP protocol 191
5.2.2 Identification of ERK2 interacting partners in mES cells 199
5.2.3 The role of p90Rsk proteins in MAPK signaling in mES cells 206
5.2.4 The effect of the p90RSK inhibitor BI-D1870 on MAPK activity and mES cell differentiation 212

5.3 Discussion 219
5.3.1 The use of IP-MS/MS as a tool to study Erk2 interactions in mES cells 219
5.3.2 Molecules that interact with Erk2 in mES cells 222
   5.3.2.1 ERK1/2 222
   5.3.2.2 TPR 224
   5.3.2.3 The HSP90-CDC37-FKBP5 complex 225
   5.3.2.4 DUSP6 and DUSP9 227
   5.3.2.5 p90RSK proteins 228
5.3.3 BI-D1870 as a tool to modulate ERK function in mES cells 231

Chapter 6: The pERKs and ERKs of MAPK signaling in mES cells 234

   6.1 Technical limitations of this study 235
   6.2 Naïve and Primed states of pluripotency 237

Appendix 239
   A1: Steady-state behavior ERK behavior in mES cells 239
   A2: Analysis of the ERK2 genetrap cell line RRZ307 242
   A3: Comparison of ERK1 activity by either FGF2 or LIF 243

References
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Serine Phosphorylation</td>
<td>2</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>The structure of p38</td>
<td>5</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>The overengineered signal relay</td>
<td>9</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Orientation of the MAPK pathway</td>
<td>10</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Sequence homology and domain architecture of MKKK family members as well as the MAPK they activate</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>MAPK sequence homology domain architecture</td>
<td>16</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Mechanism of activation of p90Rsk proteins</td>
<td>24</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Processive Vs distributive Erk activation</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.9</td>
<td>Interactions within the MAPK core</td>
<td>31</td>
</tr>
<tr>
<td>Figure 1.10</td>
<td>Early murine embryogenesis</td>
<td>37</td>
</tr>
<tr>
<td>Figure 1.11</td>
<td>Sequence alignment of the N-terminal of Erk1 and Erk2</td>
<td>44</td>
</tr>
<tr>
<td>Figure 1.12</td>
<td>A typical LC-MS/MS workflow</td>
<td>67</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Cartoon depicting the Southern blot transfer</td>
<td>101</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Cartoon depicting the western blot transfer</td>
<td>105</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Transgenic rescue of Erk2 deficiency</td>
<td>124</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>The block in neural induction cannot be transgenically rescued</td>
<td>125</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Erk2 deficient cells can differentiate in embryoid bodies</td>
<td>126</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Cartoon illustrating the WT Erk2 locus, and both modified loci</td>
<td>128</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Q-PCR analysis for Erk2 exon copy number</td>
<td>131</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Southern blot analysis of parental and Erk2 +/- clones</td>
<td>131</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Genotyping of parental and Erk2 +/- clones</td>
<td>132</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Erk1/2 expression in mutated lines</td>
<td>132</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>Embryo-derived and ES cell derived Erk2 null lines have distinct morphologies</td>
<td>133</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>Semi-quantitative western blot analysis of ppErk levels in WT and Erk2 deficient cell lines</td>
<td>136</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>Western blot analysis of both Erk and p90Rsk phosphorylation following acute Fgf2 stimulation</td>
<td>136</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page No.</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Figure 3.12</td>
<td>IEG induction is attenuated in the absence of Erk2</td>
<td>137</td>
</tr>
<tr>
<td>Figure 3.13</td>
<td>Analysis of Erk induced Dusp expression</td>
<td>140</td>
</tr>
<tr>
<td>Figure 3.14</td>
<td>D-type cyclin expression is comparable in the presence or absence of Erk2</td>
<td>140</td>
</tr>
<tr>
<td>Figure 3.15</td>
<td>Western blot analysis of Erk1 and Erk2 overexpressing cells</td>
<td>142</td>
</tr>
<tr>
<td>Figure 3.16</td>
<td>IEG expression is rescued by overexpression of either Erk1 or Erk2</td>
<td>142</td>
</tr>
<tr>
<td>Figure 3.17</td>
<td>Regulatory motifs in the Ras-Raf-Mek-Erk pathway: NFA</td>
<td>148</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Erk2 deficient cells efficiently neuralize under monolayer conditions</td>
<td>154</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Gene expression analysis of cultures following 5 and 10 days of neural induction</td>
<td>155</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Multilineage differentiation of mES cells is Erk2 independent</td>
<td>156</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Erk2 deficiency results in increased expression of pluripotency markers</td>
<td>160</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Nanog-Gfp targeting</td>
<td>160</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Increased Nanog promoter activity in the absence of Erk2</td>
<td>161</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Erk2 deficiency does not affect the kinetics of early differentiation</td>
<td>161</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Erk2 depletion enhances self-renewal in the presence of Lif</td>
<td>162</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>Clonal analysis of a panel of Erk2 mutant constructs expressed in Erk2&lt;sup&gt;-/-&lt;/sup&gt; cells</td>
<td>166</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>Clonal analysis of a panel of Erk2 mutant constructs expressed episomally in E14/&lt;sup&gt;T&lt;/sup&gt; cells</td>
<td>167</td>
</tr>
<tr>
<td>Figure 4.11</td>
<td>Western blot analysis of Erk2 mutant cells stably expressing 3xFLAG-tagged Erk2-K52R</td>
<td>168</td>
</tr>
<tr>
<td>Figure 4.12</td>
<td>Expression of Erk2-K52R positively effects the Fgf/Erk pathway output</td>
<td>168</td>
</tr>
<tr>
<td>Figure 4.13</td>
<td>Srf&lt;sup&gt;+/fl&lt;/sup&gt; mES cells fail to differentiate under mesoderm monolayer conditions</td>
<td>171</td>
</tr>
<tr>
<td>Figure 4.14</td>
<td>Srf&lt;sup&gt;+/fl&lt;/sup&gt; mES cells fail to make neural tissue under monolayer conditions</td>
<td>172</td>
</tr>
<tr>
<td>Figure 4.15</td>
<td>Transgenic rescue of Srf deficiency</td>
<td>175</td>
</tr>
<tr>
<td>Figure 4.16</td>
<td>Neural commitment is not rescued by transgenic add-back</td>
<td>176</td>
</tr>
<tr>
<td>Figure 4.17</td>
<td>Derivation of Srf&lt;sup&gt;+/fl&lt;/sup&gt; and Srf&lt;sup&gt;fl/fl&lt;/sup&gt; lines</td>
<td>179</td>
</tr>
<tr>
<td>Figure 4.18</td>
<td>Drug regulated Cre mediated excision</td>
<td>179</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page No.</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Figure 4.19</td>
<td>Cre expression is toxic to both heterozygous and homozygous floxed Srf lines</td>
<td>180</td>
</tr>
<tr>
<td>Figure 4.20</td>
<td>Cartoon depicting three possible feedback scenarios leading to enhanced Erk1/Rsk/cFos activities</td>
<td>184</td>
</tr>
<tr>
<td>Figure 4.21</td>
<td>RMCE targeting approach to study Srf function</td>
<td>188</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Efficiency of Erk2 immunoprecipitation</td>
<td>195</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>A comparison of 4 elution methods</td>
<td>195</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Comparison of ‘In-Gel’ versus ‘Gel-free’ methods of immunoprecipitate analysis.</td>
<td>196</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Peptides identified in Erk2 IPs by biological process</td>
<td>197</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>Gel-filtration analysis of Erk1/2</td>
<td>197</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Analysis of GFP-Trap® beads</td>
<td>180</td>
</tr>
<tr>
<td>Figure 5.7</td>
<td>Analysis of eluted fractions from control and FLAG-Erk2 large scale immunoprecipitates for LC-MS/MS analysis</td>
<td>202</td>
</tr>
<tr>
<td>Figure 5.8</td>
<td>Erk2 and p90Rsk stably interact following Fgf stimulation</td>
<td>204</td>
</tr>
<tr>
<td>Figure 5.9</td>
<td>PDK acts to restrain Erk activity in mES cells</td>
<td>209</td>
</tr>
<tr>
<td>Figure 5.10</td>
<td>PDK1 is dispensable for mesoderm and endoderm specification <em>in vitro</em></td>
<td>201</td>
</tr>
<tr>
<td>Figure 5.11</td>
<td>Neural differentiation is PDK1 dependent</td>
<td>211</td>
</tr>
<tr>
<td>Figure 5.12</td>
<td>Neural differentiation is p90Rsk independent</td>
<td>211</td>
</tr>
<tr>
<td>Figure 5.13</td>
<td>BI-D1870 treatments activates Erk1/2 in mES cells</td>
<td>214</td>
</tr>
<tr>
<td>Figure 5.14</td>
<td>BI-D8170 treatment only slightly increases Erk phosphorylation following Fgf stimulation</td>
<td>214</td>
</tr>
<tr>
<td>Figure 5.15</td>
<td>BI-D1870 treatment facilitates increased induction of IEGs following Fgf2 stimulation</td>
<td>215</td>
</tr>
<tr>
<td>Figure 5.16</td>
<td>BI-D1870 attenuates the self-renewal of mES cells</td>
<td>215</td>
</tr>
<tr>
<td>Figure 5.17</td>
<td>BI-D1870 enhances Sox1-GFP promoter activity</td>
<td>218</td>
</tr>
<tr>
<td>Figure 5.18</td>
<td>BI-D1870 treatment has little effect on the efficiency of mES cell neuronal differentiation</td>
<td>218</td>
</tr>
<tr>
<td>Figure 5.19</td>
<td>Fgf4 induced Dusp expression in mES cells</td>
<td>228</td>
</tr>
<tr>
<td>Figure A1</td>
<td>Map of pIPC37 used for general purpose transgene expression</td>
<td>240</td>
</tr>
<tr>
<td>Figure A2</td>
<td>Behavior of steady state Erk signaling in mES cells</td>
<td>241</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page No.</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Figure A3</td>
<td>Selection and validation of Erk2 genetrap cell line</td>
<td>244</td>
</tr>
<tr>
<td>Figure A4</td>
<td>Southern blot analysis of pGT2lx integration in RRZ307</td>
<td>245</td>
</tr>
<tr>
<td>Figure A5</td>
<td>pETV</td>
<td>246</td>
</tr>
<tr>
<td>Figure A6</td>
<td>AGS 703</td>
<td>246</td>
</tr>
<tr>
<td>Figure A7</td>
<td>Rosa26.1</td>
<td>247</td>
</tr>
<tr>
<td>Figure A8</td>
<td>pETVH</td>
<td>248</td>
</tr>
<tr>
<td>Figure A9</td>
<td>pL451</td>
<td>249</td>
</tr>
<tr>
<td>Figure A10</td>
<td>pETVN</td>
<td>250</td>
</tr>
<tr>
<td>Figure A11</td>
<td>pGT2lx</td>
<td>251</td>
</tr>
<tr>
<td>Figure A12</td>
<td>Lif and Fgf activate Erk1 for different durations</td>
<td>251</td>
</tr>
<tr>
<td>Figure A13</td>
<td>pCAG-FKBP-L106P Srf-IH</td>
<td>253</td>
</tr>
<tr>
<td>Figure A14</td>
<td>Low-level expression of Srf allows differentiation at clonal destiny</td>
<td>253</td>
</tr>
<tr>
<td>Figure A15</td>
<td>Plasmid used to express FKBP-Cre-ErT2 fusion protein</td>
<td>254</td>
</tr>
<tr>
<td>Figure A16</td>
<td>Analysis of the distribution of +ve and -ve background</td>
<td>257</td>
</tr>
</tbody>
</table>

**List of Tables**

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Distribution of kinase subfamilies by species</td>
<td>3</td>
</tr>
<tr>
<td>Table 1.2</td>
<td>Fgf4 stimulated Srf IEG targets in mES cells</td>
<td>34</td>
</tr>
<tr>
<td>Table 1.3</td>
<td>Tissue specific expression profile of MAPK isoforms based on EST data</td>
<td>46</td>
</tr>
<tr>
<td>Table 1.4</td>
<td>Differential Erk compensatory activation</td>
<td>47</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>List of Erk2 mutants generated for this study</td>
<td>163</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Proteins identified by LC-MS/MS as putative Erk2 interactors judged by peak intensity values</td>
<td>202</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>Proteins identified by LC-MS/MS as putative Erk2 interactors judged by peptide abundance</td>
<td>203</td>
</tr>
<tr>
<td>Table 5.3</td>
<td>Post-translationally modified peptides identified in this study</td>
<td>204</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
<td>Page No.</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Table 5.4</td>
<td>A comparison of Erk2 interactors identified in this study, with proteins identified by Von Kregsheim et al., (2009)</td>
<td>222</td>
</tr>
<tr>
<td>Table 5.5</td>
<td>Antibodies used to detect phospho-Erk1/2 in mES cell studies</td>
<td>224</td>
</tr>
<tr>
<td>Table 5.6</td>
<td>Abundance of Tpr, Nup50, Nup155, Nup153 and Nup214 transcripts during murine development based on EST data</td>
<td>225</td>
</tr>
<tr>
<td>Table A1</td>
<td>A list of common FLAG-agarose binding proteins in mES cells</td>
<td>254</td>
</tr>
</tbody>
</table>
Chapter 1 - Introduction

1.1 Signal transduction

Signal transduction can be loosely defined as the biological mechanism by which information is transmitted from external sources through a cell or from one subcellular location to another. In actuality it consists of multiple sets of highly complex and overlapping protein interaction networks whose function are to interpret specific signals by discerning their characteristics i.e. point of initiation, magnitude, duration, etc, and convert this information into a functional response, usually a change in cellular homeostasis or a cell fate decision such as differentiation, proliferation, apoptosis, or migration. Information is relayed throughout such networks by the activities of enzymes that catalyze the post-translational modifications of their substrates, thus altering their activity or protein stability. Post-translational modification usually involves the addition of a specific chemical moiety onto an amino acid side chain, which can act as a docking site to facilitate the interactions with other proteins or to alter the local charge of the protein thus altering protein conformation and function, or both. These modifications are usually reversible and therefore for every enzyme that catalyses a reaction in one direction, there is a corresponding enzyme to catalyse the reaction in the opposite. It is the controlled balance between these two actions that allows such signaling networks to be controlled and from which highly complex behaviors can arise (reviewed in Kholodenko et al., 2010, and in Tyson et al., 2003).

1.1.1 Phosphorylation

Phosphorylation is one of the most common, and best understood forms of post-translational modification. It involves the covalent but reversible transfer of a phosphate group (PO$_4^{3-}$) from the terminal phosphate of adenosine-5’- tri-phosphate (ATP), to the amino acid side chain of serine (Ser), threonine (Thr) and tyrosine (Tyr), and sometimes histidine (His), residues in eukaryotes (reviewed in Cossone et al., 1988) (Figure 1.1). The presence of phosphoproteins in cells was first inferred at
the turn of the twentieth century by bulk analysis of protein cleavage products (Levene et al., 1906). By the 1950s phosphate metabolism was well-documented, and was also known to be increased in various tumors (Williams-Ashman et al., 1952), however it wasn’t till several years later that the enzymatic phosphorylation of proteins, in an ATP-dependent manner, was described (Burnet et al., 1954) and subsequently the functional significance of such modifications elucidated (Krebs and Fischer., 1962), which won Edmond Fischer and Edwin Krebs the Nobel prize for Physiology or Medicine in 1992.

Figure 1.1: Serine phosphorylation. Cartoon depicting the enzymatic phosphorylation of the hydroxyl group of a serine side-chain in a peptide (R1,R2). A kinase (blue) binds a molecule of ATP and catalyses the transfer of γ-phosphate of ATP to the hydroxyl oxygen of the serine, resulting the production of ADP. The reaction is catalyzed in the reverse direction by a phosphatase (green) resulting in the release of a molecule of inorganic phosphate.

1.1.2 Kinases

Kinases belong a large family of proteins that account for approximately 2% of the human genome, corresponding to 518 distinct genes (Manning et al., 2002). Most protein kinases belong to a single superfamily containing a eukaryotic protein kinase (ePK) catalytic domain, with a second family composed of atypical protein kinases (aPK) (Table 1.1). Despite the fact that all kinases catalyze essentially the same phosphoryl transfer reaction, they display remarkable diversity in their
structures, substrate specificity, and in the number of pathways in which they participate. However a common and defining feature of all typical kinases is the kinase domain (Cheek et al., 2002).

<table>
<thead>
<tr>
<th>Families</th>
<th>Subfamilies</th>
<th>Yeast</th>
<th>Worm</th>
<th>Fly</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGC</td>
<td>14</td>
<td>21</td>
<td>17</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>CAMK</td>
<td>17</td>
<td>33</td>
<td>21</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>CK1</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>95</td>
<td>10</td>
</tr>
<tr>
<td>CMCG</td>
<td>8</td>
<td>24</td>
<td>21</td>
<td>49</td>
<td>33</td>
</tr>
<tr>
<td>Other</td>
<td>37</td>
<td>39</td>
<td>38</td>
<td>67</td>
<td>45</td>
</tr>
<tr>
<td>STE</td>
<td>3</td>
<td>13</td>
<td>14</td>
<td>25</td>
<td>18</td>
</tr>
<tr>
<td>Tyrosine Kinase</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>90</td>
<td>32</td>
</tr>
<tr>
<td>Tyrosine Kinase-like</td>
<td>7</td>
<td>13</td>
<td>0</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>RGC</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>Atypical-PDHK</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Atypical-Alpha</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Atypical-RIO</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Atypical-Atypical-Atypical-A6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Atypical-Other</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Atypical-ABC1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Atypical-BRD</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Atypical-PIKK</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>134</td>
<td>201</td>
<td>130</td>
<td>454</td>
<td>240</td>
</tr>
</tbody>
</table>

Table 1.1: Distribution of kinase subfamilies by species. Adapted from Manning et al. (2002), http://kinase.com/

1.1.2.1 The kinase domain

Most protein kinase domains are between 250-300 amino acids (a.a.) long and are generally located towards the carboxyl terminal of the protein (Hanks et al., 1988). Classical kinase domains consists of an N-terminal lobe of mainly β-sheets and a C-terminal lobe of mainly α-helices (Figure 1.2). They typically share a conserved arrangement of secondary structure elements that are arranged into 12 subdomains that fold into a bi-lobed catalytic core structure with ATP binding in a deep cleft located between the lobes (Johnson et al., 2001). ATP binds in the cleft so that the adenosine moiety is buried in a hydrophobic pocket with the phosphate backbone orientated outwards. The protein substrate binds along the cleft and a set of conserved residues within the kinase catalytic domain catalyse the transfer of the terminal γ-phosphate of ATP to the hydroxyl oxygen of the Ser/Thr or Tyr residue of the substrate (Ubersax et al., 2007, and references therein).
Many kinases possess a small regulatory polypeptide region known as the activation lip/loop, that lies between N- and C-terminal lobes of the kinase domain, that is involved in regulating the kinase activity of the protein (Adams, 2003). Kinases can be phosphorylated in this region, either in an autocatalytic manner, or by the action of another kinase. Such modifications in the activation loop cause local conformation changes and exposes the substrate binding pocket thus acting as phosphorylation-sensitive switch controlling both substrate binding and indirectly kinase activity (Hubbard et al., 1997, Mohammadi et al., 1997).

Various mutational studies have identified a highly conserved residue in the β5 strand, distal to the active site and adjacent to a hinge region that connects the N- and C-terminal domains of the kinase. This residue has been termed the ‘gatekeeper’ residue as it regulates ATP binding in the kinase domain (Emerik et al., 2006 and references therein). In most enzymes the side-chain of this residue is bulky and mutation to residues with smaller side-chains, such as alanine (Ala) or Gly, allows enzymes to bind and utilise ATP analogues such as N6-cyclopentyl ATP (Elben et al. 2003). Work from the lab of Natalie Ahn has also demonstrated a previously unappreciated role for the gatekeeper residue, inhibition of enzyme auto-activation (Emerik et al., 2006). Mutation of Q103 to either alanine or glycine resulted in an appreciable increase in the basal activity of Extracellular signal-regulated kinase-2 (ERK2), which was due to intramolecular phosphorylation of the ERK2 TEY activation motif. However, this increase in auto-phosphorylation is only seen at basal activity, and the activities of WT and Q103A/G are similar following stimulation.
1.1.2.2 Docking motifs

It has been estimated that approximately 30% of all cellular protein are phosphorylated on at least one residue (Cohen, 2000), however most protein kinases exhibit a high degree of substrate specificity ranging from only a few, to several hundred substrates (Ptacek et al., 2005). This substrate specificity is in part mediated by certain surface motifs on the kinase, both within and outside of the kinase domain, that facilitate specific interactions with substrates that may be present at low abundance, as well as kinase-specific recognition sequences on the substrate. Other mechanisms such as protein scaffolding also contribute to substrate specificity and will be discussed later.

The first point of substrate of discrimination arises from the architecture of the catalytic cleft itself. Protein kinases can be broadly subcategorised into those that exhibit a preference for Ser/Thr or Tyr residues, however this is not absolute (Zhu et al., 2000). As the side-chain of Tyr residues contain an aromatic ring and are
therefore bulkier than Ser/Thr residues, Tyr-directed kinases have a catalytic cleft sufficiently deep to accommodate the Tyr side-chain, but too deep for the side-chains of Ser/Thr residues to span the distance between the γ-phosphate of ATP and the kinase backbone (Anderson et al., 2001). Moreover, residues directly adjacent to the phosphorylation site also contribute to substrate specificity by regulating electrostatic interactions between the phosphorylation site and residues within the catalytic cleft of the kinase.

As a kinase swims in a sea of potential substrates that are composed of many phosphorylatable residues (Echols et al., 2002) it is obvious that discriminatory mechanism other than phosphorylation site motif must be at play. Interactions between docking sites on substrates and interaction domains on kinases have been shown to play an important part in conferring specificity to the interaction (Holland et al., 1999, Tanuoe et al., 2000). Docking motifs have been extensively studied in the function of mitogen-activated protein kinase (MAPK) family members (Tanuoe et al., 2000). Many MAPK family members contain a cluster of negatively charged residues (Aspartic or Glutamic acid, (Asp) and (Glu), respectively) located near the C-terminal of the protein that interacts with positively charged residues (usually Arginine (Arg) or Lys) known as a D-Domain, on not only their substrates (reviewed in Sharrocks et al., 2000), but activating kinases and inactivating phosphatases. This motif is known as the common docking (CD) domain, it is evolutionarily conserved and is thought to act by increasing the localized concentration of interaction proteins thus increasing the efficiency and specificity of enzymatic reactions. However, mutation of the CD domain from ERK2 to that of p38 did not crosswire their substrate specificity and therefore though the CD domain can increase substrate specificity it does not define it (Tanuoe et al., 2000).

Docking motifs can be both constitutive, i.e. encoded in the primary sequence of the protein, or conditional, i.e. regulated by post-translational modifications (Cohen, 2000). Conditional docking motifs represent an important switch-like regulatory mechanism controlling pathway activity, or may prime certain substrates for interaction thus increasing the efficiency of the reaction. Conditional
docking motifs regulate the activity of, and substrate specificity of polo-like kinase (PLK) 1 (Elia et al., 2003), as well the interaction of Src homology (SH) 2 domain proteins with phosphorylated Tyr (Russell et al., 1992).

Recent work from Mody and colleagues has helped to shed light on how protein kinases evolved such mechanisms of substrate specificity and also indicated a novel approach to determining regions essential for protein:protein interactions (Mody et al., 2009). In the context of the yeast MAPKs FUS3 and HOG1, orthologues to mammalian ERK and p38 proteins, respectively, Mody and colleagues demonstrated that MAPKs are modular in nature, that specific interactions are distributed across many solvent-exposed residues, and that by switching modular domains between MAPK proteins it was possible to cross wire the FUS3 and HOG1 responses.

1.1.3 Protein Phosphatases

As stated before, protein phosphatases are molecules that catalyse the removal of a phosphate group from it substrate resulting in the production of a molecule of inorganic phosphate (Figure 1.1). They act in direct opposition to kinases and have similar preferences for phosphorylated residues in that different classes of phosphatases will dephosphorylate either Ser/Thr or Tyr phosphorylated peptides (reviewed in Tonks, 2006, and in Shi, 2009). However, a large and diverse family of dual specificity phosphatases (DUSPs) are capable of dephosphorylating Ser/Thr and Tyr residues (Brondello et al., 1997). Phosphatases had historically be thought of as passive housekeeping enzymes however as our understanding of the mechanisms of signal transduction grew, roles for phosphatases in positively regulating signaling events (Saxton et al., 1997), as well as negatively regulating signal duration became apparent (Hornberg et al., 2005).

In humans 107 putative Tyr phosphatases have been identified, with only approximately 30 Ser/Thr phosphatases (Alonso et al., 2004). As Ser is by far the most commonly phosphorylated residue in eukaryotes it seems counterintuitive that there should be fewer Ser than Tyr phosphatases, however as Ser phosphatases
generally exist as part of large holoenzymes whose behaviour and specificity is
determined by the combination of various catalytic and regulatory subunits (Shi,
2009). In contrast Tyr phosphatases tend not to function in multimeric enzymatic
complexes and the number of kinases to phosphatases is more closely matched
(Tonks, 2006).

## 1.1.4 Signal transduction pathways

Over the past 30 years great improvements have been made in our
understanding of how signaling networks are constructed, the interconnectivity of
various components, and the processes governed by such pathways and how such
‘design principals’ could have evolved (Soyer et al., 2006). A combination of
classical biochemical techniques and targeted genetic manipulation has led to the
prevalent notion that signal transduction pathways are shaped in cascades whereby
activation of an upstream molecule, usually a cell surface receptor, results in the
sequential activation of downstream components of the pathway culminating on a
terminal affector(s) (Figure 1.3-A). The traditional view of a signaling pathway is
similar to that of a Rube Goldberg machine, where an initial stimulus is relayed
through a complex set of intermediaries to perform a very simple task in a very
complex fashion (Figure 1.3-B). As our understanding of the biophysical nature of
signal transduction pathways grows the need for such network topologies become
apparent:

- As there are far fewer signaling pathways than the biological processes they
govern, having multi-step, multi-component pathways allows for multiple points of
crosstalk with other pathways that can both coordinate and modify responses to a
certain stimulus (Li et al., 2008).
- Redundancy between certain signaling components can confer robustness to a
pathway (Tomaiuolo et al., 2008)
- Having a protracted route from input to output allows the system to discriminate
between transient ‘noise’ and bona fide signals, which is increasingly relevant to
multicellular organisms where cells exist in a graded sea of growth factors and
have to discriminate between the degree of pathway activation, yet can still respond appropriately to slight fluctuations in growth factor concentrations (reviewed in Kholodenko et al. 2010).

- It allows signal amplification at each step of the pathway (Birtwistle et al., 2011 and references therein).
- It allows unique behaviours to arise from varying the concentrations of components of the pathway (O'Shaughnessy et al., 2011).

As a comprehensive overview of the many signal transduction pathways known to exist is beyond the focus of this study, I will now focus predominantly on the MAPK pathway, however, many of the design principals it exhibits are shared by other signal transduction pathways.

Figure 1.3: The overengineered signal relay. A) Schematic representation of the path a signal (receptor activation) takes before it is translated into a change in gene expression (cFos). B) Cartoon depicting that path an input (tensile force from a flying kite) takes before it is translated into an effect (a sharpened pencil). Cartoon by Rube Goldberg.

1.1.5 The MAPK Pathway

MAPK pathways are among the most ancient eukaryotic signal transduction pathways and are widely used throughout evolution in many physiological processes and is conserved as a three tiered module from yeast to man (Wideman et al., 1999). The MAPK module includes three kinases that establish a sequential activation pathway comprising a MAPK kinase kinase (MKKK), MAPK kinase (MKK), and MAPK (Figure 1.4). In humans 21 genes have been identified
that code for MKKKs, 7 that code for MKK’s and 14 that code for MAPKs (10 typical and 4 atypical) that can be activated by discrete and overlapping signals (Uhlik et al., 2004). Other than the three-tiered structure of the MAPK module, defining features are the conserved dual phosphorylation of Ser and Thr residues as well as a distributive mode of activation whereby each site is phosphorylated by a single enzyme substrate interaction. Therefore, dual phosphorylation and hence full activation is dependent on two separate protein interaction events, however the generality of this mechanism, as well as the system behavior it produces is still a much debated point (Schilling et al., 2010, Aoki et al., 2011).

1.1.5.1 MKKKs

To date all MKKKs identified possess a similar kinase domain, although they display a high degree of variation in the surrounding sequence, with simple proteins such as c-MOS composed of just a kinase domain, whilst MEKK1 is divided into at least 3 functionally distinct domains (Figure 1.5). The abundance of MKKK isoforms as well as their position at the top of the MAPK tier indicates that they represent the primary point of signal integration into the MAPK core (reviewed in Cuevas et al., 2007). Although it is only a guide, and there are many exceptions to it, MKKKs can be divided in groups based on which Mek/MAPK isoforms the activate.
c-Jun NH2-terminal kinase (Jnk) and p38 MAPKs are by far the most promiscuous in terms of their activating MKKKs, with many members of the mixed lineage kinase (Mlk), thousand and one amino acid kinase (TAO kinase) families, as well as MEKK4 and MEKK5, being implicated in their activation (Cuevas et al., 2007 and references therein). RAF family MKKKs display the most restricted preference for downstream substrates almost exclusively activating ERK1/2 (reviewed in Wellbrock et al., 2004), however there is evidence to suggest that ERK5 activity is RAF-1 dependent in NIH 3T3 cells, though this is mediated by interactions independent of RAF-1 kinase activity (English et al., 1999).

The diversity of MKKK isoforms, as well as their overlapping MAPK targets allows for a high degree of diversity in the input-output relationship of the core MAPK module. However, understanding how stimulus-dependent activation of MKKK isoforms results in a specific response still remains a major challenge. Genetic studies have helped to shed some light on the major functions governed by individual MKKK isoforms, and to address the issue of redundancy between closely related proteins (Camarero et al., 2006). However genetic-background dependent phenotypes, as seen in mutational studies in RAF family members (Wojnowski et al., 1998, Pritchard et al., 1996), further confound our understanding of the specific roles of MKKK family members.

Figure 1.5: Sequence homology and domain structure of MKKK family members as well as the MAPK they activate. Sequence homology dendrogram was made using ClustalW and represents mouse a.a. sequence alignment. Domain structure was determined.
by searching the Scansite database. MAPK activation profiles are from (Cuevas et al., 2007). The pink square under Erk5 represents indirect activation by Raf-1.

1.1.5.2 MKKs

Like MKKKs, MKKs are also dual-specificity kinases and preferentially phosphorylate their substrates on Ser/Thr residues. Of the 7 MKK proteins currently identified, all display a preference for the TxY motif located in the activation lip of their target MAPK, a mechanism initially discovered for ERK2 (Robbins et al., 1993). Unlike MKKKs, MKKs are highly restricted in their substrate specificity. Specificity can be achieved through interactions between N-terminal region of the MKK and the docking groove present on the surface of the MAPK, distant from the catalytic active site (Chang et al., 2002), which may be a mechanism of recognition common to all MKK:MAPK interaction (Tanuoe et al., 2000). Moreover a novel MKK interaction domain termed the DVD has been shown to be essential for MKK:MEKK binding, and possibly confers an additional level of specificity to MKK activation (Takekawa et al., 2005).

1.1.5.2.1 MEK1/2

MEK1 and MEK2 have been generally viewed as functionally redundant molecules owing to their high degree of sequence similarity, 80% homology and their ability to activate both ERK1 and ERK2 in response to mitogenic stimuli (Shaul et al., 2007). This view has been further substantiated by the observation that only combined Mek1/2 ablation was able to block oncogenic Rat Sarcoma (RAS) induced hyperplasia (Scholl et al., 2009), with intermediate responsiveness observed upon ablation of either gene individually. This being said there is evidence to suggest that MEK1 alone can serve as a target of ERK mediated negative feedback regulation due to phosphorylation of Thr292 on MEK1 that interferes with MEK1/2 heterodimerisation and activity (Catalanotti et al., 2009), however, differences in their regulatory mechanisms may not necessarily indicate a lack of functional redundancy. Other than the RAF interaction domain (Catling et al., 1995) and the
ERK docking site (Fukuda et al., 1997), MEK1/2 possess a unique pocket adjacent to the ATP binding core that has been shown to regulate the enzymatic activity when bound to a certain class of small molecule inhibitor (Delaney et al., 2002, Ohren et al., 2004).

1.1.5.2.2 MEK3/6

MEK3 and MEK6 are the second most closely related MKKs with 76% a.a. sequence homology that preferentially activate p38 kinases on their TGY motif. While MEK6 activates all p38 isoforms, MEK3 is somewhat more selective, as it preferentially phosphorylates α/β/γ isoforms (Cargnello et al., 2011). MEK3/6 are activated in response to a host of stimuli including environmental stresses and inflammatory cytokines, including oxidative stress, UV irradiation, hypoxia, ischemia, interleukin-1 (IL-1), and tumor necrosis factor alpha (TNF-α) (reviewed in Cuadrado et al., 2011). Single knockout mice for either Mek3 or Mek6 are viable, however compound mutants die mid-gestation (~E11.5) with impaired placental growth and embryonic vasculature, similar to p38α mutant mice (Bramcho et al., 2003 and references therein). Interestingly, MEK6 is subject to a unique mechanism of negative feedback regulation (Ambrosino et al., 2003). Many of the documented negative feedback mechanisms governing MAPK pathway activity are known to occur at the level of post-translational modification (reviewed in Birtwistle et al., 2011), however p38α negatively regulates the Mek6 mRNA stability through its 3’ UTR, specifically and without affecting Mek3 transcript half-life. This isoform specific feedback mechanism may act to dampen the more promiscuous nature of MEK6 activity.

1.1.5.2.3 MEK4/7

Similar to MEK3/6, MEK4/7 are activated by predominantly stress stimuli such as ionizing radiation, heat shock, DNA damage, as well as by inflammatory cytokines and growth factors (Rmana et al. 2007). MEK4/7 show a main preference in activating members of the JNK family of MAPKs, although activation of p38
family members by MEK4 has been demonstrated (Bramcho et al., 2003). MEK4/7 phosphorylate their substrate MAPKs on their conserved TPY motif although each MEK family member displays a different a.a. preference with MEK4 preferentially phosphorylating Tyr and MEK7 for Thr, and thus it has been proposed that both isoforms act synergistically (Lawler et al., 1998). However, recent experiments in cells deficient for either, or both isoforms, showing different responses to cadmium chloride treatment may challenge this notion. Signal specificity is brought about again in part by specific docking motifs in the N-terminal region of MEK4, however, this motif serves not only as a docking site for JNK proteins, but also for MEKK1 to form a complexed MAPK signaling molecule (Xia et al., 1998).

1.1.5.2.4 MEK5

MEK5 one of the recently discovered MKKs (English et al., 1995) and has the most divergent primary structure of all MKKs with homology ranging from 39% with MEK6 and 25% with MEK7, and is expressed as alternatively spliced transcripts that code for a cytoskeletal α- and ubiquitous β-isoforms. MEK5 binds and exclusively phosphorylates EK5 on its TEY motif in response to mitogenic and stress stimuli. Targeted disruption of the Mek5 locus results in embryonic lethality at ~E10.5 owing to abnormal cardiac development as well as reduced proliferation and enhanced apoptosis in the developing heart (Wang et al., 2005). The Mek5 mutant largely phenocopies the Erk5 mutant mouse reinforcing the notion that MEK5 and ERK5 represent a separate MAPK module (Regan et al., 2002).

MEK5 possess a Phox/Bem1p (PB1) domain in its C-terminal extension that mediates interactions with its activating MKKKs, MEKK1/2 (Nakamura et al., 2003). Furthermore, the C-terminal PB1 domain of MEK5 also binds ERK5 forming a ternary signaling complex of MKKK, MKK and MAPK (Nakamura et al., 2007). In quiescent cells MEK5 is preferentially complexed with MEKK2. Once stimulated with sorbitol, MEK5 activates ERK5 and disassociates from the complex, freeing the MEKK2 PB1 domain to interact with MEK7 and activate JNK family members, thus
showing how docking sites on both MKKKs and MKKs can coordinate parallel signaling pathways (Nakamura et al., 2007).

1.1.5.3 MAPKs

MAPKs are considered the terminal effectors of the MAPK pathway and are broadly divided by their mechanism of activation into typical and atypical. MAPKs can be further subdivided into seven separate groups: ERK1/2, p38α/β/δ/γ, JNK1/2/3, ERK5, ERK3/4, ERK7/8, NLK (Figure 1.6) (Cargnello et al., 2011). MAPKs were first discovered in yeast, Fus3p in the pheromone response pathway, and later in mammalian tissue, ERK1/2 as microtubule-associated kinases. Sequence homology between these molecules indicated that they were members of a newly identified protein kinase family (Courchesne et al., 1989, Boulton et al., 1990). Extensive research in the years since has identified several common features of MAPK family members. They are all activated by dual phosphorylation in the regulatory loop of their kinase domain: typical MAPKs all possess a TxY motif, whilst atypical activation motifs differ substantially. They also share a high degree of sequence and structural similarities within their kinase domain (reviewed in Pearson et al., 2001). Furthermore, MAPKs display preference for specific amino acid residues surrounding the phosphorylation site. Specifically, a Pro at the +1 position is favoured as it can be accommodated by the hydrophobic pocket of the kinase domain, however Gly can also be accommodated possibly because it lacks a side chain (reviewed in Lewis et al., 1998). Though this preferential motif provides some predicative power when searching for putative MAPK targets, a primary focus of the MAPK field is in developing novel reliable methods to determine context dependent MAPK substrates (reviewed in Powell et al., 2005).
1.1.5.3.1 ERK1/2

The prototypical mammalian MAPKs, ERK1 and ERK2 share over 80% sequence homology and are activated by dual phosphorylation on their TEY motif. They are ubiquitously expressed, though their relative abundance can vary from tissue to tissue, and are activated in response to a myriad of both extracellular and intracellular stimuli (Yoon et al., 2006). Erk1 mutant mice are viable but exhibit defects in thymocyte maturation (Pages et al., 1999), whereas Erk2 knockout mice are embryonic lethal owing to defects in trophoblast function (Saba-El-Leil et al., 2003), mesoderm induction (Yao et al., 2003) or placental development (Hatano et al., 2003). The discrepancies between knockout phenotypes is often sited as proof of lack of redundancy between isoforms, both arguments for and against will be discussed later.

ERK1/2 activation is predominantly initiated by membrane receptors such as receptor tyrosine kinases (RTKs) (Lemond et al., 2010), G protein–coupled receptors (GPCRs) (Naor et al., 2000), and ion channels (reviewed in Rane, 1999). Stimulation of RTKs results in the recruitment of various adapter proteins such as fibroblast growth factor receptor substrate (FRS) 2 and growth factor receptor-bound protein 2 (GRB2). GRB2 appears to be constitutively bound to the guanine

Figure 1.6: MAPK sequence homology and domain architecture. Sequence homology dendrogram was made using ClustalW and represents mouse a.a. sequence alignment. Domain structure and activation motif were derived from the literature.
nucleotide exchange factor (GEF) son of sevenless (SOS). Once recruited the membrane SOS catalyses the exchange of guanosine diphosphate (GDP) for guanosine triphosphate (GTP) in RAS family proteins facilitating interactions with RAF-MKKKs (aRAF, bRAF or RAF-1). RAF recruitment to the plasma membrane coincides with its dephosphorylated on S259 (RAF-1), and de-inhibition due to release of 14-3-3 protein binding (Rodriguez-Viciana et al., 2006). RAF activation results in MEK1/2 phosphorylation and subsequently ERK1/2 activation.

Once activated ERK1/2 translocates various cellular compartments where it phosphorylates its substrates and coordinates gross alterations in cellular homeostasis. ERK1/2 target both nuclear and cytoplasmic proteins and regulate processes such as transcription by activating ternary complex factors (TCF) ELK-1, SAP-1 and NEN (Cruzalegui et al., 1999), nuclear transport through TPR and NUP proteins (Vomastek et al., 2008), and chromatin dynamics through PARP-1 (Cohen-Armon et al., 2007), and BAF155 (Kosako et al., 2009). Furthermore, ERK1/2 activity can regulate protein synthesis by targeting eukaryotic translation initiation factors (eIFs) (reviewed in Graff, 2002) as well as ribosome biogenesis (Mariappan et al., 2010 and references therein). The activation kinetics, regulation, and targets of ERK1/2 will be discussed in more detail later.

1.1.5.3.2 p38 α/β/γ/δ

p38 was initially identified as a Tyr phosphorylated protein in lipopolysaccharide (LPS)-stimulated macrophages (reviewed in Raman et al., 2007). There are currently 4 different p38 genes (p38 α/β/γ/δ), and differential splicing has resulted in the identification of multiple p38 isoforms. p38 proteins are both cytoplasmic and nuclear in quiescent cells, and accumulate in the nucleus upon activation. They are activated by MEK3/6 (TGY phosphorylation), in response to various cellular stresses such as UV radiation, osmotic shock, pro-inflammatory cytokines and in some instances growth factors. However non-canonical, MEK-independent mechanisms of auto-activation have been described (Mittelstadt et al., 2005). Unlike other MAPK subfamilies, p38 family members exhibit a high degree
of sequence variation (~40% homology), thus implying a degree of functional diversification. Indeed, in PC12 cells, \( p38 \gamma \) is selectively activated under hypoxic conditions and results in cyclin D (\( Ccnd \)) down-regulation (Conrad et al., 2000), as well as cell cycle regulation in response to ionizing radiation (Wang et al., 2000). \( p38 \delta \) represents a point of crosstalk with the ERK1/2 pathway as these proteins co-immunoprecipitate (Efimova et al., 2003). Moreover, the ubiquitous expression pattern of \( \alpha \) and \( \beta \) family members, contrasted with the largely tissue-restricted expression \( \gamma \) and \( \delta \) isoforms could again imply specific functions (Jiang et al., 1996).

Genetic studies have revealed differences in the knockout phenotypes of each isoform, \( p38 \alpha \) mutation causes embryonic lethality owing to defects in placental angiogenesis, whereas \( p38 \beta \) mutant mice are viable. This difference in phenotype may not be due to redundancy in the classical sense and may be a consequence of their different expression levels, as \( p38 \alpha \) is the more highly expresses of the two during early development (source: http://www.ncbi.nlm.nih.gov/UniGene/). Both \( p38 \gamma \) and \( p38 \delta \) knockouts were viable and displayed no overt abnormalities (Sabio et al., 2005).

1.1.5.3.3 JNK1/2/3

Like p38, JNK family members are activated in response to a plethora of stress signals as well as growth factor stimulation and protein translation inhibition (reviewed in Bogoyevitch et al., 2010). JNK family members share a high degree of sequence similarity and are expressed in and overlapping manner in multiple tissues, although JNK3 is enriched in neural tissues (Bode et al., 2007 and references therein). JNK proteins are activated by MEK4/7 phosphorylation on the TPY motif on their activation lip, where upon they translocate to both the cytoplasm and nucleus to phosphorylate their targets. The canonical JNK target is c-JUN (Hibi et al., 1993), where phosphorylation by JNK proteins leads to an increase in its transcriptional activity through promotion of (Activator Protein-1) AP-1 complex formation and transcription of genes containing AP-1-binding sites, including genes that control the cell cycle, such as \( Ccnd \), as well as epithelial-to-mesenchymal transition (EMT).
associated transcripts. Moreover, sustained JNK activity can correlate with an increase in ERK2 activity, via insulin receptor substrate-2, resulting in the expression of FBJ osteosarcoma oncogene gene (cFOS), a binding partner of c-JUN, thus enhancing its activity. Such a mechanism demonstrates how JNK activity can coordinate multiple pathways to elicit a change in gene expression and cellular phenotype (Wang et al., 2009).

1.1.5.3.4 Erk5

ERK5 is unique amongst the typical MAPKs in that as well as its kinase domain it also possesses a C-terminal domain with a nuclear localisation signal (NLS) and a putative transactivation domain (Kaiser et al. 2000). It is activated in response to calcium flux, growth factors, as well as members of the TNF superfamily (Zou et al., 2006). ERK5 is phosphorylated on multiple residues, both MEK5 dependent TEY phosphorylation on its activation loop, and MEK5 independent, cell cycle regulated, phosphorylation of residues in its C-terminal extension (Diaz-Rodriguez et al. 2010).

There is a growing body of evidence to suggest that ERK5 and ERK1/2 may have overlapping targets and thus exhibit a degree of functional redundancy. Work from the lab of Melanie Cobb has shown that the ERK5 can bind and activate p90 ribosomal s6 kinase (p90RSK) proteins, previously thought to be exclusively ERK1/2 dependent (Ranganathan et al., 2006). ERK5 can also activate the ERK1/2 substrate myelocytomatosis oncogene (c-MYC) (English et al., 1998), as well as induce immediate early genes (IEG) c-Fos and c-Jun. Moreover, ERK5 can contribute to epidermal growth factor (EGF) induced cell proliferation in MCF10A cells (Kato et al., 1998). However the fact that the Erk5 knockout mouse largely copies the Mek5 mutant, whose phenotype is distinct from either Erk1 or Erk2 mutants, implies that there may be a developmental requirement for Erk5 that does not overlap with Erk1/2 (Regan et al., 2002).
1.1.5.3.5 Atypical MAPKs

ERK3/4, ERK7/(8) and nemo-like kinase (Nlk) belong to the atypical MAPK subfamily whose regulation and general functions are poorly understood. These MAPKs are considered atypical because they either lack the conventional TxY activation motif (ERK3/4 and NLK) or because there is yet no MEKK or MEK known to activate them (reviewed in Cargnello et al., 2011). Genetic studies have indicated roles for Erk3 in embryonic growth (Klinger et al., 2009), which is not compounded by Erk4 knockout, which itself is viable, although mice display mild depression-like behaviors, arguing for non-redundant roles of the two closely related proteins (Rousseau et al., 2011). Nlk knockout mice show genetic background-specific phenotypes, on a C57Bl/6 background the mutated allele results in late embryonic lethality, where on a 129/Sv background the allele is viable, however mice exhibit various neurological abnormalities, as well as defects in stromal differentiation (Kortenjann et al., 2001). No mutant has yet been reported for Erk7, although experiments in HCT15 human colorectal cancer cells indicates that it is capable of activating c-JUN and promoting neoplastic transformation (Xu et al., 2010).

1.1.5.4 MAPK-activated protein kinases

Although in the classical sense MAPKs represent the terminal kinase in the MAPK pathway, often times they exert their function through the activation of a family of related proteins, the MAPK-activated protein kinases (MAPKAPKs). MAPKAPKs belong to a family of 11 related protein kinases that share similarities in their kinase domains. Based on sequence homology and activation requirements Mapkapks can be divided into 5 subgroups, p90RSK, mitogen- and stress-activated protein kinases (MSK), MAP kinase-interacting serine/threonine kinases (MNK), MAP kinase-activated protein kinases (MK) 2/3, and MK5 (reviewed in Gaestel, 2008). Like MAPKs, MAPKAPKs show a minimal substrate consensus motif of Pro
directed kinases. Specificity of each MAPKAPK for its upstream MAPK is achieved in part by differences in the MAPK CD domain (Tanuoe et al., 2001).

MK2/1 appear to be exclusively activated by p38 family members, as chemical inhibition of both p38α and β result in complete abrogation of MK1/2 activity (Clifton et al., 1998). Although initially described as a p38 MAPKAPK, MK5 doesn’t appear to stably bind p38 proteins, however, yeast-2-hybrid screens have identified the atypical MAPKs ERK3/4 as stably binding partners. Further *in vivo* studies showed that Mk5 could only be co-precipitated with ERK3, however as Mk5 mutant mice are viable it is not the sole mediator of ERK3 function (Shi et al., 2003).

MNK1 was identified in a yeast-2-hybrid screen using ERK2 as bait (Waskiewicz et al., 1997 and references therein) and subsequent identification of MNK2 showed they exhibited a high degree of homology in their kinase domain (~70%). Both genes are expressed as long and short splice variants, the short isoforms lack a C-terminal extension that is essential for interactions with the CD domains of ERK1/2 as well as p38, however MNK2 appears to show a preference for ERK1/2 interactions (Waskiewicz et al., 1997). Double knockout of Mnkl/2 showed that their combined activity is essential for constitutive activation of eIF4E (Ueda et al., 2004).

MSK1/2 proteins were first discovered in a yeast-2-hybrid as p38 interacting proteins and have since been shown to be activated in a p38-dependent manner in response to various cellular stresses, as well as by growth factor stimulated ERK1/2 (reviewed in Hague et al., 2006). MSK proteins posses two distinct kinase domains, a C-terminal kinase domain (CTKD) involved in autophosphorylation and an N-terminal kinase domain (NTKD) for substrate binding that belongs to the AGC (PKA, PKG and PKC) family of kinase domains (Hanks et al., 1995). Binding of active MAPK to a docking site on the C-terminal of MSK results in phosphorylation of several residues within the C-terminal kinase domain, as well as the hinge region. Phosphorylation of the CTKD results in its activation and
subsequent phosphorylation of several residues in the hinge region as well as the NTKD which leads to full MSK activation (reviewed in Pearce et al., 2010).

MSK proteins are predominantly nuclear in stimulated and quiescent cells and once activated they phosphorylate both activating transcription factor (ATF)-1 as well as cAMP response element-binding (CREB). Both ATF1 and CREB have been implicated in the regulation of the IEGs Egr1, cFos and JunB, with double Msk1/2 knockout mouse embryonic fibroblasts (MEF) showing a 50% reduction in their expression in response to cellular stress. However a less severe reduction was observed following mitogen stimulation indicating that MSK proteins show varying degrees of context dependent redundancy in control of IEG expression (Wiggin et al., 2002). Interestingly double knockout of Atf1 and Creb results in peri-implantation lethality due to developmental arrest. Ex vivo culture of mutant embryos (E3.5) show that they fail to hatch from the zona pellucida and eventually undergo necrosis. However, Atf1+/− Creb−/− embryos are capable of hatching and support trophoblast expansion, but no cells of the inner cell mass (ICM) were evident after 4 days ex vivo culture indicating a possible dose dependent function (Bleckmann et al., 2002).

Similar to MSK proteins, p90RSK also possess dual kinase domains of similar structure and function (Fisher et al., 1996). At present 4 p90RSK proteins have been identified that share ~75% sequence homology. Rsk1-3 transcripts have been detected in every tissue tested, although their relative abundance differs dramatically. Whereas Rsk4 displays a more restricted expression pattern in the adult, but is ubiquitously expressed at low levels throughout development (Kohn et al., 2003).

p90RSK proteins have a more complex mechanism of activation that that of MSK and require input of both ERK1/2 and 3-phosphoinositide dependent protein kinase (PDK)-1 signals (Figure 1.7). In quiescent cells a pool of ERK1/2 is complexed with p90RSK proteins through interactions with a docking motif at the C-terminal end of the CTKD. Following stimulation and activation ERK1/2 phosphorylates the hinge region and the CTKD on several residues that are conserved from worms to man (Dalby et al., 1998). This set of phosphorylation
events both activate the CTKD as well as relaxes the hinge region. CTKD phosphorylation of Ser380 in the hinge region acts as a docking site for PDK1. PDK1 then phosphorylates Ser221 in the NTKD that results in its activation and subsequent NTKD-dependent phosphorylation of Ser749 in the ERK1/2 docking domain, causing ERK1/2 to disassociate. The kinase activity of p90RSK is dependent entirely on both ERK and PDK inputs as no p90RSK activity is detected in either Pdk1 null or MEK inhibited cells (Williams et al., 2000, Mody et al., 2001). Such a mode of regulation results in dynamic control of both p90RSK, and indirectly ERK activity: in cells where PDK1 is inactive, the ERK1/2 signaling output might be reduced as a portion of both enzymes will be complexed with inactive p90RSK due to a lack of NTKD dependent phosphorylation of the ERK docking site. However upon full PDK1-dependent p90RSK activation phosphorylated ERK is released from the p90RSK complex thus increasing the concentration of free active ERK. Interestingly experiments into the function of the Drosophila p90RSK homologue, S6KII, indicated that it functioned as a negative regulator of ERK activity during development by acting as a cytoplasmic anchor for ERK (Kim et al., 2006). S6KII deletion resulted in hyperactivation and a developmental phenotype reminiscent of ERK gain of function, whilst overexpression resulted in cytoplasmic retention of ERK. In mammalian cells inhibition of p90RSK NTKD activity due to Pdk1 deletion also resulted in hyper ERK phosphorylation indicating that p90RSK proteins may be involved in a negative feedback loop (Williams et al., 2000 and references therein). Furthermore, experiments in p90Rsk2 knockout mice showed that p90RSK2 depletion also coincided with ERK1/2 hyperphosphorylation in skeletal muscle extracts (Dufresne et al., 2001).

p90RSK proteins exert their functions through interactions with multiple pathways. p90RSKs regulate a number of transcription factors, including CREB, serum response factor (SRF), nuclear factor-κB (NF-κB), NFATc4, NFAT3 and the transcription initiation factor TIF1A (reviewed in Frodlín et al., 1999). The canonical function of p90RSK in guiding cFos expression was identified in fibroblasts derived from patients with Coffin-Lowery syndrome that lack Rsk2 (De Cesare et al., 1998).
However p90RSK activity is also required for cFOS protein stabilization indicating a general role in cFOS function (Murphy et al., 2002). p90RSK proteins have also been associated with the mitogen induced increase in translation (Angenstein et al., 1998) possibly by activation of eIF4B and increasing the rate of cap-dependent translation (Holz et al., 2005). Due to the proposed high degree of redundancy between p90RSK family members the actual physiological significance of functions is not well understood. Rsk1/2/3 triple knockout mice have been made and are reported to be viable, however no analysis of these animals has to date been published (Dumont et al., 2005).

![Diagram of p90Rsk proteins activation](image1)

**Figure 1.7: Mechanism of activation of p90Rsk proteins.** In quiescent cells p90Rsk and Erk are bound in an inactive conformation. Following Erk activation a series of phosphorylation events converging on the hinge region and CTKD result in CTKD activation, PDK1 binding and exposure of the NTKD. The NTKD is then phosphorylated by PDK1 resulting in its activation and the phosphorylation of p90Rsk downstream targets.

### 1.1.5.5 The use of small molecule inhibitors to probe MAPK function

Over the past 2 decades enormous progress has been made in the inhibition of kinases by small molecule inhibitors (reviewed by Cohen, 1999). It is estimated that that about 30% of the research and development programs of the pharmaceutical industry are currently focused on protein kinases (Bian et al., 2003) which is not surprising as aberrant phosphorylation underlines many common malignancies such as cancers as well as diseases such as rheumatoid arthritis. High-throughput library
screens (Raj et al., 2010) as well as structure based design (Cohen et al., 2007) have resulted in the discovery of a broad range of ‘specific’ kinase inhibitors.

Many protein kinase inhibitors bear structural similarities to ATP, although cannot be metabolized by protein kinases, therefore they inhibit kinase function by competing with ATP for kinase domain binding or by binding the kinase domain and locking the activation loop in an inactive conformation (reviewed Bogoyevitch et al., 2005 and Zhang et al., 2009), which leads to several important considerations as to both their in vivo specificity and efficacy. The compound must be able to freely diffuse across cell membranes to maintain an active cellular concentration and once inside a cell it must be able to discriminate between the ATP-binding sites of all protein kinases, as well the other proteins that utilise ATP, in an environment rich in intracellular ATP. As the specificity of many small molecules is tested in vitro against panels of recombinant kinases the exact efficacy and specificity of the molecule in vivo cannot be gauged. This is confounded by the fact that kinase specificity is generally a function of its relative concentration (Bian et al., 2003), which can be affected by cell to cell differences in the expression of many kinases. This being said however, the development and implementation of approaches such as Kinobead™ profiling will help to address the issue of in vivo specificity (Bantscheff et al., 2007).

Another class of kinase inhibitor acts in an allosteric manner, binding to a region site outside of the ATP-binding site to regulate kinase function. This class of inhibitor is typified by the second-generation MEK inhibitors PD184352 and PD0325901 (Bian et al., 2003 and references therein). These compounds seem to bind inactive more readily than active MEK1/2 and inhibit MEK function by locking the molecules into a closed and inactive conformation (Ohren et al., 2004). Although allosteric inhibitors of MEK1/2 are highly specific, this specificity is still concentration dependent, as culture medium concentrations in the range of 0.5-2µM of PD0325901 inhibited MEK5 activity following EGF stimulation, though concentrations of PD184352 in the range of 10µM were needed to elicit a similar effect, however, PD184352 is also a less potent inhibitor of MEK1/2. This being said when a panel of 65 kinase inhibitors were tested in vitro against ~80 protein kinases,
allosteric MEK inhibitors were the only compounds that showed exclusivity for their reported target (Bian et al., 2003). However, if a phenotype is observed for a specific molecule, it is prudent to confirm this by using a second, structurally unrelated, molecule targeting the same kinase.

1.2 The RAF-MEK-ERK pathway

Activation of the RAF-MEK-ERK pathway can follow many routes and result in dramatically different outputs depending on the cellular context, the input signal, and the subcellular localisation of the activated module. Moreover, this core module can be ‘shared’ between different receptors, activation of which result in opposing responses such as seen in the PC12 model whereby EGF (through epidermal growth factor receptor (EGFR)) induced ERK activity results in proliferation, whilst nerve growth factor (NGF) (through transforming tyrosine kinase protein A (TrKa)) induction of ERK results in differentiation (Marshall et al., 1995). As our ability to accurately describe the components of a pathway, as well as the kinetic parameters in which they operate increases, so does our understanding of the mechanisms that confer signal specificity and reproducibility in the face of biological variation.

1.2.1 Receptor activation

ERK1/2 phosphorylation is observed following activation of a host of various membrane-associated receptors, some of the most well studied being the growth factor receptors: fibroblast growth factor receptors (FGFR) EGFR, TrK, and insulin receptors, as well as integrins, and cytokine receptors such as glycoprotein 130 heterodimers. Activation of the receptor by its cognate ligand results in a complex series of conformation changes and phosphorylation events that lead to recruitment and activation of RAS proteins and subsequently the MAPK module. The recruitment of RAS proteins is mediated by combined activities of adapter proteins such as GRB and FRS. It has been shown that the different affinities of these adapters for various receptors can have profound effects on the duration of ERK
activity (Sasagawa et al., 2005). As the signal is amplified along the cascade, approximately three-fold at every step (Schilling et al., 2010), even small differences in adapter:receptor affinities can translate into a large difference in the pathway output.

Another early point of signal control is at the rate of receptor turnover: ligand-stimulated endocytosis of occupied receptors and subsequent intracellular degradation of both ligand and receptor molecules (Sorkin et al., 2009). Receptor-ligand complexes are internalized into clathrin-coated pits wherein the lower pH of the endosomal lumen cause receptor-ligand disassociation. From here the receptor can be either targeted for degradation or recycled to the membrane, either of which will affect membrane receptor occupancy and therefore the quality of the signal. Moreover, active receptors can continue to signal from endosomes in the cytoplasm where the substrates they encounter may be different from those at the plasma membrane, thus again augmenting the initial signal output (Choudhary et al., 2009).

Another important aspect of receptor activation that can affect the MAPK output is the co-activation of parallel pathways and the cross talk between them. In the PC12 system, NGF stimulation results in both MAPK activation but also in activation of protein kinase B (PKB) by PI3K (Von Kriegsheim et al., 2009). PKB phosphorylates phosphoprotein enriched in astrocytes (PEA) 15 which is a negative regulator of ERK nuclear import. PKB dependent phosphorylation of PEA-15 on Ser116 coincides with ERK:PEA disassociation and nuclear accumulation of ERK, a process that was blocked by the PI3K inhibitor LY294002.

1.2.2 Scaffolds

Another way by which signaling specificity is achieved is by arranging the MAPK module into multi-protein complexes that can be tethered to a specific cellular location or regulated en-masse by targeting the scaffolding proteins themselves. Such scaffolds can act as regulators of signal magnitude, timing, as well as serve as points of crosstalk with other pathways (reviewed by Kolch, 2005). Signal regulation through scaffolds play a pivotal role in the *S. cerevisiae* mating
pheromone response (Strickfaden et al., 2007) where phosphorylation of the MAPK scaffold STE5 by cell cycle kinase cyclin dependent kinase (CDK) actively inhibits its association with active receptors, thus denying receptors access to MAPK substrates and blocking α-factors induced signaling. Such scaffold targeting ensures that MAPK signaling is restrained to a discrete period of the cell cycle and thus acts to integrate both internal and external cues.

In mammalian cells, interactions between Ras proteins, the scaffold kinase suppressor of RAS (KRS) and impedes mitogenic signal propagation (IMP) act to regulate signal magnitude, without affecting signal duration or timing (Matheny et al., 2004, Matheny et al., 2006). In quiescent cells KSR and IMP associate with 14-3-3 and the MAPK module in an inactive complex. Receptor activation and the subsequent activation of RAS by the SOS-GRB complex results in IMP auto-ubiquitination, release of the KSR-MAPK complex from its sequestered subcellular compartment and translocation to the membrane where upon interactions between Ras and Raf result in pathway activation.

As discussed earlier, dual phosphorylation of ERK by MEK is generally thought to require two separate interactions whereby each residue on the TEY motif is phosphorylated individually and thus the reaction proceeds in a distributive manner, as opposed to processive activation where both residues are phosphorylated in a single interaction (Schilling et al., 2010). A distributive mechanism of activation has been shown to result in bistable, switch-like ERK responses (Markevich et al., 2004), where processive activation results in a more graded, linear Erk response (Figure 1.8). Both linear and nonlinear ERK responses have been described experimentally (Xiong et al., 2003, Mackeigan et al., 2005), with linear and nonlinear activities identified within the same cell type but arising from different stimuli (Santos et al., 2007). It has been proposed that the activation of scaffolds concentrating MEK and ERK can result in ‘quasi-processive’ ERK activation: though the inherent biochemical reaction is distributive because the time between disassociation and re-association becomes negligible as the local concentration of
both MEK and ERK increase, the system behaves as a linear processive response (Levchenko et al., 2000).

Taking a focused quantitative mass spectroscopy (MS) approach, work from Ursula Klingmuller lab showed that ERK activation followed a distributive mechanism in colony-forming unit erythroid stage (CFU-E) cells following erythropoietin stimulation, with Tyr being phosphorylated first, and then Thr (Schilling et al., 2010). However, in HeLa cells, Aoki and colleagues argue that the apparent distributive ERK activation, i.e. the presence of both pTEY and pTEpY isoforms, arises from the action of phosphatases dephosphorylating pTEpY peptides and that in the absence of phosphatase activity, ERK activation appears processive and a consequence of molecular crowding rather than scaffold intervention (Aoki et al., 2011). Therefore the involvement of scaffold proteins in the nature of ERK activation remains a much debated subject and is most likely to be cell-type and stimulus dependent.

Figure 1.8: Processive Vs Distributive Erk activation. Theoretical Erk activation profiles based on the kinetics of Mek binding.

1.2.3 Interactions within the MAPK core

Both direct and indirect regulatory interactions within the MAPK core i.e. between RAF-MEK-ERK have been proposed to be a major determinate of signal quality and duration (Santos et al., 2007). Employing a novel conceptual tool, modular response analysis (MRA) that allows the connectivity of several points in a network to be deduced by calculating the change in activity of each point following
perturbations to another at steady state (Kholodenko et al., 2002), Santos and colleagues showed that positive and negative feedforward and feedback connections existed within the core MAPK module, and that the wiring of these connections changed with both stimulus and time. For example, they show that EGF stimulation resulted in appreciable negative feedback from ERK to RAF at 5 minutes post stimulation, whilst NGF stimulation resulted in positive feedback at the same timepoint, the magnitude of which increased with time. However, coincident with the positive feedback to Raf, negative feedback from ERK to MEK was also observed, an order of magnitude greater for NGF stimulation than EGF. Although the MRA approach cannot give any information as to whether a connection is direct or indirect, experimental data has shown that ERK1/2 can directly inhibit MEK function (Catalanotti et al., 2009), or indirectly enhance MEK activity through RAF Kinase Inhibitor Protein (RKIP) (Shin et al., 2009 and references therein). Similarly, RAF activity is subject to direct negative regulation (Fritsche-Guenther et al., 2011), indirect negative regulation (Douville et al., 1997), and direct positive regulation (Balan et al., 2006) (Figure 1.9).

Several eloquent arguments have been made for the purpose of such feedback and feedforward regulations (Shin et al., 2009, Birtwistle et al., 2011, Kholodenko et al., 2010) all with a common theme: that due to the stochastic nature of cell signaling processes, coupled with the signal noise that cells must deal with in a multicellular organism such signaling loops are essential to:

- discriminate between transient and sustained stimulation
- create a time-lag to allow integration of other signals
- constrain a signal within certain parameters

A fundamental challenge facing the field is to determine how multiple regulatory mechanisms, sometimes with opposing functions, are integrated to form a predictable response. Moreover, from a practical point of view, as most of this behavior is assessed on a population level, determining ERK responses to various signals on a single cell level, even at the subcellular level, is bound to aid our understanding of the molecular mechanisms underpinning signal specificity.
1.2.4 The Immediate Early Gene paradigm

Immediate early gene (IEG) is a very loose term for a gene whose expression is rapidly and transiently induced by a stimulus, and that expression of this gene does not require new protein synthesis. IEG up-regulation can be observed within minutes after stimulation and tends to return to basal levels between 45 and 90 minutes (Nakakuki et al., 2010). The rapid expression of IEGs has been linked to their genomic structure, as in general IEGs have fewer introns and are expressed as shorter transcripts that other induced genes (Tullai et al., 2007). IEGs are distinct from delayed early, or delayed primary response genes (DEG) because of this rapid expression, whereas there is an appreciable lag between stimulation and expression of DEG transcripts (30-120 minutes) (Tullai et al., 2007). Late response genes (LRG) on the other hand require new protein synthesis and are generally regulated by IEGs (Vincent et al., 1993). Which genes fall into which categories is stimulus and cell
type dependent as is the functional response of the cell. This being said there appears to be some conceptual generalities that explain the evolution for such a multiphasic transcriptional response to a stimulus. Nagashima and colleagues showed using the MCF-7 system that the magnitude of the transcriptional response was directly proportionate to the extent of receptor activation (Nagashima et al., 2007), and proposed that although many signaling pathways converge on overlapping IEGs, it is the magnitude of expression that regulates the overall phenotypic outcome. On the other hand, expression of late response genes, such as D-type cyclins that regulate cell cycle progression and are the canonical output of a biphasic mitogenic response, are dependent on the duration of signal as opposed to its magnitude (Murphy et al., 2002). LRG sensitivity to signal duration is achieved, in the PC12 system at least, because ERK activity is required for IEG induction but also for stabilization for many of their protein products such as cFOS and possibly MYC (Sears et al., 2000). When ERK is transiently stimulated its activity returns to basal levels by the time IEG transcripts are translated and thus IEG protein products are degraded and there is no LRG expression. However, when ERK stimulation is sustained IEG products are stabilized and continue to guide the second wave of transcription. This form of regulatory network may function as a noise filter ensuring only sustained signals are converted into functional responses. However, it is possible that oscillatory signals could be interpreted by this network in a similar manner, depending on the time lag between each oscillation.

### 1.2.4.1 IEG expression through Serum Response Factor

SRF is an evolutionary conserved single-copy transcription factor with a MCM-1, AMAGOUS, DEFICIENS, AND SRF (MADS) box. SRF binds as a dimer to the consensus sequence (CC(A/T)₆GG), known as a CArG box and regulates the expression of many genes involved in cell cycle regulation, cytoskeletal dynamics, and survival (Norman et al., 1988, Pellegrini et al., 1995 Medjkane et al., 2009, Ebisuya et al., 2008, Posern et al., 2002). The protein is divided into two primary domains: the N-terminal MADS that contains the SRF DNA binding and also
mediates interactions with various binding partners, and a C-terminal transactivation domain (Pellegrini et al., 1995). Regulation of IEGs by SRF is dependent on mitogen activation of accessory cofactors of the Ets family of proto-oncogenes, namely ELK-1, SAP-1 and NET-1, which are collectively referred to as Ternary Complex Factors (TCFs) (Whitmarsh et al., 1995, and reviewed in Dalton et al., 1993). Ets proteins bind to a purine rich sequence with a core motif of GGAA/T that is located immediately adjacent to many CArG boxes in the promoters of immediate early genes. DNA binding and ternary complex formation is dependent on MAPK phosphorylation of multiple sites on TCFs (Gille et al., 1992).

Recently a whole genome in silico screen for novel CArG box containing promoters has been employed and has identified several novel SRF target genes (Sun et al., 2005). The authors identified 87 novel SRF target genes, many of which were validated by EMSA, Luciferase reporter assays, or mRNA expression following Srf knockdown. This study further reinforced the role for SRF in regulating several IEGs, but also the expression of many cytoskeletal components.

Induction of the IEGs cFos and Egr1 by either serum or 12-O-tetradecanoylphorbol 13-acetate (TPA) appears to be largely SRF dependent in mES cells (Schratt et al., 2001). Moreover, Srf null mES cells do not display any obvious proliferation defect indicating that SRF induction of IEGs is dispensable for cell cycle progression. However, many IEGs induced by FGF4 in mES cells are known to possess SRF binding motifs in their promoters (Table 1.2). As FGF4 activity is essential for mES cell differentiation (Kunath et al., 2007), SRF may be involved in lineage specification, rather than proliferation, of mES cells.
Table 1.2: FGF4 stimulated Srf IEG targets in mES cells. Microarray analysis of gene expression responses to FGF4 stimulation. FGF4-/- mES cells were stimulated with recombinant hFGF4 (5ng/ml) for the indicated times. Expression was considered significant if it showed 1.5 fold over background at 1 hour, and then declined below the significance threshold by 6h (T. Kunath, unpublished).

<table>
<thead>
<tr>
<th>Fold 1h</th>
<th>Fold 6h</th>
<th>Symbol</th>
<th>Known IEG</th>
<th>Srf Target</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.80</td>
<td>0.63</td>
<td>Egr1</td>
<td>Yes</td>
<td>Yes</td>
<td>Sukhatme et al. 1988</td>
</tr>
<tr>
<td>5.08</td>
<td>1.24</td>
<td>Fos</td>
<td>Yes</td>
<td>Yes</td>
<td>Mullen et al. 1984</td>
</tr>
<tr>
<td>2.26</td>
<td>1.34</td>
<td>Ctgf</td>
<td>Yes</td>
<td>Yes</td>
<td>Ryseck et al. 1991</td>
</tr>
<tr>
<td>2.18</td>
<td>1.33</td>
<td>Cyr61</td>
<td>Yes</td>
<td>Yes</td>
<td>O’Brien et al. 1990</td>
</tr>
<tr>
<td>2.12</td>
<td>1.49</td>
<td>LOC620695</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>2.12</td>
<td>1.34</td>
<td>Jun</td>
<td>Yes</td>
<td>Yes</td>
<td>Ryseck et al. 1988</td>
</tr>
<tr>
<td>2.11</td>
<td>1.38</td>
<td>Arc</td>
<td>Yes</td>
<td>No</td>
<td>Lyford et al. 1995</td>
</tr>
<tr>
<td>1.99</td>
<td>1.48</td>
<td>Postn</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.96</td>
<td>1.33</td>
<td>Vglil3</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.94</td>
<td>1.06</td>
<td>Junb</td>
<td>Yes</td>
<td>Yes</td>
<td>Ryder et al. 1988</td>
</tr>
<tr>
<td>1.89</td>
<td>1.10</td>
<td>Dsg1a</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.89</td>
<td>1.05</td>
<td>EG629991</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.85</td>
<td>0.98</td>
<td>Egr2</td>
<td>Yes</td>
<td>Yes</td>
<td>Chavrier et al. 1988</td>
</tr>
<tr>
<td>1.81</td>
<td>1.05</td>
<td>Sct</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.78</td>
<td>1.27</td>
<td>ler2</td>
<td>Yes</td>
<td>Yes</td>
<td>Latinkic et al 1994</td>
</tr>
<tr>
<td>1.78</td>
<td>1.25</td>
<td>Fem1b</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.76</td>
<td>1.01</td>
<td>Hpca</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.67</td>
<td>1.02</td>
<td>Khlcd8a</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.65</td>
<td>0.94</td>
<td>Itga10</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.64</td>
<td>1.11</td>
<td>Pdch19</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.62</td>
<td>1.03</td>
<td>A330076H08Rik</td>
<td>No</td>
<td>No</td>
<td>Barone et al. 1994</td>
</tr>
<tr>
<td>1.61</td>
<td>0.86</td>
<td>Usp42</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.57</td>
<td>1.04</td>
<td>Tmem117</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.57</td>
<td>0.99</td>
<td>Nab2</td>
<td>Yes</td>
<td>No</td>
<td>Svaren et al. 1996</td>
</tr>
<tr>
<td>1.55</td>
<td>0.59</td>
<td>Calccoc2</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>1.52</td>
<td>0.96</td>
<td>Ank</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
1.3 ERK activity during early embryogenesis

Signaling between various tissues in development is essential for establishing cellular identity and the correct body plan of the developing animal (reviewed in Tam et al., 2007). As an embryo develops cells specialise and progressively loses the ability to adopt different identities, a process known as lineage restriction (Figure 1.10). There is considerable data to suggest that MAPK function, specifically FGF activation of the RAF-MEL-ERK pathway is crucial for some of the earliest stages of lineage specification, as well as maintenance of those lineages, in the developing mouse embryo.

1.3.1 Early murine development

Following fertilization the embryo undergoes several rounds of early cleavage divisions to generate multiple cells of decreasing size, known as blastomeres, without changing the overall size of the embryo. At the one cell stage in development the proteins of the embryo are derived from maternally provided transcripts however by the end of the cleavage cycles the zygotic genome is activated and the residual maternal transcripts are degraded (Paynton et al., 1988). The cells of the 8-cell embryo then become increasingly more adherent, expressing E-cadherin, in a process known as compaction. During compaction the blastomeres begin to polarize exhibiting apical and basolateral domains and by the 16 cell morula stage two spatially distributed cellular populations are apparent: outside polar and inside nonpolar (Johnson et al., 1981). At the 32 cell stage a fluid filled cavity known as the blastocoel begins to form, which coincides with the first lineage segregation event and the formation of the inner cell mass (ICM) and the trophectoderm (TE). The ICM will continue to form all the tissues of the embryo proper, as well as some extraembryonic tissues (Gardner et al., 1988) whereas the TE will produce the tissues of the trophoblast lineage, including the ectoplacental cone, extraembryonic ectoderm, and giant cells of the early embryo, which will later contribute to the fetal portion of the placenta.
The mutually exclusive expression patterns of two transcription factors, octamer-binding transcription factor 4 (OCT4) and caudal type homeobox 2 (CDX2), are thought to be important in establishing this lineage segregation (Strumpf et al., 2005, Nichols et al., 1998). *Oct4* mutant embryos express TE markers through the ICM and fail to support ICM derived lineages *in vivo*, or ICM outgrowths *ex vivo* (Nichols et al., 1998). Conversely, *Cdx2* mutant embryos express ICM markers OCT4 and NANOG throughout the ICM as well as the cells of the putative TE, whereas *ex vivo* outgrowths from *Cdx2* null embryos fail to attach or generate trophoblast giant cells.

Evidence suggests that *Cdx2* expression regulated by the interactions of TEA domain family member 4 (TEAD4) and Yes-activated protein-1 (YAP1) and that the positional regulation of YAP1 activity may be a contributing factor in ICM and trophoblast lineage segregation (Yagi et al., 2007). YAP1 subcellular localisation is thought to be controlled by the Hippo signaling cascade which results in nuclear exclusion of YAP1 from the nuclei of the inner cells of the developing embryo and thus helps to establish the CDX2:OCT4 expression domains (Nishioka et al., 2009).
As the embryo develops the blastocoel cavity enlarges and by E4.5 a third embryonic lineage has emerged, the primitive endoderm (PE), which forms a monolayer between the ICM/epiblast cells and the blastocoel cavity. A similar scenario to ICM:trophoblast segregation is seen in epiblast:PE lineage segregation with each lineage identifiable by the mutually exclusive expression patterns of specific transcription factors, namely NANOG (epiblast) and GATA6 (PE) (reviewed in Cockburn et al., 2010). Prior to PE segregation many cells of the ICM co-express both NANOG and GATA6 (Plusa et al., 2008). As the ICM develops the expression pattern of NANOG and GATA6 expression becomes more exclusive and each is expressed in a ‘salt and pepper’ pattern. By E4.5 expression domains have been established and distinct NANOG expressing epiblast and GATA6 expressing PE domains are evident. *Gata6* knockout embryos do develop a PE layer, however they exhibit defects in PE-derived tissues such as the visceral endoderm (Koutsourakis et al., 1999) and the lack of a more severe PE phenotype may be due to redundancy.

Figure 1.10: Early murine embryogenesis.
between GTAT6 and GATA4. Furthermore, cells injected with a dominant-negative GATA6 tend to reside in the center of the ICM and exhibit poor contribution to the PE lineage (Morris et al., 2010). Knockout of Nanog is peri-implantation lethal owing to defects in the ICM (Mitsui et al., 2003). At E3.5 Nanog null and WT embryos appear morphologically indistinguishable, however outgrowths from mutant embryos fail to support the epiblast and differentiate into parietal endoderm-like cells (Mitsui et al., 2003), in contrast to Oct4 mutant outgrowths where trophectoderm differentiation was observed thus implying a transcriptional hierarchy for early embryo lineage segregation where OCT4 acts to maintain ICM identity during trophoblast specification and subsequently NANOG has a similar role during PE specification. However the recent finding that Nanog mutant mES cells can contribute to somatic tissues of chimeric embryos, but not the germ line, would imply that in the correct context Nanog mutant cells do not assume a default PE fate (Chambers et al., 2007).

Around this stage of development the blastocyst hatches from the zona pellucida, the glycoprotein membrane that surrounds the embryo, and implants into the uterine wall. Post-implantation the embryo continues to grow and lineages subdivide. The polar trophectoderm forms the ectoplacental cone and extraembryonic ectoderm, whereas the cells of the mural trophectoderm become polyploid and form the trophoblast giant cells. The extraembryonic ectoderm acts as an important source of positional signals that help to define the embryonic anterior-posterior axis. Both visceral and parietal endoderm are clearly distinguishable at this stage with the parietal endoderm lining the inside of the mural trophectoderm/giant cells and the visceral endoderm lining the outside of the epiblast cells and extraembryonic ectoderm (Figure 1.10). The parietal endoderm will generate supportive tissue for the growing embryo such as the Reichert's membrane, whereas the visceral endoderm will further subdivide into anterior- and posterior visceral endoderm, each contributing important positional cues that help to establish the anterior-posterior axis (reviewed in Tam et al., 2007). By this stage (E5.5-E6.5) the epiblast has formed a cavitated columnar epithelium surrounding the pro-amniotic cavity. Positional cues
from the extraembryonic ectoderm and the anterior visceral endoderm are important for establishing the primitive streak (Stuckey et al., 2011), the presence of which marks the beginning of gastrulation. The primitive streak forms at the posterior end of the embryo where cells appear to undergo an epithelial to mesenchymal transition in so losing local adhesion and become motile. Cells invaginate and move through the primitive streak to form mesoderm or to be incorporated as definitive endoderm (reviewed Tam et al., 2007). By E7.5 (late gastrulation) the three germ layers have been formed and organogenesis and fetal development ensues.

1.3.2 The role of FGF-MEL-ERK signaling in the developing embryo

Several genetic studies over the past two decades have implicated FFG signaling to be an important factor in both the survival and correct lineage segregation of the early embryo. Early studies into the function of Fgf4 demonstrated that homozygous mutant embryos were present at normal Mendelian ratios until E3.5, but following that there was a sharp decrease in mutant embryo survival with disorganized mutant embryonic structures visible at E5.5 indicating FGF4 is necessary for peri-implantation development (Feldman et al., 1995). Ex vivo culture of mutant embryos showed that trophectoderm formation was not compromised however survival of the ICM as well as primitive endoderm differentiation appeared to be dependent on FGF4. The lack of primitive endoderm differentiation is consistent with previous studies that showed that ex vivo ICM culture in the presence of recombinant hFGF4 increased the proportion of primitive endoderm cells (Rappolee et al., 1994). Interestingly, FGFR2 is heterogeneously expressed in the early blastocyst with the most robust expression in cells of the trophoblast, but also cells lining the blastocoel cavity where the PE resides, indicating that paracrine FGF4 signaling through FGFR2 may be involved in PE specification/maintenance (Haffner-Krausz et al., 1999).

A subsequent report that GGB2 deficient blastocysts display aberrant ICM lineage segregation indicated that FGF4 may be acting through the RAS MAPK pathway at this stage of development (Cheng et al., 1998). Grb2 mutant embryos
appear normal until E3.5, and though they are capable of implanting only empty deciduae were detected at later time-points. *Ex vivo* culture showed that GRB2 was not necessary for trophectoderm differentiation in culture, however mutant blastocysts outgrowths lacked expanded ICM cells and contained very few endodermal cells. Experiments with *Grb2* mutant mES cells showed that expression of a constitutively active HRAS mutant was largely able to rescue the *Grb2* phenotype indicating the involvement of MAPK signaling (Cheng et al., 1998). Re-examining the *Grb2* phenotype some years later, work from Janet Rossant’s lab showed that GRB2 activity was necessary for proper segregation of GATA6 and NANOG expressing cells in the ICM. GRB2 deficient E3.5 embryos lacked expression of several primitive endoderm markers and ubiquitously expressed the epiblast marker NANOG. However, there was no significant difference in the total number of ICM cells between WT and mutant embryos arguing that the lack of PE population was due to a specification defect rather than a maintenance defect. The presence of a layer of CDX2 expressing cells indicated again the trophoblast specification was not affected.

Using small molecule inhibitors of both FGFR and MEK1/2 Nichols and colleagues, showed that segregation of the ICM into PE and epiblast was dependent on MEK-ERK activity (Nichols et al., 2009). Subsequent work demonstrated that this segregation could be driven, in a reversible manner, towards the PE lineage by treating early blastocysts with high doses of recombinant FGF4 (Yamanaka et al., 2010). Culture of embryos from the 8-cell stage in either the MEK inhibitor PD0325901 alone, or in combination with the FGFR inhibitor PD173074 resulted in a block in PE specification at E3.5. Moreover, similar to experiments with *Grb2* mutant embryos, cell numbers were largely consistent between control and drug-treated embryos again indicating a defect in specification rather than maintenance. ppERK1/2 activity has been reported in the cells of the E3.5 blastocyst (Yang et al., 2006) which appears ubiquitous and cytoplasmic. However, it has been reported that there is a shift from predominantly nuclear to cytoplasmic ppERK1/2 between E2.5 and E3.5 (Wang et al., 2004) which may correlate with the initial signal that
separates ICM cells into PE and epiblast fates. However, Erk1/2 double knock-out blastocysts have not been described at present, therefore there is insufficient genetic data to firmly implicate ERK1/2 activity in this process.

Genetic studies into the specific functions of Erk1 and Erk2 revealed gene specific phenotypes with only Erk2 being essential for embryogenesis. Disruption of the Erk2 allele results in post-implantation lethality, attributed to either failed mesoderm induction, trophoderm defects, or placental abnormalities (Saba-El-Leil et al., 2003, Yao et al., 2003, Hatano et al., 2003). Both Saba-El-Leil and Yao disrupted Erk2 by knocking-out a crucial exon that codes for part of the kinase domain (Saba-El-Leil: exon 3, Yao: exon 4) and observed embryonic lethality between E6.5 and E8.5. Hatano and colleagues chose to disrupt Erk2 gene function by knocking a gene trap cassette into intron 2 thus trapping the endogenous transcript, also within the kinase domain. Erk2 depletion by this method resulted in embryonic lethality between E11.5 and E12.5. As both Hatano et al. and Saba-El-Leil et al. used similar breeding schemes to generate their homozygous mutant embryos and genetic backgrounds were matched between both experiments the discrepancy between both phenotypes may be due to low frequency splicing around the gene trap insertion cassette and may represent a hypomorphic allele.

Erk2 mutant embryos fail to generate an ectoplacental cone as well as extraembryonic ectoderm and show signs of resorption by E8.5. In morula aggregation experiments Erk2 mutant embryos contribute exclusively to the epiblast (Saba-El-Leil et al., 2003). At E6.5 high levels of ppERK1/2 are detectable in the extra-embryonic ectoderm of WT embryos (Corson et al., 2003), however, Yao and colleagues failed to detect phosphorylated ERK1 in this region in Erk2 mutant embryos (Yao et al., 2003), even though Erk1 mRNA is detectable at this stage of development (Saba-El-Leil et al., 2003). However this discrepancy may be due to the different sensitivities of each approach.

Yao and colleagues attributed the cause of embryonic lethality to defects in mesoderm differentiation as mutant embryos failed to express the mesoderm marker brachyury (T) (Yao et al., 2003). Embryos lacking FgfR1 gene function also exhibit
mesoderm related defects at this stage of development (Deng et al., 1994). Mutant embryos are capable of primitive streak initiation, however epiblast cells tend to accumulate in an inwardly protruding streak region. Ex vivo culture of embryos in the presence of the FGFR1 inhibitor indicated that the ERK activity observed within the streak may not be FGFR dependent (Corson et al., 2003) and therefore it is difficult to conclude whether the observed defect in mesoderm induction observed by Yao and colleagues is cell-autonomous or a consequence of improper patterning signals derived from the extraembryonic ectoderm, where ERK1/2 is also highly active (Corson et al., 2003). Recent work from the lab of Sylvain Meloche has helped to address this issues (Voisin et al., 2010). Using tetraploid complementation assay (Nagy et al., 1993), whereby Erk2−/− embryos (8-cell) were aggregated with 4n WT 2-cell embryos Voisin and colleagues were able to generate live embryos at E12.5 where the embryo proper was derived from ERK2 deficient cells, albeit at a low frequency (2 of 36). This data would suggest that ERK2 is not necessary for the development of the early epiblast, as well as formation of the 3 primary germ layers, and that the embryonic lethality of ERK2 mutant embryos can be most likely traced back to defects in the extraembryonic ectoderm.

Mutations in Erk1 on the other hand do not appear to severely affect the developing embryo (Pages et al., 1999, Atler et al., 2010, Selcher et al., 2001, Mazzucchelli et al., 2002). Erk1 knockout mice are born at the expected Mendelian ratios, are fertile and of normal size, however thymocyte maturation beyond the CD4+CD8+ stage was reduced by half in mutant mice, with a similar reduction in the thymocyte subpopulation expressing high levels of T cell receptor (Pages et al., 1999). ERK1 depletion has also been reported to have an enhancement of striatum-dependent long-term memory (Mazzucchelli et al., 2002). Tissue specific double knockout models for Erk1 and Erk2 have been described (Satho et al., 2011, Voisin et al., 2010) and as such represent a valuable tool to study ERK redundancy in vivo. Specifically it would be most interesting to analyse Erk1/2 double knockout early embryos and compared the phenotypes to the pharmacological data presented by Nichols and colleagues (Nichols et al., 2009).
1.3.3 Functional redundancy of Erk1 and Erk2

The case for isoform specific functions of ERK1 and ERK2 is most often founded on the different phenotypes of the knockout mice. At present there is insufficient information available to be able to make a definitive statement either way, however it is still a vehemently argued topic (Indigro et al., 2010).

Although ERK1 and ERK2 are 83% homologous and exhibit similar activation profiles in vitro (Robbins et al., 1993), there are considerable differences in their N-terminal sequence. Recent work has shown that this sequence may act as a regulatory motif for nuclear transport. Marchi and colleagues showed that ERK1 shuttled between the cytoplasm and the nucleus at a slower rate than ERK2, and that deleting the divergent N-terminal sequence from ERK1 conferred ERK2-like transport kinetics (Marchi et al., 2008). They also proposed that this slower nucleo-cytoplasmic shuttling was functionally significant as it made ERK1 less efficient at transmitting signals to the nucleus, and went on to show that ERK1, or a fusion of the N-terminal sequence from ERK1 with the kinase domain of ERK2, was less efficient at facilitating oncogenic RAS transformation. Interestingly, sequence analysis shows that this N-terminal sequence is only slightly conserved from fish to man (Figure 1.11). Furthermore, this region is the most divergent a.a. stretch between man and mouse contributing 8 of the 11 divergent residues between the homologues. It would therefore be of interest to validate this model in cell lines of other species to determine how general this phenomenon may be.

The observation that MEFs from Erk1 deficient mice exhibited enhanced ERK2 activity, and proliferation rates, led to the notion that ERK1 acted as an ERK2 antagonist, or partial agonist to fine tune ERK2 signaling (Indigro et al., 2010, Vantaggiato et al., 2006). It was shown that knockdown of Erk1 enhanced NIH 3T3 proliferation, whereas similar knockdown of Erk2 inhibited it. It was also shown that expression of either Erk1 or a kinase defective mutant inhibited RAS induced tumour formation in mice (Vantaggiato et al., 2006). From this data the authors proposed a model where ERK1 competes with ERK2 for MEK activation, however ERK1 is a
‘weaker’ signal transducer and therefore this competition results in a reduction of pathway activity.

Reciprocal control over ERK isozyme activation has been proposed in other systems, however it has been argued to occur at the level of differential negative feedback loops rather than competition for MEK binding (Schilling et al., 2010). Sensitivity analysis on a model of EPO stimulated ERK activity in CFU-E cells showed reciprocal control coefficients between ERK isoforms, and that this control was exerted through inhibition of SOS. The authors continued to demonstrate experimentally that overexpression of either ERK isozyme led to a dramatic reduction in the proliferation of CFU-E, as determined by isotopic thymidine incorporation, and there appeared to be a degree of isoform specificity involved, however a concomitant increase in differentiation was observed upon ERK overexpression.

It has been recently shown that differences in the relative concentrations of the various MAPK cascade components can generate a high degree of variation in pathway behavior without affecting the topology of the pathway (O’Shaughnessy et al., 2010). By recreating a tunable mammalian MAPK pathway in yeast cells O’Shaughnessy and colleagues were able to demonstrate that different concentrations of the MAPK core components resulted in quantitatively different pathway outputs.

![Sequence alignment of the N-terminal region of Erk1 and Erk2.](image)

**Figure 1.11:** Sequence alignment of the N-terminal region of Erk1 and Erk2. Red box denotes the putative nuclear transport signal.
such that concentration of MKKK, 3nM; MEK, 1200nM; MAPK, 330nM resulted in high-ultrasensitivity, whereas concentrations of MKKK, 41nM; Mek, 37nM; MAPK, 470nM exhibited low-ultrasensitivity. Although this phenomenon was only theoretically demonstrated it does indicate that changing the relative components of the MAPK pathway, i.e. through gene knockdown or overexpression, can result in behaviors that may not be physiologically relevant to the cell type/process being studied and therefore care should be taken when generalizing observations from such studies. How this model fits with the observation that cells exhibit a high degree of cell-to-cell variation in protein abundance is unclear (Sigal et al., 2006).

As discussed previously, the lack of phenotypic overlap between Erk1 and Erk2 knockout models is often sited as proof of lack of redundancy and there are some instances where this may be the case (Lips et al., 2004). A factor that significantly confounds interpretation of the various Erk family member knockout models is that the ratio of ERK1 and ERK2 expression levels is highly variable across many tissues (Table 1.3), and therefore manifestation of a phenotype upon depletion of one isoform may simply be a consequence of insufficient expression of the other, i.e. the total ERK dose is now below a functional threshold.

Two recent papers argue for such a model (Lefloch et al., 2008, Voisin et al., 2010). Lefloch and colleagues showed that single silencing of Erk1 had a negligible impact on the proliferation rates of NIH 3T3 fibroblasts, however a considerable reduction was observed upon Erk2 depletion. In this cell line the expression ratio of ERK1 to ERK2 is approximately 1:4. The authors noticed that only when ERK2 levels were severely depleted additional knockdown of Erk1 resulted in a reduction in proliferation. The difference between combined and single Erk silencing may occur because in this cell line ERK levels are in excess of what is required to efficiently transmit the mitogenic stimulus. Therefore when ERK1 is reduced, there is sufficient levels of ERK2 remaining to compensate, however, when ERK2 activity is limiting, then further reducing the pool of ERK by Erk1 knockdown has an additive affect on proliferation.
Taking a genetic approach Voisin and colleagues assessed the consequence of isozyme specific depletion on the proliferation of genetically defined MEF lines (Voisin et al., 2010). In this study depletion of Erk1 had a negative impact on proliferation in MEFs from a predominantly CD-1 background, but not on MEFs from a predominantly C57Bl/6 background. Importantly, the ratio of ERK1 to ERK2 was different between backgrounds with CD-1 MEFs expressing the isoforms (Erk1:Erk2) at 1:1.4 whilst C57Bl/6 MEFs exhibiting a ratio of 1:1.9 underscoring not only the importance of genetic background on MAPK pathway topology, but also resolving the discrepancies between the relative contribution of ERK1 to fibroblast proliferation.

Another confounding issue is the possibility of tissue specific modes of feedback regulation. Multiple studies have looked at isozyme specific functions of ERK1/2. As outlined above, this is generally accomplished by disrupting gene function by gene targeting or by gene knockdown under which circumstances compensatory activation of the remaining isoform is often observed. Analysis of the effect of knockdown on the activation level of the other isoform in various tissue types, genetic backgrounds, and species revealed highly variable instances of compensatory phosphorylation (Table 1.4). The reason for such differences are not obvious, and are likely to be due to a combination of both biological and

Table 1.3: Tissue specific expression profile of MAPK isoforms based on EST data.
Source http://www.ncbi.nlm.nih.gov/unigene. Spot intensities based on transcripts per million. Empty cells indicate tissues for which no sequence was detected. Expression data was not available for murine bRaf.
experimental factors. However, EST data indicates that RAF and MEK isozymes exhibit variation in their tissue specific expression patterns (Table 1.3). Moreover, there is evidence that suggests isoforms of each family are affected differentially by ERK dependent negative feedback (Eblen et al., 2004, Balan et al., 2006). Therefore it is tempting to suggest that feedback regulation may contribute to cell type specific ERK compensation.

<table>
<thead>
<tr>
<th>Isoform Depleted</th>
<th>Method Used</th>
<th>Compensatory Phosphorylation</th>
<th>Tissue</th>
<th>Species</th>
<th>Genetic Background</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erk1</td>
<td>shRNA</td>
<td>Yes</td>
<td>NIH 3T3</td>
<td>Mus Musculus</td>
<td>Swiss</td>
<td>Lefloch et al., 2008</td>
</tr>
<tr>
<td>Erk2</td>
<td>shRNA</td>
<td>Yes</td>
<td>NIH 3T3</td>
<td>Mus Musculus</td>
<td>Swiss</td>
<td>Lefloch et al., 2008</td>
</tr>
<tr>
<td>Erk1</td>
<td>shRNA</td>
<td>No</td>
<td>FAO</td>
<td>Rattus norvegicus</td>
<td>ND</td>
<td>Basserd et al., 2006</td>
</tr>
<tr>
<td>Erk2</td>
<td>shRNA</td>
<td>No</td>
<td>FAO</td>
<td>Rattus norvegicus</td>
<td>ND</td>
<td>Basserd et al., 2006</td>
</tr>
<tr>
<td>Erk1</td>
<td>shRNA</td>
<td>Yes</td>
<td>Primary Hepatocytes</td>
<td>Rattus norvegicus</td>
<td>Sprague-Dawley</td>
<td>Franlin et al., 2007</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>Yes</td>
<td>MEF</td>
<td>Mus Musculus</td>
<td>ND</td>
<td>Vantaggiato et al., 2006</td>
</tr>
<tr>
<td>Erk2</td>
<td>shRNA</td>
<td>No</td>
<td>MEF</td>
<td>Mus Musculus</td>
<td>ND</td>
<td>Vantaggiato et al., 2006</td>
</tr>
<tr>
<td>Erk2</td>
<td>Gene targeting</td>
<td>No</td>
<td>mES</td>
<td>Mus Musculus</td>
<td>CS7B6-129 (F1)</td>
<td>Kunath et al., 2007</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>Yes</td>
<td>mES</td>
<td>Mus Musculus</td>
<td>129/0la</td>
<td>Yao et al., 2003</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>No</td>
<td>Hippocampus</td>
<td>Mus Musculus</td>
<td>129/SvImJ (98%)</td>
<td>Secher et al., 2001</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>Yes</td>
<td>Spinal Cord</td>
<td>Mus Musculus</td>
<td>CS7B6/1</td>
<td>Alber et al., 2010</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>No</td>
<td>MEF</td>
<td>Mus Musculus</td>
<td>CS7B6/1</td>
<td>Bost et al., 2005</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>Yes</td>
<td>ex vivo cultured Stratum</td>
<td>Mus Musculus</td>
<td>CS7B6/1</td>
<td>Mazuccelli et al., 2002</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>Yes</td>
<td>ex vivo cultured Cortex</td>
<td>Mus Musculus</td>
<td>CS7B6/1</td>
<td>Mazuccelli et al., 2002</td>
</tr>
<tr>
<td>Erk1</td>
<td>Gene targeting</td>
<td>No</td>
<td>Antigenous</td>
<td>Thromocyte</td>
<td>Mus Musculus</td>
<td>ND</td>
</tr>
<tr>
<td>Erk2</td>
<td>Gene targeting</td>
<td>Yes*</td>
<td>Multiple Neural tissues</td>
<td>Mus Musculus</td>
<td>CS7B6/1</td>
<td>Saito et al., 2009</td>
</tr>
<tr>
<td>Erk2</td>
<td>Gene targeting</td>
<td>No</td>
<td>ES.5 ExE</td>
<td>Mus Musculus</td>
<td>CS7B6-129 (F1)</td>
<td>Yao et al., 2003</td>
</tr>
</tbody>
</table>

**Table 1.4: Differential Erk compensatory activation.** *Compensatory hyperphosphorylation of the remaining pool of Erk2 following partial Erk2 knockdown.

### 1.4 mES cells

mES cells are derived from the ICMs of pre-implantation embryos and represent a clonal pluripotent cell line that retains the ability to differentiate into all cell types of the developing embryo and adult (Evans and Kaufman, 1981, Martin, 1981, Brook and Gardner 1997). In addition to pluripotency mES cells exhibit a second defining characteristic in that they are capable of self-renewing in culture, possibly indefinitely, without acquiring karyotypic abnormalities, and still maintaining the ability to contribute to the three germ layers, as well as the germ line when reintroduced back into the developing embryo (Robertson et al., 1986).
mES cells appear to represent a transient population of cells within the ICM of the pre-implantation that can be effectively ‘captured’ in culture under the correct conditions (Silva and Smith, 2008) However, historically the derivation of mES cells was an inefficient process restricted largely to the 129 mouse strain (Buehr et al., 2003). As our knowledge of the regulatory mechanisms that support the pluripotent state has developed, so too has our ability to control mES cells which has led to the now routine derivation of mES cells with high efficiency (Nichols et al., 2009) as well as derivation from refractory genetic backgrounds (Ying et al., 2008), and species (Buehr et al., 2008, Li et al., 2008).

mES cells exhibit several characteristics that make them an excellent cell line in which to study many cellular processes:

- they are easy to grow and require relatively simple growth media.
- they double rapidly (12-15h) which can be useful for biochemical studies where large amounts of tissue are needed.
- they exhibit a high frequency of homologous recombination ($10^{-5}$-$10^{-6}$) (Tempelton et al., 1997) which makes them useful in genetic studies.
- they can be differentiated into various cell types with high efficiency.

The popularity of mES cell research has led to a wealth of large and small scale studies on the regulatory mechanisms that act to maintain the pluripotent state in vitro and it is now clear that a complex interplay between intrinsic and extrinsic factors exist.

### 1.4.1 Extrinsic factors that regulate pluripotency

#### 1.4.1.1 Leukemia inhibitory factor

The initial derivation of mES cells was dependent on soluble molecules present in serum, or secreted by a layer of mitotically inactivated embryonic fibroblasts, or Buffalo rat liver cells (Evans and Kaufman, 1981, Martin, 1981). The subsequent identification of the leukemia inhibitory factor (LIF) as the peptide secreted by feeder cells, necessary for the maintenance of undifferentiated mES cells
in culture, began the rational interrogation of the molecular mechanisms regulating pluripotency (Smith et al., 1988, Williams et al., 1988). Although LIF is expressed by mES cells (Chambers et al., 2003), the levels to which it is expressed appear insufficient to maintain robust mES cell propagation.

LIF belongs to the interleukin family of cytokines and activates a heterodimeric receptor consisting of two related receptors, the LIF receptor (LIFR), and the GP130 receptor (Davis et al., 1993) resulting in the activation of both Janus Kinase/Signal Transducers and Activators of Transcription (JAK/STAT) and MAPK pathways (reviewed in Brudon et al., 2002). The primary STAT molecule activated by LIF signaling is STAT3 and overexpression of a constitutively active version of this molecule allows mES cells to self-renew in the absence of LIF (Matsuda et al., 1999). Although LIF is required in vitro to derive and preserve mES cells, in vivo mouse embryos that lack LIF can develop to a stage subsequent to mES cell derivation, similar to Stat3−/− embryos (Takeda et al., 1997), indicating there must be other factors active in vivo that help to maintain pluripotency, or that since pluripotent populations are transient in the developing embryo LIF signaling is only necessary for the capture and maintenance of mES cells in culture, similar to the involvement of LIF in diapause (Nichols et al., 2001). Moreover, LIF is insufficient to maintain pluripotency in defined culture conditions (Ying et al., 2003), where the emergence of SOX1+ neural precursor cells is readily observable in the presence of LIF. This indicates that there are other soluble factors present in serum that are essential for the maintenance of the undifferentiated state.

ChIP:Chip analysis showed that STAT3 occupies the promoters of many genes involved in self-renewal and pluripotency, as well as lineage specific genes (Kidder et al., 2008). Comparison of STAT3 promoter occupancy with expression analysis showed that in the undifferentiated state the promoters of pluripotency associated genes bound by STAT3 were highly expressed, whereas the opposite was true of lineage specific promoters bound by STAT3 indicating that LIF signaling through STAT3 may be involved in maintaining the expression of mES cell related
transcripts whilst simultaneously repressing lineage specific expression profiles (Kidder et al., 2008).

The proto-oncogene \emph{cMyc} is another important target of Lif-Jak/Stat signaling in mES cells (Cartwright et al., 2005). STAT3 binds the promoter of \emph{cMyc} in mES cells and positively regulates its expression, whereas cMyc protein stability is negatively regulated by GSK3\(\beta\) phosphorylation of Thr58. Overexpression of a c\textsc{MYC} T58A mutant allows Lif independent self-renewal and blocks differentiation in embryoid body (EB) assays (Cartwright et al., 2005).

1.4.1.2 Bone morphogenic protein 4

Bone morphogenetic protein (BMP) signaling as a key component in the maintenance of an undifferentiated state (Ying et al., 2003). Members of the TGF-\(\beta\) superfamily of secreted signaling molecules, BMPs have important functions in many biological contexts. They bind to specific serine/threonine kinase receptors, which transduce the signal to the nucleus through Similar to Mothers Against Decapentaplegic Homologue (SMAD) proteins (reviewed in van Bubnoff and Cho, 2001). BMP activation of SMAD4 culminates in the expression of inhibitor of differentiation (ID) proteins, which act to block differentiation in a lineage specific manner, as BMPs have been shown to have powerful anti-neuorgenic ability \textit{in vivo} (Di-Gregorio et al., 2007). Overexpression of ID2 in mES cells blocks neural commitment in the presence of LIF (Ying et al., 2003), and promotes non-neural ectoderm differentiation in its absence (Kunath et al., 2007). Furthermore, mES cell derivation is greatly enhanced in the presence of BMP4 under serum-free conditions (Ying et al., 2003), whilst derivation of BMPR1 deficient mES cells appears to be impossible under standard conditions (Qi et al., 2004).

1.4.2 Transcriptional Regulation Of mES Cell Identity

Our present understanding of the transcriptional regulation of mES cell identity is that it is supported by the combined activities of a small central
transcriptional network composed of the transcription factors OCT4, NANOG and SOX2. This core transcriptional circuit is a key regulator of its own expression as well as regulating various other transcription factors and epigenetic regulators that are involved in self-renewal and pluripotency (Loh et al., 2006, Ivanova et al., 2006, Niwa et al., 2009).

1.4.2.1 Oct4

OCT4 was discovered in the late 80’s as a factor expressed in mES cells and the early embryo (Scholer et al., 1989). Subsequent work showed that it was essential for the maintenance of both (Nichols et al., 1998). Genetic ablation of Oct4 results in peri-implantation lethality with morphologically normal blastocysts at E3.5 (Nichols et al., 1998). In vitro analysis of mutant blastocysts showed a loss of pluripotent cells from blastocyst outgrowths and an abundance of trophectoderm like cells.

Structurally Oct4 consists of an octamer-binding domain separating a proline rich N-terminal and a serine/threonine rich C-terminal domain. Both octamer domain and C-terminal regions are generally referred to as the POU domain and make sequence specific contact with DNA (reviewed in Scholer, 1991). The complex nature of OCT4 function is reflected in its biochemistry with both N and C-terminal domains possessing transactivator function, with the C-terminal domain being phosphorylated in a tissue specific manner (Brehm et al., 1997). Moreover both transactivation domains are interchangeable with each other, and with other members of the OCT family, although there are some differences in target gene activation (Niwa et al., 2002).

Sequence conservation analysis has revealed four highly conserved regions in the Oct4 promoter, proximal and distal enhancers, which have been shown to recruit both transcriptional activators and repressor and to be methylated upon the onset of differentiation (Pan et al., 2002 and references therein). Both directed and unbiased ChIP experiments have shown that the Oct4 locus is bound by a myriad of factors including LRH-1 (Gu et al., 2005a), GCNF (Gu et al., 2005b), NANOG (Loh...
et al., 2006), CDX2 (Niwa et al., 2005), RXRβ and SP1 (Sylvester and Scholer, 1994) and synergistically by a complex consisting of OCT4 and SOX2 (Loh et al., 2006) all suggesting that levels of OCT4 are tightly regulated in the developing embryo and mES cells. Indeed an increase of 2 fold above endogenous levels triggers differentiation into primitive endoderm and meso/endoderm like cells, and conversely, a reduction in steady state levels results in differentiation into trophectoderm lineages (Niwa et al., 2000). Microarray analysis has shown that there are different sets of genes associated with differentiated states/specific lineages expressed at either high or low levels of OCT4 and that pluripotency genes are only expressed when OCT4 is maintained at levels endogenous to self-renewing mES cells (Niwa et al., 2006).

1.4.2.2 Sox2

OCT4 has been shown to function, in part, through binding to SOX2, a member of the SRY-related HMG box of transcription factors. This synergistic activity was first observed in mES cells at a distal enhancer in the 3’ UTR of the Fgf4 gene (Yuan et al., 1995) where both OCT4 and SOX2 guide the developmental expression of this gene. Since then a host of embryonic stem cell specific genes have been shown to have functional *cis*-acting OCT/SOX elements, including *Utf1* (Nishimoto et al., 1999), *Fbx15* (Tokuzawa, et al., 2003), *Nanog* (Kuroda et al., 2005), *Oct4* (Okumura-Nakanishi et al., 2005), whilst ChIP:Chip experiments in human ES cells have identified many other genes, including miRNA clusters, bound and possibly regulated by OCT4:SOX2 heterodimers (Boyer et al., 2005). The necessity for SOX2 in maintenance of the pluripotent state is again underscored from genetic studies where ablation of both *Sox2* alleles results in peri-implantation lethality with a complete absence of an OCT4 positive epiblast (Avilion et al., 2003).

The above data notwithstanding, there is growing evidence of a certain level of interchangeability between members of the *Sox* gene families in the context of OCT:SOX element activation (Wiebe et al., 2003, Maruyama et al., 2005). Indeed SOX4, SOX11 and SOX15 can all be ChIPed to the OCT:SOX elements of *Fgf4*,

52
Nanog, Lefty1 and Oct4 in self-renewing mES cells (Masui et al., 2007). Moreover, putative OCT:SOX target genes show only a modest decrease in expression levels up to 48h after Sox2 excision indicating that other SOX factors can, to a certain extent, substitute for SOX2 in this context.

1.4.2.3 Nanog

Nanog was identified by multiple groups: in a screen for molecules that can support self-renewal and pluripotency (Chambers et al., 2003), by degenerate PCR for homeobox-containing genes involved in early embryonic development (Wang et al., 2003), and in a bioinformatics screen for mES cell associated transcripts (ECATs) (Mitsui et al., 2003). It is a homeodomain protein with homology to members of the NK2 gene family though it is sufficiently divergent (more than 50%) to sit in a class of its own (Wang et al., 2003).

NANOG protein can be divided loosely into 3 domains; and N-terminal; a homeodomain; and a C-terminal domain that contains the well-conserved tryptophan (Trp) repeat domain (Mullin et al., 2008). Both N- and C- domains contain a transactivator function (Pan et al., 2003), whilst the Trp-repeat is essential for dimerisation, which is itself essential for normal NANOG function (Wang et al., 2008). Initial genetic evidence indicated that NANOG was absolutely essential for the maintenance of the pluripotent state (Mitsui et al., 2003), and this notion was reinforced by studies using siRNA knockdown (Ivanova et al., 2006), and whilst Nanog−/− mES cells could be derived, they rapidly differentiated into primitive endoderm. However, recent studies have shown that it is possible to generate Nanog+/− mES cells that are capable of sustained pluripotency and chimera contribution (Chambers et al., 2007). NANOG deficient cells show a decreased self-renewal efficiency and higher propensity to differentiate, which decreases in a stepwise quantitative manner from Nanog+/− to Nanog+/− to Nanog++ cells. This data, when taken together with over-expression experiments, underscore the dose dependent effect of NANOG.
Nanog expression is subjected to a similar level of complex regulation as seen at the Oct4 locus, with both functional OCT:SOX and GCNF motifs as well as binding sites for STAT3, TCF3, FOXD3, p53, T, and NANOG itself (reviewed in Pan et al., 2007). Several of the factors known to regulate the Nanog locus, such as OCT4, FOXD3 and p53, can act as either activators or repressors depending on the cofactor context indicating dynamic control of Nanog expression in mES cells.

1.4.3 Epigenetic regulation of mES cell identity

Control of chromatin dynamics, DNA methylation status and the subtle alterations in transcript stability brought about by miRNAs can have a dramatic effect on the transcriptome of mES cells. There is a host of genetic and biochemical evidence indicating an absolute requirement for several epigenetic effectors in the maintenance of a pluripotent epiblast, and in the concomitant downregulation of pluripotent transcriptional profiles and eventual lineage specification (reviewed in Surani et al., 2007). However the permissive or instructive nature of epigenetic regulation is still a debated subject (reviewed in Silva and Smith, 2008).

Perhaps the best-studied epigenetic process is DNA methylation (Bird, 2002), where approximately 70% of all CpG dinucleotides are methylated, and associated with regions of heterochromatin. The remaining 30% of unmethylated CpG dinucleotides are generally associated with the gene promoter regions and referred to as CpG islands. DNA methylation may affect gene expression by two means, methyl groups may sterically hinder transcription factors, blocking their binding to DNA, and methylated DNA can act as binding sites for methyl DNA binding proteins, that can in turn recruit histone acetyl/methyl transferases altering the chromatin structure thus occluding transcription factor access to gene promoters. The earliest stages of development are accompanied by a general erasure of both maternal and paternal methylation patterns before the onset of zygotic transcription (Mayer et al., 2000). This process does not include a certain set of genes, known as imprinted genes, whose expression patterns are stably inherited. However, there is a
programmed change in the pattern of CpG methylation during embryogenesis indicating that this process may play a role in cell fate specification, or at least the stable maintenance of gene expression profiles (Reik et al., 2001). The enzymes that govern this process are members of the DNA methyltransferase family, specifically DNMT3a/b, the de novo DNA methyltransferase involved in re-methylating the genome following implantation (Okano et al., 1999). Genetic ablation of both enzymes results in aberrant expression of both NANOG and OCT4 during differentiation (Li et al., 2007).

Along with the regulation of transcription enforced by DNA methylation, the tails of histones, the core proteins that with DNA constitute chromatin, is also subjected to a vast number of post-translational modifications that alter their association with each other and with the DNA helix, which combined lead to stale changes in gene expression (reviewed in Berger, 2007).

mES cells appear to possess a relaxed and transcriptionally permissive chromatin structure, resulting from a high level of activating histone modifications such as H3K4 tri-methylation (tri-me) (reviewed in Bibikova et al., 2008, and in Zipori et al., 2004). The observation that regions of H3K4 tri-me co-localize with areas of repressive chromatin markings, such as H3K9/27 tri-me, and that these regions are enriched at loci containing developmentally important genes, has led to the idea of bivalent domains, or pioneer loci (Bernstein et al., 2006). In this model, genes associated with specific lineages are ‘poised’ for activation by the overlapping presence of these opposing chromatin markings thus allowing for rapid activation of lineage specific transcriptional profiles.

Mutants for components of chromatin modification machinery generally exhibit a post-implantation defect indicating that the establishment and maintenance of pluripotency is not compromised in their absence (reviewed in Niwa, 2007). Although it can be argued that if a cell is not capable of differentiating it must then be considered unipotent, and therefore pluripotency has been compromised. Conversely, if a genetic mutation is compensated for, by expression of a transgene
for example, and the cell regains potency then it can be argued that pluripotency was never lost, just blocked by the mutation (Silva and Smith, 2007).

1.4.4 miRNA

microRNAs (miRNA) are small sequences of non-coding RNA that bind to complementary RNA sequences on the open reading frames and 3’ UTRs of expressed genes and act to regulate mRNA translation and degradation (Bartel et al., 2004). miRNAs are expressed either as intronic sequences or as miRNA clusters and are extensively processed in both the nucleus and the cytoplasm by distinct protein complexes resulting the liberation of a 21 nucleotide long mature miRNA. Deletion of the miRNA processing component \textit{Dgcr8} in mES cells attenuates both their ability to self-renewal as well as downregulate mES cell markers upon differentiation (Wang et al., 2008, Wang et al., 2007). \textit{Dicer}, another processing component, deficient embryos die early in development and mES cells cannot be derived from mutant ICMs (Bernstein et al., 2003), however \textit{Dicer} mutant mES cells can be generated \textit{in vitro} and they exhibit severe growth and differentiation defects and prolonged G1 and G0 phases of the cell cycle (Kanellopoulou et al., 2005).

A recent report demonstrated opposing roles for miRNA in mES cells (Melton et al., 2010). It was reported that two miRNA families, \textit{let-7} and ES cell cycle (ESCC) act in opposition to each other to indirectly regulate pluripotency associated genes. \textit{ESCC} miRNA enhance \textit{Lin28} expression which actively repressed the expression of the \textit{let-7} cluster. Exogenous expression of \textit{let-7} results in spontaneous differentiation mES cell differentiation. However \textit{let-7} and \textit{ESCC} miRNAs are never co-expressed at high levels, and therefore the authors conclude that the purpose of \textit{let-7} is not to promote differentiation but to stabilizes the switch between self-renewal and differentiation by repressing pluripotency associated transcripts. Consistent with this notion is the observation that there is a dramatic decrease in the half-life of 80% of the mES cell transcriptome when differentiation is induced by retinoic acid treatment (Sharova et al., 2009).
1.4.5 mES cell heterogeneity

Embryonic stem cells have traditionally been thought of as a homogenous population of cells corresponding to a specific developmental potential. However, recent experiments using fluorescent protein knock-in strategies for candidate pluripotent marker genes and high throughput *in-situ* hybridization screens (Toyooka et al., 2008, Singh et al., 2008, Carter et al., 2008, Hayashi et al., 2008) have shown that the expression levels of an ever increasing list of transcription factors displays a marked deviation in any given population of cells and that this heterogeneity may have functional implications for the developmental potential of the cell.

The idea of heterogeneous gene expression is not unique to ES cells; cells of the ICM have been shown to express heterogeneous and mutually exclusive levels of NANOGg and GATA6 (Chazaud et al., 2006), and that this mosaic expression may correlate with the specification of cells for either epiblast or primitive endoderm. Perhaps it isn’t that surprising that this heterogeneity is also recapitulated in culture (Chambers et al., 2007, Singh et al., 2008). Reversible heterogeneity has been shown for several pluripotency associated transcription factors, such as NANOG (Chambers et al., 2007), the zinc-finger protein REX1 (Toyooka et al., 2008) and the PGC marker STELLA (DPPA3) (Hayashi et al., 2008). Heterogeneous expression of lineage specific transcription factors such as the PE markers HEX (Canham et al., 2010) and GATA6 (Singh et al., 2006) have also been reported.

Recent work from Minoru Ko’s lab has identified many other genes that show heterogenous expression patterns in mES cells. Employing high throughput mRNA *in situ* hybridisation Carter et al. identified 4 high confidence hits (*Zscan-4, Rex1, Whsc2* and *Rhox9*) and 8 others (*Rest, Atf4, Pa2g4, E2f2, Nanog, Dppa3, Esrrb, and Fscn1*) with a more ambiguous heterogeneity (Carter et al., 2008) indicating that cell-to-cell variations in gene expression may be a common feature in mES cells.
1.5 The role of FGF4 signaling in mES cell biology

mES cells express FGF4 in an autocrine fashion that promotes the activation of FGFRs and multiple downstream pathways (Kunath et al., 2007, Stavridis et al., 2007). Expression of Fgf4 is driven by OCT4 and SOX2 through an enhancer element in the Fgf4 3′UTR that is active in cells of the blastocyst as well as mES cells and embryonic carcinoma cells (EC) (Yuan et al., 1995, Boer et al., 2005, Luster et al., 2003). mES cells appear to express Fgfr1-2-3, however knockout studies in the embryo may imply different functions for them during development (Arman et al., 1998, Deng et al., 1994, Colvin et al., 1996). Experiments using a dominant-negative FGFR2 showed that its expression could seriously attenuate mES cell differentiation in embryoid bodies and that this correlated with a reduction in ERK1/2 and PI3K activity (Chen et al., 2000). Although FGF4 depletion in the early embryo is lethal, and ex vivo cultured ICM cells fail to proliferate (Feldman et al., 1995) mES cells lacking Fgf4 are viable and display no obvious proliferative defect under self-renewing conditions, however once differentiation is induced by LIF withdrawal and retinoic acid treatment there is a dramatic reduction in cell survival (Wilder et al., 1997). More recent studies have shown that Fgf4 is essential for neural induction under monolayer conditions (Kunath et al., 2007).

Fgf:receptor interactions are enhanced in the presence of heparin sulfate (HS) in a concentration dependent and receptor specific manner (Aviezer et al., 1999). HS synthesis at the Golgi is dependent on the action of two enzymes EXT1 and EXT2 (Kraushaar et al., 2010 and references therein). Genetic depletion of Ext1 in mES cells drastically reduced their differentiation potential with an accompanying block in FGF signaling without any reported attenuation on cell survival (Kraushaar et al., 2010). Moreover, Ext1 mutant cells displayed a dramatic reduction in ppERK levels, however treatment with the pan-FGFR inhibitor PD173074 was capable of further reducing ERK phosphorylation.
1.5.1 PI3K signaling

PI3K is activated at the plasma membrane by a variety of extracellular signals such as growth factors as well as integrin signaling (reviewed in Bozulic et al., 2009). PI3K phosphorylates phosphatidylinositol (4,5) bisphosphate (PIP2) to generate phosphatidylinositol (3,4,5) triphosphate (PIP3) which acts as a second messenger, binding proteins with pleckstrin homology (PH) domains including protein kinase B (PKB, also known as AKT) and PDK1. This reaction is inhibited by phosphatase and tensin homolog (PTEN). Recruitment of both PKB and PDK1 to the plasma membrane results in phosphorylation of PKB on Thr308 by PDK1 which is facilitated by PKB Ser473 phosphorylation by the mammalian target of rapamycin (mTORC2)(Rictor containing) complex (Sarbassov et al., 2005). PKB activity inhibits both TSC1/2 mediated inhibition of mTORC1 complex activity and GSK3 activity as well as apoptotic proteins BAD and CASPASE9 to regulate many aspects of proliferation, survival, and metabolism (reviewed in Manning et al., 2007, and Alessi et al., 1998).

Several genetic and pharmaceutical lines of investigation have implicated PI3K-PKB signaling as an important regulator of mES cell proliferation and self-renewal. Overexpression of a membrane targeted PKB results in Lif-independent self-renewal in mES cells without activating STAT3 (Watanabe et al., 2006). PTEN knockout mES cells show enhanced proliferation (Sun et al., 1999), whereas pharmaceutical inhibition of PI3K with LY294002 resulted in decreased self-renewal with a collateral increase in ppERK1/2 levels (Paling et al., 2004). The exact contribution of PI3K-PKB signaling to either mES cell proliferation or self-renewal is unclear as many experimental designs do not differentiate between both processes. Moreover, as mES cells differentiate there is a dramatic increase in doubling time (reviewed in White et al., 2005), therefore if proliferation is being measured by cell counting, an increase in differentiation could appear similar to a reduction in proliferation. This being said, there is evidence for the indirect regulation of mES cells proliferation by PI3K signaling through combined positive regulation of
CCND1 mRNA expression and inhibition of GSK3 mediated CCND3 protein degradation (Jirmanova et al. 2002). Interestingly, PDK1 knockout cells, where PKB is inactive, exhibit no apparent self-renewal defect (Williams et al., 2000, Stavridis et al., 2007), however, basal PI3K activity was doubled in the absence of PDK1 indicating that there may be divergences between PI3K and PKB activity in mES cells, or that there may be subtle differences in PKB isoform activity not identified by this study (Williams et al., 2000).

PI3K is also activated by the activity of a novel RAS family ERAS that selectively binds PI3K subunit p110, but not RAF (Takahashi et al., 2003). ERAS is robustly, and apparently exclusively, expressed in mES cells where its interaction with PI3K results in Ser473 phosphorylation of PKB. Eras depletion attenuates mES cell proliferation without any obvious effect on pluripotency, as judged by OCT4 expression. There is also evidence that PI3K signaling can be activated by hypoxia in mES cells, and this also results in enhanced proliferation and CCND1 expression (Lee et al., 2008). At present the majority of data indicated a primary role for PI3K activity in maintaining mES cell proliferation and more study is required to determine the involvement of this pathway in maintaining pluripotency.

1.5.2 ERK MAPK signaling

Several lines of evidence have shown that there is an inverse relationship between ERK1/2 activity and mES cells self-renewal (Burdon et al., 1999, Kunath et al., 2007, Stavridis et al., 2007, Ying et al., 2008) and that it is dispensable for proliferation (Jirmanova et al. 2002). ERK1/2 can be activated in mES cells by both FGF4 and LIF signaling pathways however certain quantitative or qualitative differences may exist between activation by either pathway as Fgf4 mutant mES cells still exhibit appreciable levels of ERK1/2 phosphorylation but are incapable of committing to neural lineages (Kunath et al., 2007).

Genetic studies into components of the MAPK pathway in mES cells have indicated that downstream of either FGFR and GP130 activation GRB2 and SHP2 are essential for mES cell lineage commitment (Cheng et al., 1998, Qu et al.. 1998),
with a slight reduction in proliferation observed in Shp2 deficient cells in the presence of high serum concentrations (15%) (Qu et al., 1998). GRB2 is a highly promiscuous adapter protein, at least in Hek293T cells (Bisson et al., 2011), however the observation that the differentiation defect was rescued by expression of oncogenic RAS indicates that activation of RAS/MAPK signaling is its primary function on mES cells (Cheng et al., 1998).

More recent work has implied an implicit role for ERK2 in this process (Kunath et al., 2007). Analysis of the differentiation potential of Erk2 knockout mES cells derived from Erk2<sup>+/−</sup> mice showed that they exhibited a severe defect in both neural and mesoderm lineage commitment. LIF withdrawal in Erk2 deficient cells leads to the upregulation of the epiblast marker Fgf5 however without a concomitant downregulation of the mES cell marker Rex1. Interestingly this phenomenon was completely reversible upon re-addition of Lif to the culture medium. Though there was no difference in the expression levels of ERK1 in both of the Erk2 deficient cell line, there also appeared to be no detectable compensatory ERK1 phosphorylation as seen in other cell types (Table 1.4). Moreover, the remaining ERK isozyme was dramatically less responsive to stimulation by both LIF and serum possibly explaining why ERK1 was not capable of compensating for loss of ERK2 in these cell lines and may explain the differences between other studies where E14 mES depleted for Erk2 by two rounds of gene targeting were capable of generating mesoderm tissue and exhibited compensatory phosphorylation of ERK1 (Yao et al., 2003). Corroborating evidence from Stravidis and colleagues however indicated that MEK-ERK activity was essential for neural specification and that only for a short period of activity was sufficient (~24h) (Stravidis et al., 2007). However, it also appears from this study that treatment of mES cells with the MEK inhibitor PD184352 results in enhanced phosphorylation of PKB on Ser473 indicating the chemical inhibition of MEK-ERK1/2 may result in collateral activation of parallel pathways. Furthermore, expression of a membrane targeted PKB dramatically enhanced ERK1/2 phosphorylation following LIF stimulation, in both magnitude and
duration (Watanabe et al., 2004), indicating crosstalk between both pathways is active in mES cells.

ERK1/2 activity has again been implicated in the specification of PE cells in mES culture (Canham et al., 2010). mES cells appear to heterogeneously express the endodermal marker Hex (Thomas et al., 1998). Canham and colleagues isolated Hex⁺ mES cells, that still expressed the pluripotency marker SSEA1, by targeting the endogenous Hex locus with a Venus reporter and then sorted cells based on the extent of Venus florescence. Western blot analysis of sorted cells revealed elevated levels of ppERK in Venus expressing cells, compared to the Venus negative population thus indicating that ppERK1/2 levels are highly variable from cell to cell in culture, and as Venus⁺ cells appeared primed for PE specification variability may correlate with lineage specification. Interestingly, for such a scenario to hold true ERK1/2 activity would have to be sustained for a considerable period to provide such discrete segregation of ppERK1/2 between Venus⁺ and Venus⁻ populations indicating that the duration of ERK1/2 activity may also be important in mES cells. The use of sensitive reporters of ERK1/2 activity such as FRET based biosensors (Zhang et al., 2007) will undoubtedly be highly informative in studying ERK1/2 dynamics in mES cells.

1.5.3 3i culture and the ‘Ground State’ hypothesis

Following on from the initial genetic evidence that implicated FGF-ERK was the primary auto-inductive cue that caused mES cells to commit to specific lineages in culture Ying and colleagues have developed a novel mES culture condition that allows the robust propagation of mES cells, in the absence of exogenous LIF, serum, or BMP4 (Ying et al., 2008). By combined inhibition of FGFR (PD173074), Mek (PD0325901) and Gsk3β (CHIR99021), termed ‘3i’, the authors showed that not only could STAT3 deficient mES cells be derived, but also that 3i culture supported the derivation of germ line-competent mES cell lines from CBA embryos (Ying et al., 2008), as well as from non-obese diabetic (NOD) mice (Nichols et al., 2009). Furthermore such culture methods have facilitated the
derivation of ES cell lines from rats, a phenomenon previously not possible (Li et al., 2008, Buehr et al., 2008).

From this data the authors propose a model that mES cells destabilize their self-renewal by autocrine FGF signaling and in standard culture conditions exogenous stimuli must be provided to act downstream of the FGF signal to maintain pluripotency. However, if this autoinductive signal is blocked, by pharmaceutical inhibition of the FGF pathway, mES cell default to their innately programmed state of self-renewal. Although both FGFR and MEK inhibition greatly enhance the self-renewal of mES cells, combined inhibition of GSK3β is required to maintain growth and survival (Ying et al., 2008).

This being said, the generality of such a ground state is still unknown. Current hES cell lines require FGF2 activation to proliferate (Thompson et al., 1998) and no hES cell line have been reported to have been derived in 3i culture. This may be because current hES cell lines appear to more closely resemble the pluripotent cell derived from ~E5.5 epiblasts (Brons et al., 2007; Tesar et al., 2007) whose self-renewal is also dependent on Fgf signaling. However pre-X inactivated hES cell lines have been recently described when derived under hypoxic conditions indicating that the derivation of hES cells with mES-like qualities may be possible given the correct culture conditions (Lengner et al., 2010).

Whether the 3i culture phenomenon is restricted to pluripotent cells of rodent blastocysts or if a similar ground state for other mammalian ES cells exist is unsure, however it does provide a strong proof of principal. A current challenge facing the field is to understand the mechanism by which FGF signaling acts to destabilize mES cell self-renewal. Although it is now possible to derive and maintain rodent ES cells under defined conditions, the issue of inefficient and heterogeneous differentiation is still a major experimental hurdle. Therefore a mechanistic understanding of FGF-ERK signaling in mES cells is essential for the rational control of mES cell lineage specification.
1.6 Technical notes

1.6.1 Genome engineering

The ability to manipulate the murine genome with speed and precision has proven to be one of the most profound technical advances effecting many areas of experimental biology. If such a genetic modification is introduced into mES cells it can be passed to the germ line of chimeric mice (Thomas et al., 1987, Robertson et al., 1986) or if aggregated with 4n embryos can continue to generate every tissue of the adult mouse (Nagy et al., 1993). Genome targeting strategies can vary dramatically depending on the needs of the experiment but can be loosely divided into two categories: targeted or non-targeted. A targeted mutation relies on the natural cellular process of homologous recombination, where sequences are exchanged between two highly similar molecules of DNA following a double strand break. A non-targeted mutation relies on non-homologous recombination and thus can occur at many random permissive sites throughout the genome. Cells that have undergone the desired homologous or non-homologous event are enriched in the population by the incorporation of a selectable marker and can be analysed clonally.

A targeting vector is constructed in bacteria by either conventional cloning or recombineering and usually contains: a stretch of DNA homologous to the genomic location being targeted into which the desired mutation will be introduced: a mammalian selectable marker, such as a drug resistance gene, flanked by the homology regions: a negative selection marker such as diphtheria toxin A (DTA), outside of the homology region: as well as a means to select and propagate the plasmid in bacteria. Once the DNA molecule is modified in vitro it is linearized and introduced into ES cells. Stable integrants are selected for and propagated clonally and homologous versus non-homologous events can be determined by various PCR based techniques and Southern blotting/restriction fragment length polymorphism analysis. Several factors such as cell cycle duration (Udy et al., 1997), or chromosomal location (Yanez et al., 2002) can affect the targeting efficiency,
however in general the targeting frequency for any specific locus is determined empirically.

By far the most popular form of non-homologous genome engineering is gene-trap mutagenesis. Gene-trap mutagenesis generally provides a loss of function mutation as well as reporting the expression of a specific gene and is therefore an invaluable and versatile tool. Gene-trap mutagenesis involves the random genomic insertion of a selectable marker downstream of a splice acceptor sequence (Gossler et al., 1989). If this insertion event occurs within an expressed gene, usually an intron, the endogenous transcript will be trapped by the exogenous transcript and a fusion message will be produced truncating the endogenous transcript at the point of insertion. The trapping event can then be detected by 5’ rapid amplification of cDNA ends (5’RACE) analysis, or the insertion location identified by splinkerette PCR. A major caveat with such loss of function approaches is that low-frequency alternate splicing around the inserted gene-trap cassette can occur resulting in a hypomorphic allele (McClive et al., 1998). Additionally, the trapped gene must be expressed at a sufficiently high level to achieve drug resistance, however novel strategies have been developed to overcome this issue (Salminen et al., 1998, Tsakiridis et al., 2009).

1.6.2 Mass spectrometry based approaches to determine kinase interactions

Currently, identification of any simple to complex mixture of proteins will generally involve some form of mass spectrometry (MS) analysis. To this end, in the ~20 years since the implementation of MS technology in proteomics a wide range of platforms and data acquisition/analysis tools have been developed ( Aebersold et al., 2003). Figure 5.1.1 shows a typical LC-MS/MS work flow where by tryptically digested peptides are fractionated by micro-liquid chromatography and eluted in order of their hydrophobicity across an increasing organic gradient. Eluted peptides are then fed through a needle where at the tip they are vapourized and converted to ions by strong electric potential and fed into the mass analyzer through a transfer capillary from where they are manipulated by electric fields. Once within the mass
analyzer the \( m/z \) value for each peptide fragment is determined. There are multiple types of mass analyzers that differ primarily by how they control the movement of ions through the detector, however there are differences in resolution between setups. The three main types used are:

- **Quadrupole mass analyzers** where ions are passed through four electrically charged, alternately connected rods. Only ions of a certain \( m/z \) value will pass through for a given range of voltages, the trajectories of other peptides will be unstable and not reach the detector, therefore varying the voltage allows the peptides to be scanned across a range of \( m/z \) values.

- **Time of flight mass analyzers** where the \( m/z \) value is calculated as a function of the time it takes for a specific ion to pass through the flight tube to the detector.

- **Quadrupole ion traps** trap ions in both radially and axially in a static electric field using a set of quadrupole rods. Ions can then be selectively ejected to the detector based on their \( m/z \) value.

This sort of analysis generates a mass spectra from which proteins can be identified in a process known as peptide mass fingerprinting, however it is heavily reliant on the homogeneity of the analyte. For complex mixtures of peptides derived from multiple proteins tandem MS analysis is performed: a particular peptide ion is isolated, energy is imparted by collisions with an inert gas (such as nitrogen molecules, or argon or helium atoms), and this energy causes the peptide to break apart. A mass spectrum of the resulting fragments is then generated which allows determination of the primary structure of a peptide, as well as a range of possible post translational modifications by submitting the mass spectra to a database, such as Mascot, consisting of \( m/z \) values for every possible tryptically derived peptide from an organism. A commonly accepted threshold is that an event is significant if it would be expected to occur at random with a frequency of less than 5\%, this equates to a Mascot score of 20.

There are several theoretical and technical considerations that impact the quality of data derived from MS/MS experiments (reviewed in Steen et al., 2004). One such issue is that of random sampling of parent ions selected for sequencing.
during the first round of MS analysis. This can result in a decrease in reproducibility between runs especially in identifying low abundance ions, and can make quantification of protein abundance between runs difficult. The arrival of tools for relative quantification have helped to resolve this issue to some extent, and also afforded the biologist more experimental freedom when designing an MS-based project. Relative quantification between two biological samples is usually achieved by differentially labeling proteins/peptides of each sample before analysis, then mixing both samples and analyzing them simultaneously by tandem MS. Although there are currently several methods to label peptides post purification, such as ICAT (Gigy et al., 1999) and iTRAQ (Ross et al., 2004), metabolic labeling as became a popular choice when conducting studies from in vitro cultured cells. Stable isotope labeling by amino acids in cell culture (SILAC) (Ong et al., 2002) takes advantage of the fact that cells can metabolize amino acids (usually arginine or lysine as all tryptically digested peptides will contain at least one or more of each amino acid) labelled with ‘heavy’ $^{13}$C and $^{15}$N isotopes. Using a control/reference sample grown in normal media, relative quantification is achieved usually by calculating the peak intensity ratios between isotopically labelled and non-labeled peptides using software packages such as MaxQuant.

![Diagram of LC-MS/MS workflow](image)

**Figure 1.12: A typical LC-MS/MS workflow.** Protein samples are separated in one or two dimensions and bands/spots of interest are excised and digested in-gel. Tryptic peptides are fractionated by nano-HPLC and
subsequently vapourized and ionized before entering the mass spectrometer. Image adapted from Steen et al. (2004).

1.6.2.1 IMAC and phosphopeptide enrichment

A major obstacle in the study of phosphorylated proteins is that they comprise only a small fraction of the total protein in a cellular lysate. The compound issue of the substoichiometric enzyme:substrate interactions has led to the development of various techniques to selectively enrich phosphopeptides allowing the phosphoproteome to be probed to greater depths by conventional LC-MS/MS analysis/detection.

The affinity of phosphate groups for various cations such as Zn$^{2+}$, Fe$^{2+}$, Ga$^{3+}$ has been exploited in immobilized metal affinity chromatography (IMAC) (Andersson et al., 1986, Posewitz et al., 1999). IMAC protocols involve sample preparation/solubilization, usually in denaturing conditions such as 2M Urea, followed by in-solution tryptic digestion. Samples are then subjected to phosphopeptide enrichment and the bound fraction is eluted and analysed by a suitable MS method. The major limitations of this approach are that peptides that contain many acidic amino acids such as cystine or histidine tend to bind and co-elute with phosphopeptides. Moreover, the choice of metal ion used for the enrichment step can bias the peptides that are recovered. Several pre-processing steps have been developed, such as esterification of acid amino acid side chains (Ficarro et al., 2002), however such steps can add additional bias during sample preparation. Recently the use of TiO$_2$ as a phosphopeptide chelator has been described (Larsen et al., 2005). In parallel experiments TiO$_2$ outperformed IMAC in enrichment of phosphopeptides from a relatively simple protein mixture, however the issue of highly abundant phosphopeptides from complex cellular lysates saturating the enrichment step as well as dominating the MS run remains an issue.

2-dimensional differential gel electrophoresis (2D-DiGE) is a protein labeling and separation technique allowing quantitative analysis of two or more samples by optical fluorescence detection of differentially labelled proteins that are
electrophoretically separated on the same gel (Unla et al., 1997). As 2D-DiGE allows multiple samples to be ran on the same gel it minimizes the gel-to-gel variation that can hinder the reproducibility and comparability of traditional 2D electrophoresis experiments. Crude lysates are made under denaturing conditions and then minimally labelled (<5%) with a fluorescent dye, usually Cy3/Cy5, although experiments using Alexa-fluor dyes have been described (Von Eggeling et al., 2001). Differentially labelled samples such as ‘with treatment’ and ‘without treatment’ are then combined, separated by mass in one dimension followed by isoelectric focusing in the second dimension. The gel can then be scanned and analysed and differentially labelled spots excised, in-gel digested, and identified by MS/MS analysis. By combining pre-separation techniques such as 2D-DiGE with phosphopeptide enrichment, Kosako et al (2009) recently identified a plethora of known and novel Erk targets with subsequent analysis showing 13 of the 14 candidates tested to be bona fide substrates \textit{in vitro}.

1.6.2.2 IP-LC-MS/MS

By far the most popular method to identify protein:protein interactions is by immunopurification of bait protein complexes, followed by analysis by LC-MS/MS analysis (reviewed in Chen et al., 2007). Generally, this approach consists of isolating a bait protein using affinity purification techniques, such as immunoprecipitation of endogenous protein with a bait specific antibody, immunoprecipitation of exogenous tagged protein with a tag specific antibody (e.g. FLAG or HA), or pull-down of exogenously expressed tagged protein (e.g. Bio-Avi tag). Protein samples identified in such a way are separated and visualized by conventional SDS-PAGE/Coomassie analysis, specific protein bands of interest (or the entire gel) are then excised and subjected to in-gel proteolytic digestion (usually with trypsin), followed by analysis by a suitable MS based approach.

The initial, and most crucial parameter in establishing any IP/AP-LC-MS/MS experiment is the choice of bait purification method. If possible, though it is not always so, the preferred method of protein purification is by immunoprecipitation
with an antibody that specifically recognizes the endogenous protein of interest. Obviously the quality and specificity of the antibody are of primary concern, however this approach has one main advantage in that it helps to maintain a more physiological relevant experimental design. One of the most significant drawbacks to using an antibody directed against the endogenously expressed protein is that the most robust negative control for such an experiment is a cell line lacking the bait protein, which in most cases is not readily available (reviewed in Markham et al., 2007). A common solution to this problem is to use an isotype-matched control antibody that does not recognize the protein of interest. However as every antibody recognizes a specific profile of non-specific epitopes such ‘mock IP’ approaches can result in the false identification, or exclusion, of potential interacting proteins due to antibody based artifacts.

Exogenously expressed epitope tagged bait proteins help to obviate this issue as it allows the same antibody to be used for both positive and negative controls, the only difference between the two samples being the presence or absence of the bait protein. Concerns regarding this method are the possibility of the epitope tag affecting the proteins function, localization, or preferential exclusion from macromolecular complexes in the presence of endogenous un-tagged protein. For these reasons it is necessary to gauge the physiological function of the exogenously expressed protein. The correct co-localization of endogenous and exogenous protein is a good indicator that the tagged protein functions normally. Also, if there is a readout of bait activity, such as a gene known to be regulated by it, differences in gene expression levels between samples with and without the tagged protein can also be used as an indicator of tagged protein function. An ever-present concern as to the validity of data derived from overexpression studies is that of artifactual, or physiological irrelevant interactions due to increased bait levels. However, as there it yet no study to suggest that this is a generalizable phenomenon, and as recent reports indicate that the protein abundance can vary from cell to cell (Sigal et al., 2006), it is probable that moderate overexpression will have minimal physiological disruption for most proteins, although there are always exceptions (Niwa et al., 2000), and
therefore the significance of non-endogenous levels must be determined on a protein by protein basis.

Another important aspect of the experiment design is the choice of protein purification conditions. For co-IP experiments cells are usually lysed in a ‘mild’ non-denaturing buffer, such as RIPA, that contain non-ionic detergents at concentrations around 1%, with physiological salt concentrations, usually a combination of sodium and potassium salts in the millimolar range (100-200mM). Both detergent composition and salt concentrations can affect the efficiency of protein solubilisation and thus bias the outcome of any IP-MS experiment. There is generally a trade-off between the efficiency of protein solubilisation and the maintenance of some protein:protein interactions, however this must be determined empirically for each protein. Having a knowledge of the subcellular location of the protein of interest can help to determine the initial lysis conditions, i.e. if it is membrane located then the addition of zwitterionic detergents can help with protein solubilisation, or determine whether subcellular fractionation may be beneficial to the experimental design (van den Berg et al., 2010).

The choice of support matrix for antibody:bait capture can also be of importance, however this tends to be a more generalizable issue of bead based background. For most IP protocols protein A/G coated agarose beads are used to precipitate immune complexes from solution. Whenever possible it is useful to use antibodies that are covalently cross-linked to the support matrix, therefore eliminating IgG contamination of the eluate that may interfere with downstream MS analysis. Pre-clearing the lysate before immunoprecipitation with pre-immune beads has been reported to be useful in reducing non-specific binding, however as many common background proteins such as proteosomal and ribosomal subunits are highly abundant proteins, the portion depleted by pre-clearing may be minimal (Chen et al., 2007).

Though the choice of antibody can largely determine the elution strategy, it can be best to determine the optimal conditions empirically. Where possible, elution by peptide competition can be desirable as this should specifically elute antibody
bound proteins. Conversely, elution under acidic – glycine or phosphoric acid, or basic – ammonia, conditions are possible. As this step represents a crucial point of protein loss it is important to ensure maximal protein recovery, whilst attempting to preferentially elute bait-bound complexes. The use of affinity tags with proteolytic cleavage sites, such as for TEV protease, have been used, although this results in considerable contamination of the eluate with the protease that can also affect downstream MS analysis. The use of photocleavable affinity tags such as photocleavable biotin (Lim et al., 2008) can help to obviate this issue.
1.7 Aims of this thesis

The aim of my thesis is to functionally and biochemically characterize ERK2 in mES cells. Both pharmaceutical and genetic evidence have identified the FGF-MEK-ERK pathway as an important determinant of cell fate decisions in mES cells however there is little mechanistic data as to how this is achieved. This is of particular importance as the early biochemical steps in lineage segregation are currently poorly understood and therefore difficult to predict and control.

Previous work from our lab showing a degree of phenotypic overlap between FGF4 and ERK2 knockout mES cell lines indicated that a focused study on ERK2 function may give valuable information as to how the decision to differentiate is made in mES cells, and which downstream effectors are targeted in this process. Unlike the LIF-JAK/STAT pathway, our current understanding of how the FGF4 pathway integrates with the core pluripotency machinery is limited to correlative data analysed over the timeframe of days. To this end I employed an IP-MS/MS approach to identify ERK2 binding partners isolated from FLAG-ERK2 IPs. However, in establishing the initial conditions for this study I discovered that the differentiation defect exhibited by Erk2 knockout cells could not be attributed to the loss of ERK2 protein, and that this differentiation defect was context dependent.

Therefore, to readdress the function of ERK2 in mES cells, and to generate a cell line suitable for MS and functional studies, I disrupted the Erk2 allele in E14 mES cells by two rounds of homologous recombination. This allowed the functional consequences of ERK2 depletion to be analysed on the level of ERK1 activation, ERK1/2 substrate phosphorylation and transcriptional regulation, as well the contribution of this isozyme to the pluripotent state.

Because previous embryo-derived Erk2 mutant cells were unipotent when cultured under monolayer conditions they did not represent a suitable model to study factors that might influence the pluripotent state. However, Erk2 deficient cells generated for this study retained their pluripotency and were therefore used to identify ERK2 binding partners. Results from this study were further perused and the
effect a novel small molecule inhibitor to p90RSK exerted on mES cell self-renewal and the FGF-ERK signaling cascade was determined.

It is hoped that this study will go in some way to clarifying the specific role of ERK2 in mES cell biology and provide a platform for further analysis.
Chapter 2 - Materials and Methods

2.1: Materials

All chemicals were purchased from either Sigma or Fisher Scientific. All restriction enzymes were from New England Biolabs (NEB). Electrophoresis grade Agarose was from Invitrogen. Synthetic oligonucleotides were synthesized by either Sigma or Integrated DNA Technologies, with either HPLC or PAGE purification for all sequences over 60bp. DNA sequences over 100bp were synthesized by GeneArt. Protease inhibitors were from Roche (Complete EDTA Free), and phosphatase inhibitors were from Calbiochem (Phospho Set III). Gibco distilled RNase/DNase-free water was used for DNA/RNA suspension. Milli-Q filtered and deionized water (>18 megohm-cm resistance) was used for making all stock solutions. Sterile PBS (-MgCl₂, -CaCl₂) was from Gibco. Polypropylene microfuge tubes were from Sarstedt. Low protein-binding microfuge tubes were from Alpha Laboratories. Pipette tips were from Greiner Bio One. Both bottom agar and lysogeny broth were from Difco (LB Agar, Miller/LB Broth, Miller), and were provided as an in-house resource.

2.1.2: Kits and Reagents

2.1.2.1: Kits
Qiagen- QIAprep™ Spin Miniprep Kit
Qiagen- Plasmid Maxi Kit
Qiagen- RNeasy™ Mini Kit
Qiagen- Qiashredder™ Columns
Qiagen- DNeasy™ Blood And Tissue Kit
Zymo Research- DNA Clean and Concentrator™
Zymo Research- Zymoclean™ Gel DNA Recovery Kit
Millipore- Amicon Ultra-0.5 ml Centrifugal Filters for Protein Purification and Concentration
Amersham- MegaPrime™ DNA Labeling System
Amersham- Illustra ProbeQuant™ G-50 Micro Columns
Sigma- Leukocyte Alkaline Phosphatase Assay Kit
Invitrogen- SilverQuest™ Silver Staining Kit
Invitrogen- Zero Blunt® TOPO® PCR Cloning Kit

2.1.2.2: Reagents
Qiagen- Taq DNA Polymerase
Qiagen- Tri Reagent
Finnzymes- Phusion® Taq
Finnzymes- Calf Intestinal Phosphatase
Finnzymes- 1Kb DNA Size Standard
Invitrogen- SuperScript® III Reverse Transcriptase
Invitrogen-M-MLV Reverse Transcriptase
NEB- Vent DNA Polymerase
NEB- Klenow Fragment
NEB- T4 DNA Polymerase
NEB- T4 DNA Ligase
NEB- Calyculin A
Roche- LightCycler® 480 Probes Master Mix
Roche- UPL Set, Mouse
Sigma- Benzonase® Nuclease
Sigma- PerfectHyb™ Plus Hybridization Buffer
Sigma- GMEM
Invitrogen- NuPAGE® Novex® Bis-Tris Gel (4-12% gradient gel)
Invitrogen- MES SDS Running Buffer
Invitrogen- LDS Sample Loading Buffer
Invitrogen- SimplyBlue™ Safe Stain (coomassie)
Invitrogen- SeeBlue™ Prestained Protein Standard
2.1.2: Solutions

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>137mM NaCl, 2.7mM KCl, 4.3 mM Na$_2$HPO$_4$, 1.47mM KH$_2$PO$_4$</td>
</tr>
<tr>
<td>TAE</td>
<td>40mM Tris Base, 0.114% Acetic Acid, 1mM EDTA pH 8.0</td>
</tr>
<tr>
<td>‘Ear’ Buffer</td>
<td>25mM NaOH, 0.2mM EDTA</td>
</tr>
<tr>
<td>Sarcosyl Lysis Buffer</td>
<td>10mM Tris-HCl pH 7.5, 10mM EDTA, 10mM NaCl, 0.5% Sarcosyl, 1mg/ml Proteinase K (added fresh)</td>
</tr>
<tr>
<td>2x Laemili's Lysis Buffer</td>
<td>120mM Tris pH 6.8, 4% SDS, 20% Glycerol</td>
</tr>
<tr>
<td>Western Transfer Buffer</td>
<td>12mM Tris-pH 7.5, 96mM Glycine, 20% MeOH</td>
</tr>
<tr>
<td>Western Blocking Buffer</td>
<td>10mM Tris-HCL pH 7.5, 150mM NaCl, 0.1% Tween 20 5-10% Skimmed Milk Powder or 5% BSA (Fraction V)</td>
</tr>
<tr>
<td>Western Wash Buffer</td>
<td>10mM Tris-HCL pH 7.5, 150mM NaCl, 0.1% Tween 20</td>
</tr>
<tr>
<td>Western Stripping Buffer</td>
<td>72.5 mM Tris pH 7.5, 2% SDS, 125mM β-Mercaptoethanol</td>
</tr>
<tr>
<td>IHC Blocking Buffer</td>
<td>137mM NaCl, 2.7mM KCl, 4.3 mM Na$_2$HPO$_4$, 1.47mM KH$_2$PO$_4$, 0.5% Triton X-100, 2% Species specific serum</td>
</tr>
<tr>
<td>CCMB80 Salt Buffer</td>
<td>10mM KOAc pH 7, 80mM CaCl$_2$, 20mM MnCl$_2$, 10mM MgCl$_2$, 10% Glycerol, pH 6.4</td>
</tr>
<tr>
<td>20x SSC</td>
<td>3M NaCl, 340mM Tri-Sodium Citrate, pH 7</td>
</tr>
<tr>
<td>Denaturing Solution</td>
<td>500mM NaOH, 1.5M NaCl</td>
</tr>
<tr>
<td>Neutralisation Solution</td>
<td>500mM Tris, 1.5mM NaCl, pH 5.5</td>
</tr>
<tr>
<td>SSC Wash</td>
<td>0.5-2x SSC, 0.1% SDS, 0.5% Sodium Pyrophosphate</td>
</tr>
</tbody>
</table>
**IP Lysis Buffer**: 20mM Hepes pH 7.6, 100mM KCl, 2.5mM MgCl₂, 10% Glycerol, 0.02% NP40, 0.5mM DTT, 1x Complete Protease Inhibitor Cocktail, 1x Phospho Set III Inhibitor Cocktail.

**IP Wash Buffer**: 50mM Tris-HCL pH 7.5, 100mM NaCl, 0.1% NP40, 10% Glycerol, 0.5mM DTT, 1x Complete Protease Inhibitor Cocktail, 1x Phospho Set III Inhibitor Cocktail.

**IP Elution Buffer (Basic)**: 500mM NH₄OH, 0.5mM EDTA, pH 11.5

**IP Elution Buffer (Acidic)**: 50mM H₃PO₄ (pH1.8)

**IP Elution Buffer (Glycine)**: 200mM Glycine, 0.5M NaCl, 0.1% NP40 (pH 3.5)

**IP Elution Buffer (Peptide)**: 0.2 mg/mL FLAG peptide, 20mM HEPES, 10% Glycerol, 100mM KCl, 1.5M MgCl₂, 0.2mM EDTA (pH 7.4)

**N2 Supplement**: 25µg/ml Insulin, 100µg/ml Transferrin, 6ng/ml Progesterone, 16µg/ml Putrescine, 30nM Se₂Cl₂,

**6x Sucrose DNA Loading Buffer**: 0.25% Bromophenol Blue, 40% (w/v) Sucrose

**Citrate Solution**: 18mM Citric Acid, 9mM Sodium Citrate, 12mM NaCl

---

### 2.1.3: Plasmids

**pBluescript II KS⁺**: Plasmid containing a bacterial LacZ expression cassette with an intragenic multiple cloning site for blue/white screening, ColE1 and F1 origins of replication. Ampicillin resistant.
<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZeroBlunt TOPO</td>
<td>Plasmid containing both LacZ and ccdB expression cassettes for blue/white screening and negative expression, respectively, CoIE1 and F1 origins of replication. Kanamycin resistant.</td>
</tr>
<tr>
<td>pPy CAG IP (pIPC37)</td>
<td>Mammalian expression plasmid containing a CAG promoter followed by an IRES puromycin acetyltransferase selection cassette. SV40 and CoIE1 origins of replication. Ampicillin resistant.</td>
</tr>
<tr>
<td>pPy CAG IH</td>
<td>Mammalian expression plasmid containing a CAG promoter followed by an IRES hygromycin phosphotransferase selection cassette. SV40 and CoIE1 origins of replication. Ampicillin resistant.</td>
</tr>
<tr>
<td>pPy CAG IB</td>
<td>Mammalian expression plasmid containing a CAG promoter followed by an IRES blasticidin-S selection cassette. SV40 and CoIE1 origins of replication. Ampicillin resistant.</td>
</tr>
<tr>
<td>pL451</td>
<td>FRT flanked neomycin acetyltransferase selection cassette, driven by both mPGK and Em7 promoters for selection in mammalian and bacterial cells, respectively. CoIE1 and F1 origins of replication. Ampicillin resistant.</td>
</tr>
<tr>
<td>AGS 703</td>
<td>LoxP flanked hygromycin phosphotransferase-thymidine kinase selection cassette, driven by a CMV promoter. CoIE1 and F1 origins of replication. Ampicillin resistant.</td>
</tr>
<tr>
<td>AGS 684</td>
<td>Mammalian expression plasmid containing a CAG promoter, driving the expression of eGFP, followed by an IRES puromycin acetyltransferase selection cassette. SV40 and CoIE1 origins of replication. Ampicillin resistant.</td>
</tr>
</tbody>
</table>
AGS 627  Mammalian expression plasmid (pcDNA3.1) containing a CMV promoter driving the expression of the hLIF05 gene. Neomycin acetyltransferase expression is driven by a SV40 promoter. pBR322 origin of replication. Ampicillin resistant.

pSM2c  Retroviral plasmid for expression in mammalian cells containing a hU6 promoter driving expression of the human mir30 microRNA sequence and a PGK promoter driving expression of a puromycin acetyltransferase cassette. R6Kg origin of replication. Kanamycin and chloramphenicol resistant.

bMQ209m09  Bacterial artificial chromosome (129s7/AB2.2 BAC clone) containing a 200kb insert containing to the Erk2 allele.

2.1.4: Primers/Synthetic oligonucleotides

-Cloning

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BamHI Erk2 F</td>
<td>GATCGCTAGCGCGCCGGCCGGCGGCCTTCTGATCCGGAGATGG</td>
<td>Forward primer for amplifying Erk2</td>
</tr>
<tr>
<td>Erk2 Notl R</td>
<td>GCGGCGCCCTTAAGATCTGTATCTCTTGG</td>
<td>Reverse primer for amplifying Erk2</td>
</tr>
<tr>
<td>BE1F</td>
<td>CGACAAGGGATCCGGTGCGCCGGCGCGGGCCGGCCGGCCGGCCGGCCGGCCGGCCGGCCGGCCGGCCGG</td>
<td>Forward primer for amplifying Erk1</td>
</tr>
<tr>
<td>NE1R</td>
<td>TTATCGAGCGCGCGCCCTTAGGGGCCCTCTCTGCGCCGCCCTCTCTGCGCCGCCCTCTCTGCGCCGCC</td>
<td>Reverse primer for amplifying Erk1</td>
</tr>
<tr>
<td>E3PF</td>
<td>GTTTGCAAGTTCCAGTTGGGCC</td>
<td>Forward primer for amplifying 3' Erk2 southern probe</td>
</tr>
<tr>
<td>E3PR</td>
<td>CCCTGTGTATTTGTAGGCCAGCGTTTCTAC</td>
<td>Reverse primer for amplifying 3' Erk2 southern probe</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>K52R-F</td>
<td>GTTGCTATCAGGAAAATCAGTCTTTTTTGAGC</td>
<td>Forward primer for K52R mutation</td>
</tr>
<tr>
<td>K52R-R</td>
<td>CAAAAGGACTGATTTTCTGATAGCAACTCGA</td>
<td>Reverse primer for K52R mutation</td>
</tr>
<tr>
<td>Y185F-F</td>
<td>GACAGAGTTTCGTAGCCACACGTTGGTACAGA</td>
<td>Forward primer for Y185F mutation</td>
</tr>
<tr>
<td>Y185F-R</td>
<td>GGCTACGAACCTCTGTCGAAGAACCCTGTG</td>
<td>Reverse primer for Y185F mutation</td>
</tr>
<tr>
<td>Y261A-F</td>
<td>AGCTAAGAAACGCATTGCTTTTCTCTCCGCAC</td>
<td>Forward primer for Y261A mutation</td>
</tr>
<tr>
<td>Y261A-R</td>
<td>AGAAAGCAATGCCTTTCTAGCTTTTTAAATTTATTATAC</td>
<td>Reverse primer for Y261A mutation</td>
</tr>
<tr>
<td>AEF-F</td>
<td>CACAGGTTTCTTGGCAGAGTTTCGTAGCCACA</td>
<td>Forward internal primer for T183A, Y185F mutation</td>
</tr>
<tr>
<td>AEF-R</td>
<td>GTGTTGGCTACGAAACTCTGCAAGAACCCTGTGTGAC</td>
<td>Reverse internal primer for T183A, Y185F mutation</td>
</tr>
<tr>
<td>EVP-F</td>
<td>GGATCCGGTGGAGGTGGAGGTGGAATGTCAACTTTAGAAGT</td>
<td>Forward internal primer for Erk2-Vp16</td>
</tr>
<tr>
<td>EVP-R</td>
<td>TATAGTTTACGCGGCGCATTACCGGACCCTG</td>
<td>Reverse internal primer for Erk2-Vp16</td>
</tr>
<tr>
<td>EEN-F</td>
<td>GGATCCGGTGGAGGTGGAGGTGGAATGTCAACTTTAGAAGT</td>
<td>Forward internal primer for Erk2-En</td>
</tr>
<tr>
<td>EEN-R</td>
<td>GCTTATCGAGCAGGCCGGCGCATTATCGGAGGCGGCGGAGGCGGCAATGATG</td>
<td>Reverse internal primer for Erk2-En</td>
</tr>
<tr>
<td>XKE-F</td>
<td>ATGTACCCCTCGAGCAGGCCACCATGGCAGCTGC</td>
<td>Forward primer to amplify both Erk2-Vp16 and Erk2-En</td>
</tr>
<tr>
<td>ENN-R</td>
<td>TTATCGAGCAGGCCCGCTTTATCGAGGCGGCGGAGGCGGCAATG</td>
<td>Reverse external primer to amplify Erk2-En</td>
</tr>
<tr>
<td>EVPN-R</td>
<td>TTATCGAGCAGGCCCGCTTTACGGACCCCGGAAATGCAATG</td>
<td>Reverse external primer to amplify Erk2-Vp15</td>
</tr>
<tr>
<td>BSRFF</td>
<td>GGATCCGGTGGAGGTGGAGGTGGAATTACCGA</td>
<td>Forward primer for amplifying Srf</td>
</tr>
<tr>
<td>NSRFR</td>
<td>TTACCGCGCCGCTACGCTCTTTTGCTGTG</td>
<td>Reverse primer for amplifying Srf</td>
</tr>
<tr>
<td>DNOTF</td>
<td>AGGCCGGCGCGCGAGCGCCCGCGGGGGGCTACG</td>
<td>Forward primer for Δ NotI mutation</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DNOTF</td>
<td>AGCCCCCGGCGGCTGCCGCCGCCCTCAGGCCACCC</td>
<td>Reverse primer for Δ NotsI mutation</td>
</tr>
<tr>
<td>RVSGL</td>
<td>TTGATATCGAGGGTGAGGTGGAGAAGACTAAAGGACGACGACAAG</td>
<td>Forward primer for amplifying Srf to clone into the FKBP expression plasmid</td>
</tr>
<tr>
<td>E6GGF</td>
<td>CCAGGATACAGATCTGGATCCGCCGCCGCCGCTGGAGGAGGAGG</td>
<td>Forward internal primer for Erk2-eGFP</td>
</tr>
<tr>
<td>E6GGR</td>
<td>GCTCACCATGGCCGCCGCGCCGCGCCGCTGGATCCAGATCTGGG</td>
<td>Reverse internal primer for Erk2-eGFP</td>
</tr>
<tr>
<td>NGFPR</td>
<td>AGCGGCGCTTATTACTTGTTCAGCTCGTCCATGCGAGAG</td>
<td>Reverse external primer to amplify Erk2-GFP</td>
</tr>
<tr>
<td>XKFPF</td>
<td>AGAATCCCTCGAGCCGCCACCATTAGGAGAGAGCACCGAGACG</td>
<td>Forward primer to amplify both PDK1 containing a XhoI site, Kozac sequence, and a single FLAG sequence</td>
</tr>
<tr>
<td>NPDKR</td>
<td>ATCGAGGCACCCTCTCACTCGACAGCAGCACTGGGATTGCTCTGGTACTG</td>
<td>Reverse primer for amplifying PDK1</td>
</tr>
<tr>
<td>PKGF</td>
<td>AATATGCTATTTGGAGGGGAAACGAGGACGACGACAATGCAG</td>
<td>Forward primer for PDK1 K114G mutation</td>
</tr>
<tr>
<td>PKGR</td>
<td>CTCCAGAATTCCATAGCATATTCTTGTCTCG</td>
<td>Reverse primer for PDK1 K114G mutation</td>
</tr>
</tbody>
</table>

**-Sequencing**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T3</td>
<td>AATTAACCCTCCTACTAAAGGGAAA</td>
<td>Primer against T3 promoter</td>
</tr>
<tr>
<td>T7</td>
<td>GTAATACTGACTCTATATAG</td>
<td>Primer against T7 promoter</td>
</tr>
<tr>
<td>IC37</td>
<td>CAGCTCCTGGGCAACGGTGCTGG</td>
<td>Forward primer for pCAG promoter</td>
</tr>
<tr>
<td>M13F</td>
<td>AGCGGATAAAATTTCCACACAGG</td>
<td>Forward primer for LacZ gene</td>
</tr>
<tr>
<td>M13R</td>
<td>GTAAACGACAGGCCAGT</td>
<td>Reverse primer for LacZ gene</td>
</tr>
<tr>
<td>Erk2 Internal F</td>
<td>CGTAGCCACAGGTGAGCAGCATCTC</td>
<td>Forward primer for Erk2 coding region</td>
</tr>
<tr>
<td>Erk2 Internal R</td>
<td>CCAAGTCCATGATTTGGTCTCTCGG</td>
<td>Reverse primer for Erk2 coding region</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>PDK1 Internal F</td>
<td>AGATGTGTGATGTACGCCTGG</td>
<td>Forward primer for PDK1 coding region</td>
</tr>
<tr>
<td>PDK1 Internal R</td>
<td>GAGGTGTCTGTGGTGCAAA</td>
<td>Reverse primer for PDK1 coding region</td>
</tr>
</tbody>
</table>

**-Q-PCR**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’→3’</th>
<th>Description</th>
<th>UPL Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGUPLF 97</td>
<td>CTAACGTTCTGCACCCGTGAC</td>
<td>Erk2 exon 3 genotyping primer (forward)</td>
<td>97</td>
</tr>
<tr>
<td>EGUPLR 97</td>
<td>GATCACAAGTGTGTTCCAGAGCAG</td>
<td>Erk2 exon 3 genotyping primer (reverse)</td>
<td>97</td>
</tr>
<tr>
<td>Spry1F</td>
<td>TCCACCTCGAGGTTCTG</td>
<td>Spry1 intron 1 genotyping primer (forward)</td>
<td>78</td>
</tr>
<tr>
<td>Spry1R</td>
<td>GGCATGCAGTATATCTCAGC</td>
<td>Spry1 intron 1 genotyping primer (reverse)</td>
<td>78</td>
</tr>
<tr>
<td>Oct4F</td>
<td>GTTGGAGAAGGTGGAACCAA</td>
<td>Oct4 intron spanning (forward)</td>
<td>95</td>
</tr>
<tr>
<td>Oct4R</td>
<td>CTCTTCTGCAGGGGCTTTC</td>
<td>Oct4 intron spanning (reverse)</td>
<td>95</td>
</tr>
<tr>
<td>CCND2F</td>
<td>CACTACCAGTTCCCACCTCAG</td>
<td>CCND2 intron spanning (forward)</td>
<td>45</td>
</tr>
<tr>
<td>CCND2R</td>
<td>CTGTGCAATTACACCGACAAC</td>
<td>CCND2 intron spanning (reverse)</td>
<td>45</td>
</tr>
<tr>
<td>NestinF</td>
<td>CTGCAGGCCACTGAAAAGT</td>
<td>Nestin intron spanning (forward)</td>
<td>2</td>
</tr>
<tr>
<td>NestinR</td>
<td>TTCCAGGATCTGAGCGATCT</td>
<td>Nestin intron spanning (reverse)</td>
<td>2</td>
</tr>
<tr>
<td>StellaF</td>
<td>GATGCACAACGATCCAGATT</td>
<td>Stella intron spanning (forward)</td>
<td>73</td>
</tr>
<tr>
<td>StellaR</td>
<td>TGGAAATTAGAAAGCTACACTCAA</td>
<td>Stella intron spanning (reverse)</td>
<td>73</td>
</tr>
<tr>
<td>NanogF</td>
<td>CCTCCAGCAGATGCAAGAA</td>
<td>Nanog intron spanning (forward)</td>
<td>25</td>
</tr>
<tr>
<td>NanogR</td>
<td>GCTTGCACTTCATCCTTGG</td>
<td>Nanog intron spanning (reverse)</td>
<td>25</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence 5′→3′</td>
<td>Description</td>
<td>UPL Probe</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------</td>
<td>-------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>TBPF</td>
<td>GGGGAGCTGTGATGTGAAGT</td>
<td>TBP intron spanning (forward)</td>
<td>97</td>
</tr>
<tr>
<td>TBPR</td>
<td>CCAGAAATAATTCTGGCTCA</td>
<td>TBP intron spanning (reverse)</td>
<td>97</td>
</tr>
<tr>
<td>Egr1F</td>
<td>CCTATGAGCACCTGACCAC</td>
<td>Egr1 intron spanning (forward)</td>
<td>22</td>
</tr>
<tr>
<td>Egr1R</td>
<td>TCGTTTGCTGGGATAACTC</td>
<td>Egr1 intron spanning (forward)</td>
<td>22</td>
</tr>
<tr>
<td>Gata6F</td>
<td>GGTCTCTACAGCAAGATGAATG</td>
<td>Gata6 intron spanning (forward)</td>
<td>40</td>
</tr>
<tr>
<td>Gata6R</td>
<td>TGGCACAGGACAGTCCAAG</td>
<td>Gata6 intron spanning (forward)</td>
<td>40</td>
</tr>
<tr>
<td>PdgfR1αF</td>
<td>GTCTTGACCTGCACTGGGA</td>
<td>PdgfR1α intron spanning (forward)</td>
<td>80</td>
</tr>
<tr>
<td>PdgfR1αR</td>
<td>CCAGCATGGTGATACCTTTGT</td>
<td>PdgfR1α intron spanning (forward)</td>
<td>80</td>
</tr>
<tr>
<td>Erk2 UPL1F</td>
<td>CAGAGCTCCAGAAATTATGTGA</td>
<td>Erk2 intron spanning (forward)</td>
<td>54</td>
</tr>
<tr>
<td>Erk2 UPL1R</td>
<td>GGGGAAGATAGGCTTGTGG</td>
<td>Erk2 intron spanning (reverse)</td>
<td>54</td>
</tr>
<tr>
<td>Erk1 UPL1F</td>
<td>TGGAAAGCCATGAGATGTAGT</td>
<td>Erk1 intron spanning (forward)</td>
<td>94</td>
</tr>
<tr>
<td>Erk1 UPL1R</td>
<td>GTCAGCCTGGCTGGCTTTA</td>
<td>Erk1 intron spanning (reverse)</td>
<td>94</td>
</tr>
<tr>
<td>cFosF</td>
<td>GGGACAGCCTTTCTACTACC</td>
<td>cFos intron spanning (forward)</td>
<td>67</td>
</tr>
<tr>
<td>cFosR</td>
<td>GATCTGCAGCAAAGTCCCTGT</td>
<td>cFos intron spanning (reverse)</td>
<td>67</td>
</tr>
<tr>
<td>Egr2F</td>
<td>CTACCGGTGAAAGACCTC</td>
<td>Egr2 intron spanning (forward)</td>
<td>60</td>
</tr>
<tr>
<td>Egr2R</td>
<td>AATGTTGATCATGCCCATCTCC</td>
<td>Egr2 intron spanning (reverse)</td>
<td>60</td>
</tr>
<tr>
<td>Dusp6F</td>
<td>CTGGTGAGAGTCCGGTCCT</td>
<td>Dusp6 intron spanning (forward)</td>
<td>66</td>
</tr>
<tr>
<td>Dusp6R</td>
<td>CGGCTTGGAACCTACTGAAG</td>
<td>Dusp6 intron spanning (reverse)</td>
<td>66</td>
</tr>
<tr>
<td>Dusp9F</td>
<td>AGTCGCCTGCCTCTGGAT</td>
<td>Dusp9 intron spanning (forward)</td>
<td>67</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence 5’—3’</td>
<td>Description</td>
<td>UPL Probe</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------</td>
<td>------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Dusp9R</td>
<td>AGGCATGACCGACTCAGACT</td>
<td>Dusp9 intron spanning (reverse)</td>
<td>67</td>
</tr>
<tr>
<td>NeuroD3F</td>
<td>CCTTCTTTGTGACTGGCTCA</td>
<td>NeuroD3 intron spanning (forward)</td>
<td>83</td>
</tr>
<tr>
<td>NeuroD3R</td>
<td>CCCTTTCTCAACACCACACTG</td>
<td>NeuroD3 intron spanning (reverse)</td>
<td>83</td>
</tr>
<tr>
<td>Soc3F</td>
<td>ATTTCGCTTTCCGGACTAGC</td>
<td>Soc3 intron spanning (forward)</td>
<td>83</td>
</tr>
<tr>
<td>Soc3R</td>
<td>AACTTGCTGTGCGGTACCAT</td>
<td>Soc3 intron spanning (reverse)</td>
<td>83</td>
</tr>
<tr>
<td>Mash1F</td>
<td>TCTCCTGGGAATGGACTTTG</td>
<td>Mash1 intron spanning (forward)</td>
<td>74</td>
</tr>
<tr>
<td>Mash1R</td>
<td>CGTTGGCGAGAACACTAAAG</td>
<td>Mash1 intron spanning (reverse)</td>
<td>74</td>
</tr>
<tr>
<td>Ngn2F</td>
<td>ACATCTGGAGGCGCGGTAG</td>
<td>Ngn2 intron spanning (forward)</td>
<td>69</td>
</tr>
<tr>
<td>Ngn2R</td>
<td>CCCAGCAGCAGGATCAGTACCTC</td>
<td>Ngn2 intron spanning (reverse)</td>
<td>69</td>
</tr>
<tr>
<td>TF</td>
<td>CAGCCCACCTACTGGCTCTA</td>
<td>T intron spanning (forward)</td>
<td>100</td>
</tr>
<tr>
<td>TR</td>
<td>GAGCCTGGGGTTGATGTTA</td>
<td>T intron spanning (reverse)</td>
<td>100</td>
</tr>
<tr>
<td>Rex1F</td>
<td>GCACACTCACTCTATTGAGAGA</td>
<td>Rex1 intron spanning (forward)</td>
<td>16</td>
</tr>
<tr>
<td>Rex1R</td>
<td>CAGCTCCTGCACACAGAAGA</td>
<td>Rex1 intron spanning (reverse)</td>
<td>16</td>
</tr>
<tr>
<td>CCND1F</td>
<td>TCTTCCAGAGTCAAGGCTG</td>
<td>CCND1 intron spanning (forward)</td>
<td>72</td>
</tr>
<tr>
<td>CCND1R</td>
<td>GACTCCAGAAGGGCTTCAATC</td>
<td>CCND1 intron spanning (reverse)</td>
<td>72</td>
</tr>
<tr>
<td>CCND3F</td>
<td>GGAAGATGCTGGCACTTTG</td>
<td>CCND3 intron spanning (forward)</td>
<td>88</td>
</tr>
<tr>
<td>CCND3R</td>
<td>GGTAGCGATCCAGGATGTTCA</td>
<td>CCND3 intron spanning (reverse)</td>
<td>88</td>
</tr>
<tr>
<td>SrfF</td>
<td>CTGAGCAGCAGGAGGAAC</td>
<td>Srf intron spanning (forward)</td>
<td>10</td>
</tr>
<tr>
<td>SrfR</td>
<td>GCTGGGCTGTGCTGAGAT</td>
<td>Srf intron spanning (reverse)</td>
<td>10</td>
</tr>
<tr>
<td>Klf4F</td>
<td>CGGGAAGGAGAAGACACT</td>
<td>Klf4 intron spanning (forward)</td>
<td>62</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence 5’→3’</td>
<td>Description</td>
<td>UPL Probe</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------</td>
<td>-------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Klf4R</td>
<td>GAGTTCCCTCACGCAACG</td>
<td>Klf4 intron spanning (reverse)</td>
<td>62</td>
</tr>
<tr>
<td>Tbx3F</td>
<td>TTGCAAGGGTTTTCGAGAC</td>
<td>Tbx3 intron spanning (forward)</td>
<td>51</td>
</tr>
<tr>
<td>Tbx3R</td>
<td>TGCAGTGAGCTGCTTTCT</td>
<td>Tbx3 intron spanning (reverse)</td>
<td>51</td>
</tr>
<tr>
<td>Foxa2F</td>
<td>GAGCAGCAACATCACACCAG</td>
<td>Foxa2 intron spanning (forward)</td>
<td>77</td>
</tr>
<tr>
<td>FoxA2R</td>
<td>CGTAGGCCTTGAGGTCCAT</td>
<td>Foxa2 intron spanning (reverse)</td>
<td>77</td>
</tr>
<tr>
<td>Gata4F</td>
<td>TTCGCTTTCCTCCCTCAAG</td>
<td>Gata4 3’ UTR non-intron spanning (forward)</td>
<td>60</td>
</tr>
<tr>
<td>Gata4R</td>
<td>CAATGTTAACGGGTGTGGA</td>
<td>Gata4 3’ UTR non-intron spanning (reverse)</td>
<td>60</td>
</tr>
<tr>
<td>Acta2F</td>
<td>CCAGCACCATGAAGATCAAG</td>
<td>Acta2 intron spanning (forward)</td>
<td>58</td>
</tr>
<tr>
<td>Acta2R</td>
<td>TGGAAGGTAGACAGCGAAGC</td>
<td>Acta2 intron spanning (reverse)</td>
<td>58</td>
</tr>
</tbody>
</table>

### 2.1.5: Antibodies

<table>
<thead>
<tr>
<th>Target</th>
<th>Species</th>
<th>Clone</th>
<th>Vendor</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gapdh</td>
<td>Mouse monoclonal</td>
<td>G8795</td>
<td>Sigma</td>
<td>WB: 1:35,000</td>
</tr>
<tr>
<td>Erk1/2</td>
<td>Rabbit polyclonal</td>
<td>9102</td>
<td>Cell Signaling</td>
<td>WB: 1:1000, IHC: 1:100</td>
</tr>
<tr>
<td>pErk1/2</td>
<td>Rabbit polyclonal</td>
<td>9101</td>
<td>Cell Signaling</td>
<td>WB: 1:500</td>
</tr>
<tr>
<td>pErk1/2</td>
<td>Mouse monoclonal</td>
<td>E10</td>
<td>Cell Signaling</td>
<td>WB: 1:1000, IHC: 1:100</td>
</tr>
<tr>
<td>Target</td>
<td>Species</td>
<td>Clone</td>
<td>Vendor</td>
<td>Dilution</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------</td>
<td>-------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>RSK1/2/3</td>
<td>Rabbit polyclonal</td>
<td>9355</td>
<td>Cell Signaling</td>
<td>WB: 1:1000</td>
</tr>
<tr>
<td>pRSK (T359/S363)</td>
<td>Rabbit polyclonal</td>
<td>9344</td>
<td>Cell Signaling</td>
<td>WB: 1:1000</td>
</tr>
<tr>
<td>pRSK (S573)</td>
<td>Rabbit polyclonal</td>
<td>9335</td>
<td>Cell Signaling</td>
<td>WB: 1:1000</td>
</tr>
<tr>
<td>Srf</td>
<td>Rabbit polyclonal</td>
<td>sc-335</td>
<td>Santa Cruz</td>
<td>WB: 1:1000</td>
</tr>
<tr>
<td>pSrf (S103)</td>
<td>Rabbit polyclonal</td>
<td>4262</td>
<td>Cell Signaling</td>
<td>WB: 1:1000</td>
</tr>
<tr>
<td>β III Tubulin</td>
<td>Mouse monoclonal</td>
<td>Tuj-1</td>
<td>R&amp;D Systems</td>
<td>IHC: 1:1000</td>
</tr>
<tr>
<td>Nanog</td>
<td>Rabbit polyclonal</td>
<td></td>
<td>Gift from Ian Chambers</td>
<td>WB: 1:1000, IHC: 1:2000</td>
</tr>
<tr>
<td>FLAG</td>
<td>Mouse monoclonal</td>
<td>M2</td>
<td>Sigma</td>
<td>WB: 1:2000, IHC: 1:1000</td>
</tr>
</tbody>
</table>

2.1.6: Cell Lines

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Selection</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E14tg2a</td>
<td>Wild type, hypoxanthine phosphoribosyltransferase deficient, mouse embryonic stem cell line.129/Ola background (Clonal)</td>
<td></td>
<td>Hooper et al., 1987</td>
</tr>
<tr>
<td>Ju09</td>
<td>Wild type, hypoxanthine phosphoribosyltransferase deficient, mouse embryonic stem cell line derived from E14tg2a chimeric embryos. 129/Ola background (Clonal)</td>
<td></td>
<td>This Study</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
<td>Selection</td>
<td>Reference</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>E14/T</td>
<td>E14tg2a derived cell line expressing SV40 Large T antigen for episomal maintenance of plasmid DNA containing a SV40 origin of replication (Clonal)</td>
<td>G418</td>
<td>Gassmann et al., 1995</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aubert et al., 2002</td>
</tr>
<tr>
<td>B1</td>
<td>Erk2 homozygous null mouse embryonic stem cell line. 129/Ola-C57B6 mixed background (Clonal)</td>
<td>G418</td>
<td>Saba-El-Leil et al., 2003</td>
</tr>
<tr>
<td>Fl3</td>
<td>Srf heterozygous floxed mouse embryonic stem cell line. 129/Ola-C57B6 mixed background (Clonal)</td>
<td></td>
<td>This Study</td>
</tr>
<tr>
<td>Sfh9</td>
<td>Srf homozygous floxed mouse embryonic stem cell line. C57B6j background (Clonal)</td>
<td></td>
<td>This Study</td>
</tr>
<tr>
<td>PDK1/-/-</td>
<td>E14tg2a derived PDK1 homozygous null cell line (Clonal)</td>
<td></td>
<td>Williams et al., 2003</td>
</tr>
<tr>
<td>PDK1/-/-</td>
<td>PDK1/-/- derived cell line expressing FLAG tagged PDK1 cDNA (Population)</td>
<td>Puromycin</td>
<td>This Study</td>
</tr>
<tr>
<td>PDK1/-/-</td>
<td>PDK1/-/- derived cell line expressing FLAG tagged mutated PDK1-K114G cDNA (Population)</td>
<td>Puromycin</td>
<td>This Study</td>
</tr>
<tr>
<td>JuD2/</td>
<td>Ju09 derived cell lines heterozygously depleted for exon 3 of Erk2 (Clonal)</td>
<td>Hygromycin B</td>
<td>This Study</td>
</tr>
<tr>
<td>JuD7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2N4/</td>
<td>JuD2/JuD7, respectively, derived cell lines homozgyously depleted for exon 3 of Erk2 (Clonal)</td>
<td>Hygromycin B</td>
<td>This Study</td>
</tr>
<tr>
<td>D7N2</td>
<td></td>
<td>G418</td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
<td>Selection</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>D2WT/D7WT</td>
<td>D2N4/D7N2, respectively, derived cell lines expressing 3xFLAG tagged Erk2 cDNA (Clonal)</td>
<td>Hygromycin B</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G418</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Puromycin</td>
<td></td>
</tr>
<tr>
<td>D2TNG/D7TNG</td>
<td>D2N4/D7N2, respectively, derived cell lines where an eGFP cDNA has been knocked into the ATG codon of the endogenous Nanog allele (Clonal)</td>
<td>Hygromycin B</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G418</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Puromycin</td>
<td></td>
</tr>
<tr>
<td>D2TNG-WT/D7TNG-WT</td>
<td>D2TNG/D7TNG, respectively, derived cell lines expressing 3xFLAG tagged Erk2 cDNA (Clonal)</td>
<td>Hygromycin B</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G418</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Puromycin</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blasticidin S</td>
<td></td>
</tr>
<tr>
<td>D7E1</td>
<td>D7N2 derived cell line expressing 3xFLAG tagged Erk1 cDNA (Population)</td>
<td>Puromycin</td>
<td>This Study</td>
</tr>
<tr>
<td>D7E2</td>
<td>D7N2 derived cell line expressing 3xFLAG tagged Erk2 cDNA (Population)</td>
<td>Puromycin</td>
<td>This Study</td>
</tr>
<tr>
<td>RZZ307</td>
<td>E14tg2a derived Erk2 genetrapped cell line (Clonal)</td>
<td>Neomycin</td>
<td></td>
</tr>
<tr>
<td>Cos7</td>
<td>African Green Monkey SV40-transfected kidney fibroblast cell line</td>
<td>Jensen et al., 1964</td>
<td></td>
</tr>
</tbody>
</table>
2.1.7: Culture Medium

2.1.7.1: Bacterial culture medium

Terrific broth (TB) (1L) 12g Tryptone, 24g Yeast extract, 4ml Glycerol, autocalved and 10% (v/v) sterile salt solution (0.17M KH$_2$PO$_4$ and 0.72M K$_2$HPO) was added before use.

SOB medium (1L) 20g Tryptone 5g Yeast extract 0.5g NaCl, 2.5mM KCl, autoclaved pH adjusted to 7.0. 10mM sterile MgCl$_2$ was added before use.

SOC medium SOB medium containing 20mM Glucose

2.1.7.2: Mammalian cell culture medium

mES cell growth medium Glasgow modified Eagle’s medium (GMEM), 10% foetal calf serum, 1mM sodium pyruvate, 1x MEM non-essential amino acids, 2mM L-glutamine, 0.1mM β-mercaptoethanol, 1000U Lif.

N2B27 medium 100ml Neurobasal, 100mL DMEM/F12 (without L-glutamine), 0.1mM β-mercaptoethanol, 2mM L-glutamine, 1ml N2 supplement, 2ml B27 supplement

2i medium N2B27 supplemented with 3µM CHIR99021 and 1µM PD0325901, 1000U Lif

Mesoderm medium α- MEM supplemented with 10% FCS and 0.2mM β-mercaptoethanol

COS7 Medium mES medium without Lif

2.1.8: In-silico resources

<table>
<thead>
<tr>
<th>Software</th>
<th>Application</th>
<th>Developer</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SeqBuilder</td>
<td>DNA annotation and manipulation tool</td>
<td>DNASTar Inc.</td>
<td><a href="http://www.dnastar.com/">http://www.dnastar.com/</a></td>
</tr>
<tr>
<td>4Peaks</td>
<td>Sequence trace viewer and editor</td>
<td>Mekentosj</td>
<td><a href="http://mekentosj.com/science/4peaks/">http://mekentosj.com/science/4peaks/</a></td>
</tr>
<tr>
<td>Software</td>
<td>Application</td>
<td>Developer</td>
<td>URL</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Align</td>
<td>Pairwise Alignment Algorithm</td>
<td>EBI</td>
<td><a href="http://www.ebi.ac.uk/Tools/emboss/align/">http://www.ebi.ac.uk/Tools/emboss/align/</a></td>
</tr>
<tr>
<td>ClustawlW</td>
<td>A general purpose multiple sequence alignment program for DNA or proteins</td>
<td>EBI</td>
<td><a href="http://www.ebi.ac.uk/Tools/msa/clustalw2/">http://www.ebi.ac.uk/Tools/msa/clustalw2/</a></td>
</tr>
<tr>
<td>iRNAi</td>
<td>RNAi design tool</td>
<td>Mekentosj</td>
<td><a href="http://mekentosj.com/science/irnai/">http://mekentosj.com/science/irnai/</a></td>
</tr>
<tr>
<td>RNAi Selection</td>
<td>RNAi design tool</td>
<td>Whitehead Institute</td>
<td><a href="http://jura.wi.mit.edu/bioc/siRNAext/">http://jura.wi.mit.edu/bioc/siRNAext/</a></td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool for DNA or protein sequences</td>
<td>Ensembl</td>
<td><a href="http://www.ensembl.org/Multi/blastview">http://www.ensembl.org/Multi/blastview</a></td>
</tr>
<tr>
<td>Oligocalc</td>
<td>Oligo analysis tool</td>
<td>Northwestern University</td>
<td><a href="http://www.unc.edu/~cail/biotool/oligo/">http://www.unc.edu/~cail/biotool/oligo/</a></td>
</tr>
<tr>
<td>Tm Calculator</td>
<td>Melting temperature calculator</td>
<td>Finnzymes</td>
<td><a href="http://www.finnzymes.fi/tm_determination_old.html">http://www.finnzymes.fi/tm_determination_old.html</a></td>
</tr>
<tr>
<td>UPL assay design center</td>
<td>Tool for UPL Q-PCR oligo selection</td>
<td>Roche</td>
<td><a href="https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=uplct_030000">https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=uplct_030000</a></td>
</tr>
<tr>
<td>Scansite</td>
<td>Protein domain prediction tool</td>
<td>MIT</td>
<td><a href="http://scansite.mit.edu/">http://scansite.mit.edu/</a></td>
</tr>
<tr>
<td>Mascot</td>
<td>Protein primary sequence databases</td>
<td>Matrix Science</td>
<td><a href="http://www.matrixscience.com/">http://www.matrixscience.com/</a></td>
</tr>
</tbody>
</table>
2.2 Methods

2.2.1: DNA Techniques

2.2.1.1: DNA isolation from bacteria
2.2.1.1.1: Plasmid isolation from DH5α

LB broth, 4ml and 100ml, containing the appropriate antibiotic was inoculated with a single bacterial colony for mini and maxi preparations, respectively. Cultures were incubated at 37°C with agitation at 200 rpm for approximately 16 hours. The following day bacteria were collected by centrifugation at 4000g for 10 minutes. Plasmid DNA was isolated using QIAprep™ Spin Miniprep, or Plasmid Maxi Kits for mini or maxi preparations, respectively, according to manufacturer’s instructions.

2.2.1.1.2: BAC isolation from DH10β

LB broth (250ml) containing the appropriate antibiotic was inoculated with a single bacterial colony and incubated at 37°C, with agitation at 200 rpm, for approximately 16 hours. The following day bacteria were collected by centrifugation at 4000g for 10 minutes. Bacterial pellets were resuspended in 20ml 10mM EDTA pH 8, place on ice for 5 minutes. Cells were lysed by addition of 40ml of SDS/NaOH (1%SDS, 0.2M NaOH) and incubation on ice for 10 mins. Lysis was stopped and chromatin denatured by the addition of 30ml 3M KOAc and incubation on ice for 25 minutes. Lysates were cleared by centrifugation at 10000rpm for 30 minutes. BAC DNA was isolated from the supernatant by addition of 45ml isopropanol and centrifugation at 3200rpm for 30 minutes. BAC DNA pellets were resuspended in in 8ml 10mM Tris/50mM EDTA, pH8 and extracted by addition of 4ml 7.5M KOAc and incubation at –80°C for 30 minutes, followed by centrifugation at 3200rpm for 20 minutes to remove contaminants. BAC DNA was recovered from the supernatant by EtOH precipitation (60% EtOH final and centrifugation at 3500rpm for 10 minutes). The pellet was resuspended in 4ml 50mM tris/50mM EDTA, pH8, and
treated with 20µl DNase-Free RNase (10mg/ml at 37°C for 1 hour. DNA was extracted twice with phenol and once with chloroform, precipitated with EtOH as before and resuspended in 200µl ddH$_2$O.

2.2.1.2: DNA isolation from mammalian cells

2.2.1.2.1: DNA isolation from 96 well plate

Cells were grown to ~80% confluence in flat-bottom 96 well plates under standard conditions. Cells were washed 3 times in ice cold PBS and then lysed in 50µl Sarcosyl lysis buffer supplemented with 1mg/ml Proteinase K (added freshly), and incubated at 65°C for 3 hours in a sealed humidified box. Genomic DNA was precipitated by addition of 150µl precipitation buffer (150mM NaCl in 100% EtOH) and incubated on ice for 10 minutes. Genomic DNA was collected by centrifugation at 4000g, and excess salt was removed by washing 3 times with 200µl of 70% EtOH. DNA pellets were air-dried to remove excess EtOH and resuspended overnight in 50µl ddH$_2$O at 37°C.

2.2.1.2.2: DNA isolation from 6 well plate

Cells were grown to ~80% confluence in flat-bottom 6 well plates under standard conditions, equivalent to ~10$^6$ cells. Genomic DNA was isolated using the Qiagen- DNeasy™ Blood And Tissue Kit according to the protocol for DNA isolation from Blood. Elution buffer was warmed and passed over the column twice to increase the DNA yield.

2.2.1.3: DNA isolation from mouse ear clippings

Mouse ear clippings were taken for genotyping at weaning. Biopsies were incubated in 25µl of ‘Ear Buffer’ and boiled for 20 minutes, then cooled on ice. Samples were neutralised by addition of 25µl of 40mM Tris-HCl (pH 8).

2.2.1.4: DNA cloning

2.2.1.4.1: Restriction enzyme digestion
For general cloning purposes between 1-2µg of purified DNA was digested in 50µl volumes containing the appropriate buffer supplied by the manufacturer, and supplemented with 100ng BSA if required. Reactions were performed at a temperature appropriate for the enzyme being used. For complete DNA digestion samples were treated with 10U of enzyme and allowed to digest for ~2 hours. For partial digests samples were treated with between 0.5 and 10U of enzyme for 20 minutes, or as required. Conditions for partial digest were generally determined empirically.

2.2.1.4.2: Dephosphorylation of DNA fragment ends

5’ phosphates groups of DNA ends we removed by incubating 1-2ug of digested DNA with 1U Calf Intestinal Phosphatase (CIP) according to manufacturer’s instructions. Reactions were preformed at 37°C for 30 minutes.

2.2.1.4.3: Blunting DNA fragment overhangs

3’ and 5’ overhangs were removed or filled in, respectively, by T4 DNA polymerase treatment. 1U of T4 DNA polymerase was used per 1ug of DNA according to manufacturers instructions except that reactions were incubated at room temperature for 20 minutes, and stopped by boiling for 20 minutes.

2.2.1.4.4: DNA fragment ligation

Cohesive DNA ends were ligated using T4 DNA ligase by incubating DNA fragments, at a molar ratio of 1:3, backbone to insert, according to manufacturer’s instructions. Reactions were carried out at room temperature and allowed to proceed no longer than 20 minutes. Blunt DNA ends were also ligated using T4 DNA ligase although reactions were supplemented with PEG6000 to a final concentration of 1.5% (w/v), and incubated at 16°C overnight.
2.2.1.4.5: DNA electrophoresis

DNA was analysed and separated for purification by agarose gel electrophoresis. 0.8-2% (w/v) gels were used, depending on the size of fragments to be resolved, and contained 0.5x Sybr Safe for visualization either by eye or UV illumination. Samples were run in 1x sucrose loading buffer, alongside a 1Kb DNA marker for size comparison, at 100V for 45 minutes.

2.2.1.4.6: DNA extraction from agarose gels

DNA was extracted from gel slices using the Zymo Research-Zymoclean™ Gel DNA Recovery Kit according to manufacturer’s instructions and eluted in 20µl ddH$_2$O. For fragments 8Kb and larger, ddH$_2$O was heated to 50°C prior to elution to increase yields.

2.2.1.4.7: DNA extraction from solution

DNA was purified from solution by either phenol extraction or using the Zymo Research- DNA Clean and Concentrator kit. For phenol extractions, DNA in a volume of no less than 200µl was mixed with an equal volume of a solution of phenol:chloroform:isoamyl alcohol (25:24:1) and vortexed vigorously. Water and phenol phases were separated by centrifugation at maximum speed in a bench top microfuge for 20 minutes. The water/DNA phase (top) was recovered and to it an equal volume of chloroform was added to remove any phenol carryover. Water/chloroform phases were separated by centrifugation and DNA was recovered from the water phase by salt precipitation.

2.2.1.4.8: DNA precipitation

DNA was precipitation by incubation with 100mM NaOAc (pH6.3) in 70% EtOH on ice for 10 minutes. Precipitated DNA was recovered by centrifugation at maximum speed in a bench-top microfuge at 4°C. Excess salt was removed by washing the DNA pellet twice with ice cold 70% EtOH. DNA was resuspended in an appropriate volume of ddH$_2$O.
2.2.1.4.9: DNA quantification

DNA was quantified by UV spectrometry, using a NanoDrop® ND-1000 spectrophotometer. The Beer-Lambert equation was used to correlate the calculated absorbance with concentration, with extinction coefficients of 50 for dsDNA and 33 for ssDNA.

2.2.1.4.10: Bacterial transformation

2.2.1.4.10.1 Generation of chemically competent DH5α

A glycerol stock of DH5α E.coli cells was streaked onto an LB plate and incubated overnight at 37°C. The following day a single colony was picked and seeded in 250ml of LB broth and incubated overnight at 18°C at 250rpm. 200µl of the starter culture was then used to seed 250ml of LB broth and was incubated until the OD600 reached 0.3, approximately 24 hours. The culture was chilled on ice and collected by centrifugation at 3000g at 4°C. The bacterial pellet was washed in 40ml of ice cold CCMB80 and incubated on ice for 10 minutes. This step was repeated twice. Bacteria were again collected by centrifugation and resuspended in 10ml ice cold CCMB80 supplemented with 7.5% DMSO (w/v) just before use. The bacterial suspension was then divided in 100µl aliquots in 1.5ml microfuge tubes that had been pre-chilled on dry ice. Aliquots were stored at -80°C. This method typically generated $10^8$ CFU/µg of pUC19 plasmid transformed.

2.2.1.4.10.2: Transformation of plasmid DNA into DH5α

For routine transformations either 2µl of a ligation mix, or less than 1µg plasmid DNA, was incubated on ice for ~5 minutes with 200µl of freshly thawed competent E.coli in a 15ml Falcon tube. Bacteria were heat shocked for 30 seconds at 42°C and immediately placed on ice. 500µl of SOC medium was added to the tube and the culture was allowed recover at 37°C for ~40 minutes (depending on selection being used) at 250rpm. Between 20-100µl of the culture was plated on LB plates (with appropriate antibiotic) and incubated overnight at 37°C.
2.2.1.5: Polymerase chain reaction

2.2.1.5.1: Generic PCR

In general PCR reactions were prepared in 50µl volumes, with 200µM dNTP. The DNA polymerase concentrations used were specified by the manufacturer and reactions were made-up with the reaction buffer supplied with the polymerase. In general less than 1ng of template DNA was used/reaction. All reactions were preformed on a peltier thermal cycler (DYAD™ DNA Engine). Primer annealing temperatures were calculated using the nearest neighbor method (http://www.finnzymes.fi/tm_determination_old.html and http://www.unc.edu/~cail/biotool/oligo/). Phusion Pfu DNA polymerase (Finnzymes) was used for all general cloning PCR reactions, Vent Pfu DNA polymerase (NEB) was used for amplifying short hairpins, and Qiagen Taq DNA Polymerase was used for all other purposes. All PCR reactions contained an initial denaturation step, 1 minute at 94°C (Qiagen Taq DNA Polymerase) or 98°C (Phusion or Vent Pfu DNA Polymerase) followed by successive cycles of denaturation (15 seconds at 94°C (Qiagen Taq DNA Polymerase) or 98°C (Phusion or Vent Pfu DNA Polymerase), annealing (15 seconds at 1 degree lower than calculated the lowest primer annealing), and extension (30 seconds/Kb to be amplified (Qiagen Taq DNA Polymerase) or 15 seconds/Kb to be amplified (Phusion or Vent Pfu DNA Polymerase) at 72°C. Each cycle was repeated 30 times (Qiagen Taq DNA Polymerase) or 25 times (Phusion or Vent Pfu DNA Polymerase), followed by a final incubation at 72°C for 10 minutes. If the annealing temperature of the primer pair was over 72°C, the annealing step was omitted. If the primers being used contained non-homologous overhangs the initial annealing temperature (for the first 5 cycles) was calculated from only the homologous region of the primer, the annealing temperature for the remaining cycles was calculated from the full length primer. A negative control containing no DNA template was used for all reactions.
2.2.1.5.1.1: Site-directed DNA mutagenesis

Substitution or removal of specific nucleotides in plasmid DNA was performed using a PCR based method. Briefly, primers were designed against the region to be mutated, and contained the desired alteration. The primers were designed to that they annealed with 10bp of homology 5’ to the region to be mutated and 15bp of homology 3’ to the region to be mutated. Primer annealing temperatures were calculated for the homologous region of the primer only. All reactions were preformed using Phusion DNA Polymerase according to usual except that extension times were increased to 1 minute/Kb to be amplified (i.e. the size of the entire plasmid), and reactions were preformed for only 15 cycles. Following amplification, the reaction mixture was incubated at 37°C with 10U DpnI for 2 hours to digest the template DNA. 2µl of the reaction was transformed into competent bacteria and plated on LB agar plates with the appropriate selection. Individual colonies were screened by DNA sequencing. A negative control reaction lacking DNA polymerase was included for all experiments.

2.2.1.5.1.2: PCR fusion by overlap extension

To generate a single PCR product from 2 distinct templates (as in when cloning fragments to express fusion proteins), PCR fusion by overlap extension was employed. Primers were designed against template A and B so that the 3’ end of the reverse primer to template A contained ~20bp homology to the 5’end of the forward primer of template B, and that this region of homology had a lower annealing temperature than the forward primer to template A and the reverse primer B. Fragments A and B were amplified as usual, separated by agarose gel electrophoresis and gel extracted. 1ng of each of fragment A and B were then combined in a new PCR reaction with the forward primer to template A and the reverse primer to template B only. If the subsequent reaction produced a single discrete fragment it was purified and cloned into an appropriate vector for sequencing. This procedure was also used to insert or delete DNA sequences from plasmid DNA.
2.2.1.5.2: Blunt cloning of PCR products

PCR products were directly cloned using the Invitrogen Zero Blunt® TOPO® PCR Cloning Kit according to the manufacturer’s instructions.

2.2.1.5.3: Quantitative genomic PCR

Quantitative genomic PCR for screening clones for targeted loss of a region of DNA was preformed using the Roche UPL System. Briefly, DNA was extracted according to methods outlined in sections 2.2.1.2.1 and 2.2.1.2.2. Primers were designed using the Roche Assay Design Center (https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=uplct_030000) and selected so as to be non-intron spanning, each pair having a specific probe. Reactions were preformed using the LightCycler® 480 Probes Master Mix (Roche) and UPL Set, Mouse (Roche) as outlined by the manufacturer. Abundance of the targeted region was calculated using the ΔΔCt method normalized against a non-targeted genomic region whose amplification primers had an identical amplification efficiency to the targeted region.

2.2.1.5.4: Quantitative RT-PCR

Quantitate RT-PCR was preformed as outlined in section 2.2.1.5.3 except that first strand cDNA was used as a template, section 2.2.2.2, and relative concentrations were calculated from a standard curve and normalized against the house-keeping gene TBP.

2.2.1.6: DNA sequencing

Plasmid DNA was sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction kit at the Genepool Sequencing Center (University of Edinburgh). Samples were submitted in 6µl volumes containing 500ng of DNA template with 1µ of a 3.2pmole/µl primer solution. Sequence trace files were read using 4Peaks (http://mekentosj.com/science/4peaks/).
2.2.1.7: Southern blotting

Restriction fragment length polymorphisms were determined by Southern blotting (Southern, 1975).

2.2.1.7.1: Sample preparation

Genomic DNA was prepared as outlined in section 2.2.1.2.2. 5µg of genomic was digested in 50µl volumes with the 30U of enzyme, over night, at the appropriate temperature.

2.2.1.7.2: DNA electrophoresis and denaturation

DNA fragments were separated by agarose gel electrophoresis (0.8% gel) alongside a 1Kb size marker, overnight at 30V. The following day the gel was stained with ethidium bromide in 1xTAE at a concentration of 0.5mg/ml for 1 hour. The gel was then washed twice with 1xTAE and transferred to a UV trans-illuminator, where it was exposed for 1 minute to nick the DNA, and for photographic imaging. The efficiency of digestion was determined by the extend of DNA laddering. The gel was then incubated in Denaturing Solution for 40 minutes with gentle rocking. The gel was then washed with Neutralisation Solution for 10 minutes followed by a 5 minute wash in ddH₂O.

2.2.1.7.3: Sample transfer

DNA was then transferred overnight to a nitrocellulose membrane by capillary action (Figure 2.1). The following day the membrane was removed from the transfer apparatus and rinsed in 2x SSC. DNA was UV cross-linked to the membrane using a Stratalink oven and subsequently baked at 120°C for 1 hour.
2.2.1.7.4: Probe preparation and hybridization

The membrane was then placed in a glass hybridization and incubated at 65°C with 14.5ml PerfectHyb solution (Sigma) for a minimum of 1 hour. During this time the probe was labeled. The probe template was amplified from either plasmid DNA or BAC DNA containing the desired genomic sequence from a 129/Ola genomic library, as appropriate. Probe templates were generally between 200-700bp in length. The probe was labelled using the Amersham MegaPrime™ DNA Labeling System according to manufacturers instructions, using \([\alpha^{32}]\text{dCTP}, 3000 \text{ Ci/mmol.}\)

This method of random priming generated probe fragments of an average length of 40bp, which were then purified from unbound primer and nucleotide using the Amersham Illustra ProbeQuant™ G-50 Micro Columns, as outlined in the product manual. 500 µl of Sonicated Herring Sperm DNA was added to the purified probe and boiled at 100°C for 5 minutes and then chilled on ice for 10 minutes. The mixture was then added to the hybridisation tube and incubated with the membrane overnight, between 63-68°C. The exact hybridisation temperature was determined empirically.

2.2.1.7.5: Membrane washing and exposure

The hybridization/probe solution was disposed of and the membrane was washed twice in 2xSSC wash buffer at 65°C fro 15 minutes and once for 30 minutes
in 0.5x SSC wash buffer. The membrane was then wrapped in cling film and exposed to X-ray film (Kodak) overnight at 80°C. The following day the film was removed and developed in an automated autoradiographic film processor.

### 2.2.2: RNA Methods

For routine work with RNA all tube and solutions were RNase free. Solutions were made with DEPC treated water, microfuge tubes and pipette tips were RNase free (Star Labs) and all reactions were preformed using RNase/DNase free water (Gibco).

#### 2.2.2.1: Total RNA extraction

Total RNA was extracted using either RNeasy™ Mini Kit or Tri Reagent (Qiagen) according to manufacturer’s instructions. Briefly, cells were harvested at approximately 80% confluence in a 6 well dish and lysed in dish. If samples were processed using the RNeasy™ Mini Kit, 350µl buffer RLT with 10µl/ml β-mercaptoethanol was added directly to the culture dish and lysed on ice for 5 minutes. Lysates were transferred to Qiashredder™ columns and centrifuged to homogenise. Genomic DNA was removed by on-column purification and the flow-through was loaded onto RNeasy MinElute Spin Columns for RNA capture and washing. RNA was eluted in 20µl of RNase free water. If samples were processed using Tri Reagent cells were lysed in 1ml of reagent. 200µl of chloroform was added to the mixture and vigorously vortexed. Phases were resolved by centrifugation and RNA was precipitated from the water phase by NaOAc/isopropanol precipitation. The RNA pellet were washed twice in ice cold 70% EtOH, air dried and resuspended in 20µl of 20µl of RNase free water. Genomic DNA was digested with DNase (Promega) for 30 minutes at 37°C. The reaction was stopped by boiling the mixture in the presence of EDTA for 10 minutes. Carryover DNase was removed by re-extraction with Tri-Reagent.
2.2.2.2: RNA quantification

RNA was quantified by UV spectrometry, using a NanoDrop® ND-1000 spectrophotometer. The Beer-Lambert equation was used to correlate the calculated absorbance with concentration, with an extinction coefficient of 40 for RNA.

2.2.2.3: First strand synthesis using random hexamers

First strand synthesis was performed using 2µg of total RNA as a template. Either M-MLV or SuperScript III RT was used according to manufacturer’s instructions and the RT reaction was primed using random hexamers (Thermo Fisher) in the presence of RNase inhibitor, RNaseOUT (Invitrogen), in a reaction volume of 20µl. Following first strand synthesis the reaction mixture was diluted to 80µl, and 2µl was used per PCR reaction.

2.2.3: Protein Methods

All protein work was performed on ice, and in general, where possible, in the cold room. Protease and phosphatase inhibitors were added fresh to all buffers, and DTT was both made and added freshly.

2.2.3.1: Total protein isolation from mammalian cells

Samples were washed twice in ice cold PBS and then lysed on ice with 1x Laemmli buffer (with Bromophenol Blue). In general a semi-confluent 6 well dish was lysed in 400µl of buffer, and scaled accordingly for different size dishes. Samples were allowed to lyse on ice for ~20 minutes and then scraped into a 1.5ml microfuge tube. Samples were sonicated in a probe tip sonicator (Soniprep) for 8 seconds at 22 amplitude microns, followed by centrifugation for 10 minutes at 14000 rpm at 4°C. 10 µl lysate (equivalent to ~20ug total protein) was used for Western Blotting.
2.2.3.2: Protein quantification

Approximate protein concentration was estimated by UV spectrometry, using a NanoDrop® ND-1000 spectrophotometer, taking a $A_{280}$ reading to calculate protein concentration. Where possible, protein concentration was determined using a Bradford assay (Biorad) according to reagent guidelines. Briefly, protein standards ranging from 0.1-2mg/ml of BSA were prepared in 5µl volumes. Samples were diluted 1:5 and 1:2 again in a total volume of 5µl. 250µl of Bradford reagent was added to each sample and standard curve dilution, mixed and incubated at room temperature for approximately 10 minutes. The absorbance was read using a UV spectrophotometer at 595nm and sample concentrations were extrapolated from the standard curve derived from protein standard measurements.

2.2.3.3: Protein separation by SDS-PAGE

SDS-PAGE electrophoresis was preformed using the NuPAGE system (Invitrogen). Briefly, samples were boiled in the presence of 100mM DTT for 5 minutes, chilled on ice and collected by brief centrifugation. The electrophoresis chamber was set up as outlined by the manufacturer and both inner and outer chambers were filled with 1x NuPAGE MES buffer, ensuring there was no contact between the inner and outer buffer reservoirs. The gel wells were washed free from storage buffer and samples were loaded, usually 10µl volumes. All samples were ran against 8µl SeeBlue™ Prestained Protein Standard (Invitrogen) for size estimation. Gels were ran at 180V until the 3kDa band of the protein marker reached the end of the gel (typically 1 hour). The apparatus was then disassembled, the gel removed from the casing and washed briefly with ddH₂O in preparation for either staining or Western Blotting.

2.2.3.4: Protein detection by coomassie staining

Coomassie staining was carried out using SimplyBlue™ Safe Stain (Invitrogen). Briefly, gels were ran as outlined in section 2.2.3.3, followed by immersion in staining solution (~100ml) and incubation at room temperature, with gentle rocking, for approximately 1 hour, or until bands became visible. When
desired band intensity was reached the staining solution was disposed of and the gel was washed 3 times in 100ml ddH\textsubscript{2}O for 1 hour per wash, followed by an on overnight wash in 200ml ddH\textsubscript{2}O.

2.2.3.5: Protein detection by silver staining

Silver staining was preformed using the SilverQuest™ Sliver Staining Kit (Invitrogen) as outlined by the manufacturer. In general the rapid microwave protocol was used and the gel was allowed develop for between 2-8 minutes before adding the stop solution. The gel was then washed 3 times in 100ml ddH\textsubscript{2}O for 10 minutes per wash.

2.2.3.6: Western blotting

2.2.3.6.1: Sample preparation

For analysis for protein samples prepared as outlined in section 2.2.3.1, 10µl of lysate was loaded per well and separated as outlined in section 2.2.3.3. For all other samples types 4x LDS Sample Loading Buffer (Invitrogen) was diluted to a working concentration with sample before processing as outlined in section 2.2.3.3.

2.2.3.6.2: Protein transfer to nitrocellulose membrane

A 7x9cm piece of Hybond™ nitrocellulose membrane (Amersham) was soaked in ddH\textsubscript{2}O for 1 minute, followed by Transfer buffer for 10 minutes. For single gel procedures the transfer apparatus was assembled (Figure 2.2), taking care to avoid any air bubbles and ensuring all blotting paper and sponges were well soaked in buffer prior to assemble. For multi-strip blotting, gel slices, corresponding to the region where the protein of interest had migrated to, were excised from each gel and recombined in the transfer apparatus to ensure identical transfer efficiency, and subsequent antibody binding and signal detection across a large number of samples. Typically 36-45 samples can be analysed in parallel, dramatically decreasing the standard error of technical replicates (Kiyatkin, et al. 2009). Protein was transferred
at 380mA for 65 minutes in the cold room, following which the membrane was removed and rinsed in TBS-T.

**Figure 2.2: Cartoon depicting the western blot transfer apparatus.**

2.2.3.6.3: Membrane blocking, probing and washing

Membranes were blocked with 5% non-fat powered milk in TBST at room temperature for 1 hour, and subsequently washed in TBST for 30 minutes. Primary antibodies incubation were preformed as advised by the manufacturer, generally overnight at 6°C in TBST supplemented with either 5% BSA (Fraction V) or 5% non-fat powered milk. Following primary antibody incubations the membrane was washed 3 times for 15 minutes each with TBST at room temperature and then incubated with the appropriate secondary HRP conjugated antibody (Promega) at 1:10000 in 5% non-fat powered milk in TBST at room temperature for 1 hour. The membrane was then again washed 3 times for 15 minutes each with TBST at room temperature.
2.2.3.6: Signal detection

The antibody signal was detected using the ECL reagent kit SuperSignal™ West Pico/Femto Chemiluminescent Substrate (Pierce). Equal volumes of each reagent were mixed and incubated with the membrane in a 50ml Corning tube for approximately 1 minute. The membrane was then removed, wrapped in cling film and exposed to autoradiographic film for the minimum time needed to detect a signal, typically between 1-30 seconds. Films were developed using an automated autoradiograph developer. Films were scanned and where necessary analysed with ImageJ64 software.

2.2.3.7: Immunoprecipitation of tagged proteins

All manipulations were preformed in the cold room. All buffers were chilled on ice and supplemented with both protease and phosphatase inhibitors directly before use. All centrifugation steps were preformed at 4°C. When preforming immunoprecipitates for LC/MS-MS analysis all reagents were sterile filtered through a .22µm syringe driven filter device. Aero-guard pipette tips were used as well as low protein-binding microfuge tubes (Star Labs). All pipettes were cleaned with 70% EtOH before and during use. Nitrile gloves were used at all times. To increase the efficiency of multi-protein complex capture, all steps from lysate preparation to elution were preformed in a single day.

2.2.3.7.1: Sample preparation

Cells were washed twice with ice cold PBS and then scraped from the culture dish with a soft rubber spatula. The cell suspension was collected by centrifugation at 4000rpm and washed again with PBS. Cells were lysed in an appropriate volume of IP Lysis buffer, and solubilized on ice for 30 minutes, being passed through a 18 gauge needle three times every 15 minutes. 2µl of Benzonase nuclease was added to each sample and gently mixed for 10 minutes. Insoluble material was removed by centrifugation at 14000rpm for 30 minutes. The soluble fraction was pre-cleared with between 20-100µl IGG agarose beads (Sigma) for 30
minutes with gentle rolling. Beads were removed by centrifugation for 2 minutes at 1200rpm and 30µl of the supernatant was retained as the input fraction and for protein quantitation estimation.

2.2.3.7.2: Protein capture

The supernatant from section 2.2.3.7.1 was incubated with equilibrated antibody-bound beads. 2-3µl of bead slurry per 15cm dish was used, depending on the conjugated antibody. The protein/bead mix was incubated for 3 hours with end-over-end rolling at 6°C. Bead/protein complexes were collected by centrifugation at 1200rpm, and 30µl of the supernatant was retained as the flow-through sample. Beads were washed three time with 1ml of lysis buffer for 30 seconds each and eluted with either peptide (FLAG) or glycine (GBP), three times for 15 minutes at 6°C. The eluates were combined and concentrated using a 3kDa Amicon Ultra-0.5 ml Centrifugal Filters for Protein Purification and Concentration (Millipore).

2.2.3.8: Sample preparation for Mass Spectrometry

Wherever possible manipulations were preformed in a laminar flow hood. All reagents used were HPLC grade. Equipment was washed in ethanol before use. Glassware was washed once in detergent and again in NaOH. Non-latex powder-free gloves were used for all manipulations.

2.2.3.8.1: Gel-free digestion

Eluates (~pH 8) were reduced in 20mM DTT and then alkylated in 10mM iodoacetamide at room temperature in the dark for 30 min. Protein solution was diluted with 100µl of 200mM HEPES (pH 8.0) and trypsin (modified, sequencing grade) was added to the sample at a ratio of 1:50 enzyme/protein. Trypsin digestion was performed overnight at 37°C. After tryptic digestion, formic acid was added to 2% final concentration to stop the reaction, and incubated at 37°C for additional 4 hours. Samples were centrifuged for 30 min at 10000g to remove insoluble material.
and filtered with a 0.2µm cartridge. Filtered peptides were dried using a SpeedVac, and stored at −20°C.

2.2.3.8.2: In-gel digestion

Samples were prepared and separated by SDS-PAGE electrophoresis as described in section 2.2.3.3, and subsequently coomassie stained (section 2.2.3.4). The gel was destained overnight in water and the following day each lane was cut into 5 equal sized pieces with a sterile scalpel. Each gel piece was then cut into a further 6 pieces, taking care to avoid mincing the gel, and transferred into a polypropylene tube. Gel pieces were subjected to three rounds of dehydration and rehydration with MeOH and 50mM NH₄HCO₃ solutions, respectively, for 10 minutes per step, finishing with a dehydration step. The gel pieces were then rehydrated with a solution of 20mM DTT in 50mM NH₄HCO₃ and then incubated at 60°C for 30 minutes to reduce the protein samples. The DTT solution was then removed and the gel pieces were again dehydrated in MeOH for 10 minutes. Protein samples were then dealkylated in a solution of 50mM iodoacetamide in 50mM NH₄HCO₃ and incubated at room temperature for 1 hour in the dark. Gel pieces were again dehydrated in MeOH for 10 minutes and subsequently dried for 15 minutes using a SpeedVac.

20µl trypsin (20µg/ml) in 50mM NH₄HCO₃ was then added and incubated at 4°C until completely absorbed. Enough 50mM NH₄HCO₃ was then added to cover the gel pieces completely and the samples were incubated overnight at room temperature. The following day the pH was adjusted to 8 and the digested solution was transferred to a fresh tube. 50µl of a 1% formic acid in 50% MeOH was added to the gel pieces, incubated at room temperature for 15 minutes, and then added to the digested solution from the previous. This step was repeated again. Finally, the gel pieces were again dehydrated with MeOH for 10 minutes and the extracted solution was combined with the peptide/formic acid solution from the previous steps. This mixture was then dried in a_speedvac and stored at -80°C for LC/MS-MS analysis preformed by Dr. Thierry Le Bihan at the CSBE, University of Edinburgh using a
1200 series micro-WPS autosampler and a binary micro-pump coupled to an LTQ-Orbitrap hybrid mass spectrometer.

**2.2.3.9: Size exclusion chromatography**

Lysates were prepared as described in section 2.2.3.7.1, except that once the insoluble material was removed the supernatant was filtered through a microfuge based filtration device (0.44µm pore size). 500µg of total protein, in a volume of 500µl, was injected into an ÄKTA Explorer (GE) and passed over a Superdex 200 column (Tricorn). Fractions were eluted with 20mM HEPES, 250mM NaCl pH 7.6, supplemented with 2mM NaVO₄, at a flow-rate of 0.5ml/min. Fractions were collected in 250µl volumes and stored at -80°C for processing. Size exclusion chromatography was performed with Dr. Martin Wear and Dr. Liz Blackburn at the Edinburgh Protein Production Facility.

**2.2.3.10: TCA precipitation of proteins**

Protein samples were mixed with 100% (w/v) TCA at a ratio of 4:1 and incubated at -80°C overnight. The following day the protein was pelleted by centrifugation at 14000rpm for 30 minutes at 4°C. The protein pellets were washed twice with 1ml of ice cold acetone and subsequently allowed to air-dry. Protein pellets were resuspended in 30ml of 1x Laemmli buffer and stored at -80°C.

**2.2.4: Tissue Culture**

Cells were routinely grown at 37°C at 7.5% CO₂ in a humidified incubator. All solutions were sterile and all cell manipulations were performed in a sterile laminar flow hood. Solutions were stored at 4°C and not used for longer than one month. Prior to use, culture medium was warmed to 37°C in a water bath. For mES cell culture, sterile plastic culture dishes (Iwaki) were coated with 0.1% porcine gelatine in PBS before use. For 2i culture, culture dishes were coated either in 0.1% porcine gelatine in PBS or poly-L ornithine (Sigma) overnight followed by 1µg/ml laminin (Sigma) in PBS for 2 hours, depending on the cell line being cultured. All
stable cell lines were cultured in the absence of penicillin/streptomycin. Newly derived cell lines were kept in quarantine until they tested negative for mycoplasma and bacterial contamination. All newly derived lines and clonal lines were karyotyped before used in experiments. Antibiotic concentrations for selection were determined empirically by dose curve, and a batch to batch basis.

2.2.4.1: Routine culture of mES cells

2.2.4.1.1: Culture of 129 derived mES cells

Cells were routine cultured until they reached 80% confluence whereupon the culture medium was aspirated off and the cell monolayer was rinsed with room temperature Dulbeco’s PBS (Gibco). 1ml of trypsin at a concentration of 0.125% (w/v) in PBS was added per T25 flask, and scaled accordingly for other sized flasks. The flask was incubated at 37°C for less that 2 minutes and then tapped 3 times to disassociate the cells. 5ml of complete culture medium was then added and the cell suspension was pipetted 10 times using a 10ml glass pipette to generate a single cell suspension. 1ml of the cell suspension was added to a freshly prepared flask containing complete media and returned to the incubator. Cells were passaged accordingly every second day, and fed with 8ml of complete medium every day.

2.2.4.1.1: Culture of C57Bl/7-129 derived mES cells

Cells were cultured using a modification of the 3i culture procedure. Briefly, complete mES cell media supplemented with 3µM CHIR99021 and 1µM PD0325901 (Axon Medchem). Cells were grown to an apparent 40% confluence and than passaged as described in section 2.2.4.1.1, except that precisely 10^6 cells were plated into a new flask. Cells were passaged every 2-3 days and fed every day.

2.2.4.2: Transfection of mES cells

2.2.4.2.1: Transfection by liposomes

Transient, and sometimes stable, transfection was preformed using either Lipofectamine 2000 or Lipofectamine LTX (Invitrogen) at the DNA:reagent ratios
recommended by the manufacturer. 10^6 cells were plated per 6 well dish 1 hour before transfection, in complete medium. During this time 3µl of transfection reagent and 3µg of DNA were each suspended in 250µl mES cell media lacking FCS and incubated at room temperature for 15 minutes in polystyrene tubes. Care was taken to ensure the DNA used for transfection was of a high quality and contained low salt concentrations. For transfection controls either DNA was replaced with either water or with a plasmid expressing eGFP (AGS 684). Both DNA and Lipofection reagent mixtures were then mixed together and incubated for a further 45 minutes. Transfection mixture were added to the cell suspension in a dropwise manner and contact with the culture dish was avoided at all times. Cells were returned to the incubator and the media was changed to complete mES cell media after ~8 hours. The next day cells were passaged for down stream applications. Routine transfection efficiency of between 40-60% was achieved using this method.

2.2.4.2.2: Transfection by electroporation

Plasmid DNA was digested for a minimum of 3 hours with the appropriate enzyme. Digestion efficiency was assessed by agarose gel electrophoresis. Once complete digestion was achieved the DNA solution was salt precipitated and washed 3 times with 70% EtOH before allowing to air-dry in a laminar flow hood. The DNA pellet was resuspended in 200µl PBS. For large scale electroporations 100µg of DNA was used, for small scale 25µg of DNA was used. Cells were collected as usual and counted using haemocytometer. Cells were collected by centrifugation at 1300xg and washed once with 10ml PBS. For large scale electroporations 10^7 cells were used, for small scale 5x 10^6 cells were used. Cell pellets were resuspended in 600µl PBS, mixed with the DNA mixture, and allowed to stand for 10 minutes. The mixture was transferred to a 4mm electroporation cuvette (Biorad). For large scale electroporations cells were electroporated at 0.8kV and 10µF giving a time constant of 0.1. For small scale electroporations cells were electroporated at 0.25kV and 500µF giving a time constant of ~0.6. Cells were allowed to rest for 10 minutes and then plated in complete media at 10^6 cells per 10cm dish. The cells were returned to
the incubator and the following day antibiotic selection was started. Cultures were selected from until the emergence of discrete medium sized colonies which were then picked into a 96 well plate.

2.2.4.3: Picking mES cell colonies

20µl of trypsin solution was added to each well of a gelatinized 96 well plate. Medium was aspirated from the culture dish and cells were washed twice with 5ml PBS. 2ml PBS was left to cover the colonies. Approximately 10µl of trypsin solution was collected using a pipette, fitted with a wide bore p20 tip, in reverse operating mode. This solution was used to bathe an individual colony so that it could be easily detached from the plate. Once picked the colony was placed into the well of the 96 well plate and incubated at room temperature for 5 minutes, followed by vigorous pipetting, to dissociate the colony. 200µl of selection medium was then added to the well. 12 colonies were picked at a time in this manner. When the cells reached 80% confluence they were either passaged to 48 well dish, replica-plated and frozen, or directly processed for the desired application.

2.2.4.4: Freezing mES cells

For routine freezing of mES cells a freezing solution of complete medium containing 10% DMSO (v/v) was prepared. Cells were harvested as usual, collected by centrifugation, and resuspended in freezing medium at an approximate concentration of 2x10⁶ cells/ml. 0.5ml of this cell suspension was transferred to a cryovial (Nunc). The tube was labelled with the cell line, passage number and date and stored at -80°C overnight. The following day the vial was transferred to the liquid nitrogen cell bank. Cultures in 96 well plates were grown to 90% confluence and fed with complete media 2 hours prior to harvesting. Cells were washed with 100µl PBS and then disassociated in 30µl trypsin solution for 3 minutes. A single cell suspension was achieved by pipetting and the reaction was stopped by addition of 70µl of complete medium containing 50% FCS (v/v). 50µl of this cell suspension was then added the each corresponding well of a gelatinised replica plate, containing
200µl of selection media in each well and the plate returned to the incubator. To the remaining cell suspension 50µl of complete medium containing 20% DMSO (v/v) was added and mixed vigorously. The plate was covered, sealed with Parafilm, wrapped in aluminium foil and stored at -80 until required.

2.2.4.5: Thawing mES cells

Cells frozen in cryovials were retrieved from the liquid nitrogen cell bank and quickly thawed at 37°C. The cell suspension was then added to ~7ml of complete medium, pre-warmed to 37°C. The cells were collected by centrifugation, resuspended in 8ml of complete medium and transferred to a fresh T25 flask. The media was changed once the cells had attached, usually 5 hours. For cells frozen in 96 well plate format, the entire plate was thawed at 37°C. 100µl of cell suspension diluted in 1ml of complete medium and transferred to a gelatinised well of a 46 well plate. The media was changed once the cells had attached, usually 10 hours.

2.2.4.6: Karyotyping

In general chromosomal counting was performed by Jonathan Rans of the ISCR Tissue Culture Core Facility. Briefly, metaphase spreads were prepared by adding mES cell medium supplemented with colcemid (10µg/ml) to exponentially growing at ~70% confluence. Cultures were returned to the incubated for approximately 1 hour after which time metaphase arrested cells were collected and incubated in 2ml hypotonic solution (0.4% KCl, 0.4% Sodium Citrate) for a further 5 minutes. Cells were fixed in 2ml fixation buffer (3:1 MeOH:Acetic acid) at RT for 30 minutes. Cells were dropped onto a glass slide, flooded with fixative and allowed to dry before staining with Giemsa and chromosome counting.

2.2.4.7: hLIF05 expression in COS7 cells

hLif05 was expressed in COS7 cells, secreted, and collected as a supernatant. Briefly, COS7 cells were cultured on plastic culture dishes in mES cell media, without Lif. The day before transfection 3x10^6 cells were plated in a 10cm
dish. 16 hours later, plasmid DNA encoding a hLIF05 expression cassette AGS 627 was packaged into liposomes as described in section 2.2.4.2.1, except that 10 times the amount of DNA and Lipofectamine was used. After 6 hours the media was changed, and subsequently refreshed and collected every subsequent 24 hours, for 3 days. The collected supernatant was filtered through a 0.44µm filter and assayed for the ability to block LIF signaling at 0.01x LIF concentration. LIF activity was determined by SOCS3 expression. The minimum concentration that efficiently blocked SOCS3 expression under these conditions was used for further experiments. Supernatants were aliquoted and stored at -80 °C.

2.2.4.8: Acute stimulation of mES cells

Due to the autocrine expression of both FGF4 and LIF in mES cells, prior to stimulation cultures were treated with the FGFR inhibitor PD173074 and the LIF antagonist hLIF05. Briefly, 2.5x10^5 were plated in a 6 well dish 49 hours before stimulation. The following day media was changed to mES cell media, with or without LIF, and containing PD173074 (100nM) or hLIF05 (concentration determined empirically), or both, depending on the experiment. The following day the media was aspirated and the cell monolayer was washed 3 times with 2ml PBS and stimulated with 2ml mES cell media supplemented with either 10ng/ml recombinant FGF2 (R&D Systems), or 1x LIF (produced in-house as a supernatant), or both, depending on the experiment. Cells were returned to the incubator for the desired time and then either total protein or total RNA was harvested, as required.

2.2.4.9: Differentiation of mES cells

2.2.4.9.1: Neural monolayer differentiation

Neural induction was preformed under monolayer conditions as outlined in Ying et al. 2003, with some exceptions. The initial number of cells plated was determined empirically for each cell line. For most cell lines 1.5 x 10^4/cm^2 were plated in N2B27. In this instance the day of plating was considered day 0 of the experiment. For embryo derived Erk2^-/- mES 1.5 x 10^4/cm^2 cells were plated in
N2B27 supplemented with Lif (Kunath et al. 2007). The following day media was changed to N2B27 only. In this instance the day of plating was considered day -1 of the experiment. Cultures were fed every day with 2ml of N2B27. Cultures were assayed for neural induction between days 5 and 10. This assay generally generates neurectodermal cell types found in the developing forebrain.

2.2.4.9.2: Non-neural surface ectoderm differentiation

Non-neural surface ectoderm differentiation was performed as outlined in section 2.2.4.8.1, except that the media was supplemented with recombinant human BMP4 (R&D Systems) at a concentration of 10ng/ml. Cultures were assayed for non-neural surface ectoderm differentiation at day 6.

2.2.4.9.3: Multilineage differentiation in hanging drops

mES cells were differentiated into a mixture of mesoderm, endoderm, and ectoderm lineages in hanging drops (Wang and Yang, 2008). Cells were collected as usual and resuspended in mES cell medium without Lif at a concentration of $3.2 \times 10^3$ cells/ml. Using a multi-channel pipette, 30µl drops were placed on the inside of the lid of a 15cm dish, approximately 100 drops/lid. 10ml PBS was added to the culture dish and the lid was inverted over it and returned to the incubator for 48 hours. At day 2 the embryoid bodies were collected and resuspended in 4ml medium, again without Lif, and plated in sterile, bacterial grade 4cm dishes. Cultures were fed every 2-3 days and harvested between day 6-12.

2.2.4.9.4: Mesoderm monolayer differentiation

Mesoderm monolayer was performed according to Nishikawa, et al. (1998). Briefly mES cell were seeded in collagen type IV plates at a density of $10^3$ cells/cm² in 2ml mesoderm differentiation media. Mesoderm induction was analysed on day 4.
2.2.4.9.5: Clonal assay

mES cells were trypsinized to obtain a single cell suspension and plated at a density of 0.2x10³ cells/cm² in the presence or absence of Lif. After 6/7 days, plates were stained for alkaline phosphatase activity and colonies scored in three categories: undifferentiated, mixed, and fully differentiated.

2.2.5: Histological techniques

2.2.5.1: Alkaline phosphatase staining

Alkaline phosphatase staining was performed using the Leukocyte Alkaline Phosphatase Assay Kit (Sigma) as outlined in the product manual. Cells were washed in PBS and then fixed in a solution of acetone, citrate solution, and formaldehyde (37% v/v) (Sigma) at a ratio of 8.125 : 2.75 : 1. The fix was removed after 1 minute and the cells were washed with 3 ml PBS. Enough stain (ddH₂O, sodium nitrate, FRV alkaline solution, and Napthol AS-B1 at a ratio of 45:1:1:1) was added to cover each well, and samples were incubated in the dark at room temperature for approximately 15 minutes. The stain was aspirated and the wells washed twice with PBS.

2.2.5.2: Immunocytochemistry

mES cells were cultured in standard culture dishes and prior to fixation wells were washed twice with PBS. Cells were fixed in 4% paraformaldehyde (w/v) at room temperature for 15 minutes. The fixation was aspirated and the wells washed 3 times, for ten minutes per wash, in PBS. Permeabilization and blocking was performed in a single step by incubation for 1 hour at room temperature with a solution of 0.1% Triton-X and 2% serum in PBS. Primary antibodies were incubated as recommended by the manufacturer, for either 1 hour or overnight. Primary antibody was then aspirated and the cells washed 3 times for 30 minutes in PBS. The appropriate Alexa Fluor™secondary antibody (Invitrogen) was diluted in blocking/permeabilization and incubated with the sample at room temperature for 1 hour. The
antibody solution was aspirated and washed extensively with PBS. DNA was stained with DAPI at a concentration of 1µg/ml and immunofluorescence was visualized using an Olympus IX51 fluorescent microscope and analysed using the Velocity software package (http://www.perkinelmer.com/pages/020/cellularimaging/products/velocitoxhtml).

### 2.2.5.3 Flow Cytometry

Cells were collected by trypsinization as usual, washed once in mES cell medium, and then once in PBS. For analysis of eGFP/Venus expression, cells were resuspended in FACS Buffer (3%FCS (v/v) in PBS) and analysed immediately. For propidium iodide staining cells (~10⁶) were resuspended in 100µl PBS following trypsinization and 1ml ice-cold 70% EtOH was added, dropwise with gentle mixing. Fixed cells were collected by centrifugation and rehydrated with 2 PBS washes, for 10 minutes each. Cells were incubated with RnaseA at a concentration of 1mg/ml and incubated at room temperature for 20 minutes. Propidium iodide was then added to a final concentration of 1µg/ml and DNA content was visualized using a BD FACSCalibur flow cytometer. Data was analysed using FLOWJO flow cytometry analysis software (http://www.flowjo.com/index.php).

### 2.2.6: Mouse and embryo techniques

#### 2.2.6.1: Animal housing

Mice were housed and bred within the University of Edinburgh, according to the Animals (Scientific Procedures) Act (UK) 1986. Litters from matings were left with the mother until weaning at 3 weeks. At this time ear biopsies were taken for genotyping and the sex was determined. When the mice were 6 weeks and older they were set up for mating. The Biomedical Unit Staff of the Institute for Stem Cell Research performed all ear biopsies, mating setups, and maintenance of the mouse colony.
2.2.6.2: mES cell derivation

mES cell lines were derived by Jan Ure at the ISCR Transgenics Unit as outlined in Ying et al (2008) but with some small exceptions. Briefly, embryos were collected at the morula stage by flushing the infundibulum with M2 medium. Morulas were cultured overnight in 2i medium and the following day each single morula was transferred to a laminin coated well of a 4 well dish, allowed to hatch from the zona pellucida, and attach to the dish. Outgrowths were monitored daily and when they reached a medium size they were disassociated to a single cell suspension in Accutase and transferred to a laminin coated well of a 12 well dish. Stable lines were weaned onto serum containing medium with Lif, and subsequently weaned off 2i supplement, and transferred to gelatin coated plates. 1-2 vials of cell were frozen before each change in culture conditions.
Chapter 3 - Analysis of ERK2 regulation in mES cells

3.1 Introduction

The Mapk pathway is a crucial mediator of many cellular processes. Although noise and stochasticity are unavoidable aspects of many cellular processes, in general the phenotypic outcome of a signaling event is consistent, and reproducible. Recent reports have indicated that the cellular concentration of many proteins show a broad distribution, with a standard deviation of 20-30% of the mean (Sigal et al., 2006). This phenomenon has also been observed in the MAPK pathway, with a similar standard deviation of the mean (Cohen-Saidon et al., 2009). Concentrations of nuclear ERK2 show high variation in basal levels, and following acute stimulation, with nuclear accumulation exhibiting fold change dynamics. Moreover, the difference in Erk2 protein concentration from one cell to another can be as great as threefold. These observations have profound implications on our understanding of the regulatory mechanisms underpinning MAPK signaling given that in some cell types a reduction in ERK dose of ~70% can dramatically alter cellular responses to external cues (Lefloch et al., 2008). This being said, the phenotypic outcome of MAPK signaling is highly homogenous. In the classical PC12 model system, NGF stimulation results in ~75% of all cell adopting a neuronal fate (Santos et al., 2007).

One mechanism by which robustness to perturbations in ERK dose is achieved is through ERK mediated negative feedback to upstream signaling components (Santos et al., 2007, Schilling et al., 2009, Fritsche-Guenther et al., 2011), thus allowing compensatory phosphorylation of one ERK isozyme in response to changes in the levels of the other. This model assumes a certain degree of redundancy between ERK isozymes, and that the outcome of a signaling event is a function of the total phospho-ERK output. This is still however a much debated phenomenon and may be cell type specific (Indigro et al., 2010, Vantaggiato et al., 2006).
3.1.2 Hypothesis: The differentiation defect in ERK2 deficient mES cells may be due to impaired or absent negative feedback control.

At present the Erk2 knockout phenotype in mES cells has been analysed in 2 different studies, one in E14tg2a (129/Ola) cells where Erk2 was genetically depleted by 2 rounds of homologous recombination (Yao et al., 2003), and the other in mES cells derived from intercrosses of Erk2+/− C57Bl/6/129 F1 mice (Kunath et al., 2007). Yao et al. noted slight compensatory phosphorylation of ERK1 in ERK2 deficient cells, and observed no block in mesoderm commitment under monolayer differentiation conditions. On the other hand embryo-derived Erk2 mutant cells exhibited no compensatory phosphorylation of ERK1 (Kunath et al., 2007: Figure 4-A, and Figure 3.2.1-A of this chapter), and failed to generate mesoderm tissue under monolayer conditions. As ERK-dependent feedback control appears to be active in 129/Ola mES cells (Williams et al., 2000), it is possible that the differentiation defect observed by Kunath et al. may be due to genetic-background specific differences in the architecture of signaling components upstream of ERK, such as differential expression levels/involvement for RAF isozymes that might affect feedback control mechanisms (Fritsche-Guenther et al., 2011, Pritchard et al., 1996), thus preventing compensatory ERK1 activation in response to ERK2 depletion.

3.1.3 Aims of this chapter

The aim of this chapter was to address the discrepancy of the Erk2 knockout phenotype in mES cells and in doing so generate and validate a cell line expressing a FLAG-tagged ERK2 for downstream biochemical studies. The rational behind the approach was that if the differentiation defect observed by Kunath et al. was due to a lack of compensatory phosphorylation of ERK1, as noted by Yao and colleagues, transgenically expressing ERK2 should rescue this differentiation defect by increasing total ERK dose, and at the same time generate and validate a cell line suitable for IP/MS/MS analysis of ERK2 interacting proteins.
3.2 Results

3.2.1: Transgenic rescue of ERK2 deficiency

The lack of compensatory phosphorylation of ERK1 in B1-Erk2 mutant lines was confirmed by western blotting (Figure 3.1-A). Transgenic rescue of ERK2 deficiency was achieved through random integration of an Erk2 expression cassette with a 3xFLAG tag fused to the N-terminal of ERK2 allowing subsequent use in biochemical studies. The Erk2 coding region was amplified from cDNA derived from E14tg2a total-RNA with the primers ‘BamHI Erk2 F’ and ‘Erk2 NotI R’, that contained BamHI and NotI sites, respectively. A poly-glycine linker was included at the 3’ end of the forward primer to aid the accessibility of the FLAG epitope. The resulting PCR fragment was TOPO cloned and sequence verified. The insert was then cloned into pPy CAG IP (IPC37) (Figure A1) as a BamHI/NotI fragment, by 3-way ligation, between the FLAG and IRES sequences, and confirmed by restriction digestion.

ScaI-linearised plasmid (25µg) was then electroporated into Erk2−/− B1 mES cells, and transfectants were selected for with puromycin (1µg) for 8 days. Clones were picked, expanded, and screened for transgene expression. All clones tested expressed 3xFLAG-tagged ERK2, which was phosphorylated normally (Figure 3.1-B). Because both tagged ERK2 and endogenous ERK1 migrated with an identical apparent molecular mass (44kDa) it was difficult to quantify the exact extent of overexpression in the rescued clones by western blot. However, it was evident that the levels of exogenous ppERK were not excessive, and between 2 to 3 fold endogenous. To functionally test tagged protein, rescued clones were subjected to acute FGF2 stimulation for 1 hour and the induction of the IEG Egr1 was determined by Q-PCR (Figure 3.1-D) (see Appendix 1.1 and Methods 2.2.4.7 for an explanation of the experimental approach used). Egr1 induction was evident in WT mES cells at 1 hour, approximately 4 fold over basal levels, where induction at this time point was all but absent in the Erk2−/− cell line. This aberrant gene induction was rescued by transgenic Erk2 expression indicating that Egr1 expression is largely
ERK2 dependent in this cell line, and that neither the 3xFLAG tag, nor the increase in protein expression impeded Erk2 function.

The differentiation potential of these lines was then assessed by neural induction (Figure 3.2). Neural induction was performed as described in Kunath et al. (2007), and the extent of induction was assayed by both Q-PCR and immunocytochemistry for markers of neurectodermal lineages. Whilst WT mES cells generated morphologically distinct, β-III Tubulin positive neurons, up-regulated the pan-neural marker Nestin, and down-regulated the pluripotency marker OCT4, neither the Erk2−/− cell line, nor the rescued clones, showed any signs of neural induction, and maintained Oct4 expression. This indicated that the blockage in neural commitment was not due to the targeted disruption of the Erk2 allele.

As it has been observed within our Institute that some cell lines do not differentiate efficiently under neural monolayer conditions, the differentiation potential of the B1 Erk2−/− cell line was tested by embryoid body differentiation (Figure 3.3). After 6 days of differentiation WT, ERK2 deficient, and rescue cell lines had all formed well defined, partially cystic embryoid bodies (Figure 3.3-A [arrowheads]). Marker gene analysis showed down-regulation of the pluripotency markers Nanog, Oct4 and Stella (Figure 3.3-B), as well as up-regulation of CCND2 (proliferative, non-ES cells), Gata6 (endoderm) and PDGFR1α (mesoderm/PE) (Figure 3.3-C), indicating in vitro differentiation under these conditions was Erk2 independent.
Figure 3.1: Transgenic rescue of ERK2 deficiency. A) ppERK expression in two independent Erk2 null cell lines: B1 and B3. Gels were coomassie stained following transfer to estimate equal loading. B) ppERK levels in ERK2 over-expressing B1 (Erk2 null) clones: lane 1 top blot: upper and lower bands correspond to endogenous ERK1 and ERK2, respectively. Lanes 2-4 correspond to the overlapping signals from endogenous ppERK1 and exogenous tagged ppERK2. Gapdh was used as a loading control, bottom blot. C) Q-PCR analysis of Egr1 induction 1 hour after acute stimulation with recombinant FGF2 (10ng/ml). Values are normalized to TATA binding protein (TBP). Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments.
Figure 3.2: The block in neural induction cannot be transgenically rescued. A) Immunocytochemistry analysis of cultures following 7 days of neural induction for neural marker βIII Tubulin immunoreactivity. DNA is counterstained with DAPI. Images are taken at 20x magnification. B) Q-PCR analysis of markers of pluripotency (Oct4), and neural tissue (Nestin). Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments.
Figure 3.3: Erk2 deficient cells can differentiate embryoid bodies. A) Phase contrast images embryoid bodies following 6 days of differentiation, arrowheads indicate presumptive cysts. Images are taken at 20x magnification. B) Q-PCR analysis of the embryoid bodies in A for markers of pluripotency. C) As in B except for markers of endodermal and mesodermal lineages.
3.2.2: Genetic depletion of Erk2 in 129/Ola mES cells

The initial work with the B1-Erk2\(^{-/-}\) cell line indicated that the reported phenotype (Kunath et al., 2007) could not be attributed to Erk2 depletion, nor a lack of compensatory Erk1 activation. As the B1-Erk2\(^{-/-}\) line was derived from intercrosses of C57Bl/6/129 F1 Erk2\(^{+/−}\) mice the differentiation defect seen under monolayer culture conditions may be due to either genetic background, which is known to have a significant effect on the overall biology of mES cell lines (Ying et al., 2008), or possibly selection for certain traits, such as enhanced self-renewal, during the derivation process. As part of the overall aim of this project was to generate an ERK2 interaction network it was felt that to do so in a cell line that exhibited abnormal mES cell biology would bias any results from such studies.

Therefore, 129/Ola derived, germ-line competent mES cell were depleted for exon 3 of the Erk2 allele by two round of homologous recombination. This strategy has been used before to disrupt the Erk2 allele (Fischer et al., 2005, Saba-El-Leil et al., 2003) and disrupts the regions coding for kinase sub-domains V and VI, essential for ERK2 function. Moreover, removal of exon 3 will generate a frame shift within exon 4 a sufficient distance from the downstream exon junction complex to trigger nonsense-mediated decay of the transcript (Chang et al., 2007).

Figure 3.4 illustrates the targeting strategy and both targeting vectors cloned; one containing a floxed Hygromycin-Thymidine Kinase selection cassette: pETVH, and the other containing a Frt flanked Neomycin cassette: pETVN. pETVH was cloned by digesting pETV (Figure A5), a gift from S.Meloche, (Saba-El-Leil et al., 2003), with BglII to remove a 2.5kb fragment flanking exon 3. The vector backbone, containing 4.8kb 5’ homology and 2.8kb 3’ homology, was then ligated to the BamHI fragment from AGS 703 (Figure A6) (Vives et al., 2008), that contained the floxed CMV-HyTK cassette to generate pETVH(-)DTA. A negative selection cassette containing mPgk driven diphtheria toxin-A was taken from pRosa26-1 (Figure A7) as a HincII fragment that also contained the Amp\(^r\) cassette and ColEi ori. This fragment was then inserted into the targeting vector between the PvuII sites.
flanking the pBS backbone to generate the final targeting vector, which was confirmed by restriction fragment analysis (Figure A8).

pETVN was cloned by digesting the transition vector pETVH(-)DTA with BamHI to remove the CMV-HyTK cassette. The Frt flanked mPgk-Em7-Neo cassette from pL451 (Figure A9) was retrieved as a BamHI fragment and ligated to the targeting backbone. The resulting plasmid was then digested with NsiI and the 5’ fragment was ligated to the complimentary 3’ NsiI fragment from pETVH to add the DTA cassette and verified by restriction fragment analysis (Figure A10).

![Restriction enzyme map](image)

**Figure 3.4**: Cartoon illustrating the WT Erk2 locus, and both modified loci. 5’ and 3’ homology arms are shown as blue bars on the WT allele. Black bar indicates the location of the 3' probe. Positions and sizes of the WT and mutated alleles are as shown.

XmnI-linearised pETVH (100µg) was then electroporated into the WT Ju09 mES cells, and transfectants were selected with Hygromycin B (100µg/ml) yielding over $10^3$ colonies by day 10 of selection. Approximately 300 clones were picked and expanded to 96 well format whereupon replica plates were made and frozen at -80°C. Genomic DNA was harvested from the remaining cells and used to screen for exon 3 copy number (Figure 3.5). Q-PCR analysis detected 6 clones that showed a decrease in exon 3 copy number, giving a putative targeting efficiency of 22%. These 6 clones were then selected for further analysis. The genotype was then confirmed by Southern blot for all 6 clones (Figure 3.6). Correct targeting introduced an exogenous KpnI site into the region of exon 3 replaced by the selection cassette,
which allowed both WT and targeted alleles to be distinguished from each other. A 3’ external probe identified 2 bands, a 10kb and 7kb, corresponding to WT and targeted alleles, respectively. All clones were karyotyped and exhibited normal ploidy. Clones D2 and D7 (now referred to as JuD2 and JuD7, respectively) were selected for further manipulation.

XmnI-linearised pETVN (100µg) was then electroporated into each of the JuD2 and JuD7 clones and transfectants were selected with both Hygromycin B (100µg/ml) and G418 (150µg/ml) for 9 days yielding approximately 0.3x10^3 clones per electroporation. Clones were expanded as before and genomic DNA harvested after replica plate freezing and analysed for the presence of exon 3. Q-PCR analysis, (Figure 3.7-A), identified 4 clones that showed no signal for exon 3 indicating putative homozygous knockout lines. Introduction of a PstI site in both Neo and Hygro-TK cassettes allowed all three alleles to be distinguished by KpnI-PstI double digestion and Southern blotting. An external 3’ probe identified 3 bands, a 10kb, a 6kb and a 5kb, corresponding to the WT, hygro, and neo mutated alleles, respectively (Figure 3.7-B). All clones were karyotyped and displayed normal ploidy. Two null clones, D7N2 and D2N4, each derived from independent heterozygous clones, JuD7 and JuD2, respectively, were selected for further analysis.

To test whether the targeted mutated generated the predicted nonfunctional allele, Erk2 expression was assessed by Q-PCR and western blot (Figure 3.8). A 50% reduction of Erk2 message was evident in the heterozygous line JuD2, as well as a near absence of any message from both homozygous null lines. Erk2 expression followed a linear relationship with copy number indicating, in this instance at least, that Erk2 promoter activity is not regulated by ERK2 dose. Also, there was no evidence of a compensatory increase in Erk1 expression. This was also confirmed on the protein level by immunoblotting with a polyclonal antibody that recognizes both ERK1 and ERK2. Moreover, using a similar targeting strategy Saba-El-Leil et al. (2003) failed to detect any expression of a truncated ERK2 protein using an antibody that specifically recognizes the C-terminal of ERK1 and ERK2.
Both Erk2−/− lines exhibited similar morphological features to those seen in the parental WT line (Figure 3.9), with sub-populations of tight, light-refractory, self-renewing colonies (red arrowheads) and flattened differentiated cells (black arrowheads). This is in sharp contrast to the embryo-derived B1 Erk2−/− line in which there was scant evidence of spontaneous differentiation. Moreover, the embryo-derived mutant line exhibit an atypical morphology, being flattened, and slightly oblong in shape (blue arrowheads).
Figure 3.5: Q-PCR analysis for Erk2 exon 3 copy. Erk2 Ct values are normalized to those of a non-targeted locus (Spry1 intron 1) and then normalized to the same values from the parental cell line. Analysis reveals 6 of 27 clones tested show a reduction in a single allele of Erk2 exon 3.

Figure 3.6: Southern blot analysis of parental and Erk2 +/- clones. DNA was digested with KpnI and hybridized with a 3’ external probe. Positions and sizes of WT and mutated DNA is shown.
Figure 3.7: Genotyping of parental and Erk2⁻/⁻ clones. A) PCR Analysis comparing exon 3 dose across WT (Ju09), Erk2⁺/⁺ (JuD2), and Erk2⁻/⁻ (D7N1, D7N2, D2N3, D2N4), values relative to Spry1 intron 1. B) Southern blot analysis of Erk2⁻/⁻ lines. DNA was digested with KpnI and PstI and hybridized with a 3' external probe. Positions and sizes of WT and mutated DNA are as shown.

Figure 3.8: Erk1/2 expression in mutated lines. A) Q-PCR analysis showing reduced Erk2 transcript in the mutated lines, with no compensatory increase in Erk1 expression. Genotypes are as shown. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments. B) Western blot analysis of ERK1/2 protein levels in the mutated lines. ERK2 expression is reduced in the heterozygous line JuD2, and completely absent in the
homozygous lines D7N2 and D2N4. GAPDH is used as a loading control. Genotypes are as indicated.

Figure 3.9: Embryo-derived and ES cell derived Erk2 null lines have distinct morphologies. Morphological analysis of mES derived Erk2<sup>-/-</sup> lines showing typical mES cell morphology, as well as sub-populations of differentiated cells, in contrast to embryo-derived B1 Erk2<sup>-/-</sup> cells that form homogenous cultures of oblong, undifferentiated cells. Black arrowheads indicate differentiating cells, red arrowheads indicate undifferentiated colonies and purple arrowheads indicate atypical mES cell morphology. Images are taken at 20x magnification.
3.2.3: Analysis of Erk2 depletion on FGF/ERK signal transduction

To test whether ERK2 deficient cells were also robust to perturbations in the level of Erk dose, lysates from WT and Erk2^−/− cells were examined for ppERK immunoreactivity (Figure 3.10). In cells lacking ERK2, hyper-phosphorylation of ERK1 was detected, with over a 3 fold increase in ppERK1 levels over those in WT mES cells, substantially more than observed by Yao and colleagues (2003). However, the amount of total ppERK, i.e. ppERK1+ppERK2/GAPDH, was decreased by approximately 30% in ERK2 deficient cells (Figure 3.10-B). Furthermore, this ERK1 hyper-phosphorylation, but global decrease in ppErk levels, showed no apparent change over time following acute stimulation with Fgf2, and exhibited similar ppErk time-to-peak and decay kinetics in the presence or absence of ERK2 (Figure 3.11-A). This decrease in ppERK levels also correlated with a decrease in p90RSK phosphorylation on both ERK dependent sites: Thr359 and Ser363, (Dalby et al., 1998), indicating a reduction in the output early stage ERK signaling. The time-integrated ppERK response has been proposed to correlate with functional system output (Schilling et al., 2010). Semi-quantitate western blot analysis showed a reduction of approximately 20% in the cumulative ppERK response to FGF2 stimulation over the 120 minute window analysed (Figure 3.11-B).

As both ERK1/2 and p90RSK activity have been shown to be essential for IEG induction in other systems (De Cesare et al., 1998, Brüning et al., 1999), this attenuation in ERK signaling was further investigated at the level of transcription response to FGF stimulation. Figure 3.12 shows Q-PCR analysis of cFos, Egr1 and Egr2 IEG induction at 1 and 3 hours following acute stimulation with 10ng/ml FGF2. Stimulation in the presence of the MEK inhibitor PD0325901 (grey lines), showed that activation of both Egr1 and Egr2 promoters is entirely MEK dependent, however there was still appreciable induction of the cFos promoter. As the cFos promoter contains multiple transcription factor binding sites, including a Sis-inducible element that binds STAT 3 (Rajotte et al., 1993), it is possible that the MEK independent cFos expression observed is due to LIF activation of the JAK/STAT pathway. The decrease in ERK activity is exaggerated at the level of IEG
expression where analysis of 1h induction values showed a 60-70% reduction in IEG expression without ERK2 (Figure 3.12-D). Taken together, these experiments show that in the absence of ERK2, ERK1 is hyperphosphorylated, however this is not sufficient to completely compensate for loss of ERK2.
Figure 3.10: Semi-quantitative western blot analysis of ppERK levels in WT and ERK2 deficient cell lines. A) Hyper-phosphorylation of ERK1 in ERK2 deficient cells, ppERK1 values are normalized to those of total ERK1. B) Loss of ERK2 results in a decrease in total ppERK, although disproportionate to the decrease in total Erk levels. Cumulative ppERK1 and ppERK2 (where present) were normalized to GAPDH. Western blots were analysed with ImageJ64 software. Error bars denote standard error from 2 separate experiments. Genotypes are as indicated. (See Figure 3.11-A, lane 1 (top and bottom blots), for a representative blot).

Figure 3.11: Western blot analysis of both ERK and p90RSK phosphorylation following acute FGF2 stimulation. A) Top: WT mES cells, Bottom: ERK2 null mES cells. Blotting was performed using the multi-strip blotting method and thus band intensity is directly comparable across all samples. Blots are representative of results from 4 separate experiments. Genotypes are as indicated. B) Time integrated ppERK response in WT and
ERK2 mutant cells. ppERK levels were quantified, normalized to GAPDH and plotted as a curve. The time-integrated response was calculated as the area under each curve. Error bars denote standard error from 2 separate experiments.

Figure 3.12: IEG induction is attenuated in the absence of ERK2. A-C) Q-PCR analysis of IEG induction following acute stimulation with (10ng/ml) recombinant FGF2: Egr2, Egr1, and cFos expression in the presence of FGFR inhibitor PD173074 (0h), 1 and 3 hours following stimulation. D) Normalized peak induction, values for stimulations preformed in the presence of PD0325901 at 1h were subtracted from those of WT and ERK2 deficient samples for the same time point. Values are plotted as the per cent of WT induction. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 3 separate experiments. Genotypes are as indicated.
Auto-induced negative regulation plays a key role in the duration and subsequently the functional consequences of Erk signaling (Nakakuki et al., 2010 and reference therein). Members of the Dusp family of dual specificity phosphatases are the primary induced phosphatases that regulate the duration of ERK activity (Brondello et al., 1995, Brondello et al., 1997, Brondello et al., 1999). Both Dusp message, and protein stabilization are ERK dependent and therefore represents another regulatory mechanism that fine-tunes the ppERK response to growth-factor stimulation. IP-LC-MS/MS experiments, discussed in Chapter 5, identified DUSP6 and DUSP9 as ERK2 interactors. Figure 3.13-A shows that whilst Dusp6 and Dusp9 promoter activity is FGF sensitive, Dusp6 is positively and whilst Dusp9 negatively regulated, however this apparent negative regulation may be due to peek induction falling outside the resolution of the assay. Therefore, Dusp6 expression was used as a readout of transcriptionally induced ERK negative feedback. Q-PCR analysis shows that induction of Dusp6 is largely MEK dependent, and that the induction amplitude appears to be a function of ERK activity with a decrease of ~60% peak induction amplitude in the absence of ERK2, similar to cFos, Egr1 and Egr2 (Figure 3.13-B).

The classical transcriptional response to growth-factor stimulation is divided into two primary phases: immediate early and late response gene induction (Tullai et al., 2007 and references therein). The expression of LRG’s is dependent on the IEGs such as cMYC and cFOS (Tullai et al., 2007, Nakakuki et al., 2010), the stability of which, and thus activity, is ERK dependent (Murphy et al., 2002, Sears et al., 2000). As the DUSP dependent duration of nuclear ppERK plays a crucial role in this process (Nakakuki et al., 2010), and that DUSP expression and stability is ERK dependent (Brondello et al., 1999), it was rationalized that the attenuation of Dusp expression represented another point where perturbations in ppERK levels could be compensated for, and that this compensation could be visualized at the level of LRG expression, specifically D-type cyclin expression (Yang et al., 1997, Brown et al., 1998, Hosokawa et al., 1994).
Consistent with recent studies into the unique cell cycle architecture mES cells (reviewed in Dalton and White, 2005, and in Savatier and Afanassieff, 2002), CCND3 expression was both ERK and MEK independent. CCND1 expression was largely MEK dependent but only slightly induced by FGF2, within the timeframe analysed, CCND1 expression levels were comparable in the presence or absence of ERK2. CCND2 however, reported to be expressed at very low levels in mES cells (Savatier et al., 1996), was robustly expressed following FGF stimulation, the magnitude of which was again comparable in the presence or absence of ERK2 (Figure 3.14). These results indicate, albeit indirectly, the interdependence between DSUP and ERK activities contributes to the robustness of this system to perturbations in ERK dose. However further analysis is required to confirm this.
Figure 3.13: Analysis of ERK induced Dusp expression. A) Q-PCR analysis showing inverse responses of Dusp family members to FGF stimulation. B) Q-PCR analysis showing a decrease in Dusp6 expression levels in the absence of ERK2. WT cells were stimulated with recombinant FGF2 (10ng/ml) for the indicated times. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments.

Figure 3.14: D-type cyclin expression is comparable in the presence or absence of Erk2. Q-PCR analysis of D-type cyclin levels following FGF stimulation shows that whilst the expression of each isoform pattern differs, this difference is largely Erk2 independent. Stimulation was performed with 10ng/ml recombinant FGF2 for the indicated times. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel (CCND1, CCND3), and 2 sets of 2 separate experiments (CCND2).
To test whether the attenuation FGF/ERK signaling observed in the absence of ERK2 was isozyme specific, or a consequence of reduced ppERK dose, ERK2 depletion was rescued by transgenic expression of either ERK1 or ERK2. The ERK1 coding region was amplified from cDNA derived from E14tg2a total RNA with the primers ‘BE1F’ and ‘NE1R’, that contained BamHI and NotI sites, respectively. A poly-glycine linker was included at the 3’ end of the forward primer to aid the accessibility of the FLAG epitope. The resulting PCR fragment was TOPO cloned and sequence verified. The insert was then cloned into pPy CAG IP (IPC37) as a BamHI/NotI fragment, by 3-way ligation, between the FLAG and IRES sequences, and confirmed by restriction digestion. D7N2 Erk2 deficient cells were transfected with either of pPyCAG-Erk1-IP or pPyCAG-Erk2-IP (25µg) of Scal linearized DNA and selected for 8 days with puromycin (1µg/ml). Transfectants were expanded as a population.

Western blot analysis showed that exogenous ERK1 was expressed at comparable levels to endogenous ERK2 in WT cells, and that it was phosphorylated normally. Furthermore, analysis of endogenous ppERK1 levels in the FLAG-ERK1 expressing cells showed that it was phosphorylated to the same extent as in WT cells indicating that the hyperphosphorylation observed in ERK2 mutant cells was most likely dependent on difference in total ERK expression levels, or by ERK induced negative feedback (Figure 3.15). Interestingly, the kinetics of ERK1 activation appeared accelerated in ERK1 overexpressing cells, with a more rapid induction and decay than observed in other samples. Q-PCR analysis of IEG induction revealed that either ERK isozyme was capable of reinstating normal levels of IEG induction (Figure 3.16) arguing for a degree of redundancy between both proteins in mES cells.
Figure 3.15: Western blot analysis of ERK1 and ERK2 overexpressing cells. Genotypes are as indicated. Blotting was preformed using the multi-strip blotting method and thus band intensity is directly comparable across all samples.

![Western Blot Images]

Figure 3.16: IEG expression is rescued by overexpression of either ERK1 or ERK2. Q-PCR analysis of IEG induction following acute stimulation with (10ng/ml) recombinant FGF2: Egr2, Egr1, and cFos expression in the presence of FGFR inhibitor PD173074 (0h), 1 and 3 hours following stimulation. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
3.3 Discussion

3.3.1: Analysis of embryo-derived Erk2 mutant cells

Before this study it had been implied that ERK2 was the primary ERK isozyme involved in lineage commitment mES cells (Kunath et al., 2007). This model was based on the initial implication that MEK activity was essential for lineage specification (Burdon et al., 2002), followed by analysis of the Erk2 knockout phenotype. Experiments to rescue the block in neural commitment by transgenic expression of Erk2 failed, although expression of ERK2 reinstated normal to elevated ppERK levels as well as normal IEG induction, indicating the protein was properly expressed, folded and the N-terminal tag did not affect its function. When performing the initial rescue experiments I had some concerns that overexpression of ERK2 may have an adverse effect on the dynamics of ERK:MEK interactions. As ERK is phosphorylated in a distributive manner (Schilling et al., 2009, and Chapter 5 of this study), i.e. each phosphorylation event of the TEY motif occurs separately and therefore full ERK activation requires multiple MEK:ERK collisions, I rationalized that elevating the levels of ERK may increase the competition between un-phosphorylated and mono-phosphorylated forms of ERK, therefore acting in dominant negative fashion, as seen in scaffold overexpression experiments (reviewed in Kolch, 2005). Moreover, as many ppERK1/2 antibodies (including the three used in this study) recognize both dual TpEYp and mono TpEY motifs and abundance of mono-phosphorylated ERK would not be discriminated against by western blot analysis. Therefore I expressed Erk2 not only from pIPC37, but from a similar plasmid where puromycin resistance is expressed under its own promoter (mPGK), thus allowing robust drug resistance without selecting for high copy number, or insertion into a highly active locus, as can be the case when drug resistance is expressed from and IRES. Transgene expression from this plasmid has been known to give lower expression levels (I. Chambers, personal communication). Erk2 was also expressed using the FKBP L106P system (Banaszynski et al., 2006), where without the stabilizing drug Shld1, protein levels can range between very low to endogenous depending on the protein itself. However neither method showed any
obvious sign of rescuing the differentiation defect seen under monolayer conditions, as judged by morphology alone.

Cell density can play an important role in the behavior of cells in culture. In the initial experiments with the B1 Erk2 knockout cells neural induction was preformed with a slight but possibly important difference to Ying et al. (2003). Ying and colleagues plate cells at a density of $1.5 \times 10^4$/cm$^2$ directly into N2B27 defining the start of the assay. Erk2 null cells were pre-plated at the same density but in N2B27 supplemented with LIF, and the assay started 24 hours later by removing LIF from the culture medium (Kunath et al., 2007 and Materials and Methods 2.2.4.8.1). mES cell double approximately every 12 hours, therefore the starting density used to differentiate the Erk2 mutant cells was approximately $9 \times 10^4$/cm$^2$. Indeed when Erk2 mutant cells are differentiated without pre-plating in LIF, although attachment occurs as normal after plating, cultures exhibit a catastrophic wave of cell death at day 2, with virtually no cells surviving by day 7 (n=3). Therefore the maintenance of self-renewal seen under monolayer differentiation conditions may be due to an increase in the concentration of autocrine LIF produced at higher densities, enhancing the already high self-renewal capacity this cell line exhibits in standard culture conditions.

There are several possible explanations for the unique behavior of the embryo-derived mutant cell line. Certain traits, such as enhanced self-renewal may have been selected for during the derivation process. Although as two independently derived mutant cell lines, B1 and B3, exhibited the same behavior in culture (Kunath et al., 2007 and data not shown) this may not be the sole cause. There is also the issue of genetic background: the mutant cell line being from a 129/C57Bl/6 mixed background. Analysis of the phenotype of mES cells derived from both WT and heterozygous littermates would help to address this issue.
3.3.2: Generation of Erk2 knockout lines

Gene targeting in mouse embryonic stem cells has become a preferred tool in studying gene function, *in vivo* and *ex vivo*. At the current state of the art it is possible to make targeted alterations to the mouse genome as subtle as a single nucleotide, or as gross as several megabases. Genetrap mutagenesis has proved to be a highly fruitful approach as it allows the high-throughput generation of both constitutive and conditional null alleles (Skarnes et al., 2011). With large international consortiums dedicated to the generation of libraries of genetrap cell lines for use by the scientific community. Due to time considerations, and a lack of information as to the targeting efficiency of the *Erk2* locus, it was decided to take advantage of one of the public cell line collections of gene-trap mutated mES cells. As outlined in Appendix A2, the *Erk2* genetrap cell line RRZ307 (trapped with pGT2lxf (Figure A11)) was acquired from the BayGenomics repository, now at UC Davis. Though detailed analysis of the cell line indicated that although there was a bona fide trapping event of *Erk2* intron 1, the splicing efficiency of the cassette was most probably low and did not represent a nonfunctional allele. This was most probably due to the insertion of the gene-trap cassette as a concatemer, or due to inefficient trapping of the up-stream exon. Moreover, Southern blot analysis indicated the possible presence of a second insertion event. In general 5’RACE or RT-PCR analysis is deem sufficient to denote a proper trapping event, though this gives little information as to the efficiency of trapping, nor the presence of alternative trapping events. Analysis of the RRZ307 cell line highlights the importance in proper characterization of any gene-trap cell line.

As stated earlier, the efficiency of targeting at the *Erk2* locus was not publicly known when I began work on generating *Erk2* knockout cells. Therefore I felt it important to establish a rapid and sensitive method of prescreening drug-resistant colonies for correct targeting. Long-range PCR is generally the technique of choice for prescreening putatively targeted clones, however this method is prone to false positives that can arise due to annealing of products initiated by the different primers, leading to products consistent with appropriate targeting but that in fact
derive from clones containing random insertions (Gomez-Rodriguez et al., 2008). Therefore I established a method to sensitively detect gene copy number that was suitable for the analysis of low concentration, moderate purity gDNA, as is generally the yield from extraction methods in 96 well format. This method was able to robustly discriminate between single and double copy number and gave a false positive rate of 0 (n=10). With this method it is possible to screen up to 240 clones simultaneously, with a total hands-on work time of ~3 hours, from cell harvesting to data processing. It is likely that this method can be adapted to screen most knockout strategies as long as ~100bp or more of genetic material is removed during targeting.

3.3.3: Robustness of the MAPK pathway in mES cells

It is clear from recent single cell analysis of MAPK dynamics that signaling pathways have to function within a broad range of component parameters, yet ensuring a constantly reproducible, input specific outcome. This is achieved through a complex network of positive and negative feedback loops (Figure 3.17), with many of these loops emanating directly from ERK, thus ensuring that ERK activity is a major determinant of pathway dynamics. Various feedback mechanisms have been described in several cell systems (Waters et al., 1996, Santos et al., 2007, Douville et al., 1997, Sasagawa et al., 2005, Schilling et al., 2010, Brummer et al., 2003, Langlois et al., 1995, Eblen et al., 2004, Shin et al., 2009, Fritsche-Guenther et al., 2011), although there is evidence to indicate that such circuitry may not be common across cell types (Spakota et al., 2007). There is also indirect evidence, based on the degree of penetrance of both the aRaf and cRaf knockout phenotypes, that genetic background can affect MAPK signaling (Pritchard et al., 1996, Wojnowski et al., 1998). Moreover, as RAF proteins represent a key regulatory node in the MAPK pathway (Fritsche-Guenther et al., 2011, Shin et al., 2009), it may be possible that the differences in ERK1 activity seen between embryo-derived and mES cell-derived Erk2 mutant lines could be linked to genetic background specific alterations in the RAF signaling core.
In contrast, on a 129/Ola background it seems that there is a high degree of robustness to perturbations in ERK dose. Following ERK2 depletion, ERK1 is hyper-phosphorylated in mES cells, in agreement with studies in other cell types (Lefloch et al., 2008, Voisin et al., 2010). A recent study proposed a mechanism that explains how the topology and interconnectedness of the MAPK acts to regulate the input:output ratio. Sturm and colleagues demonstrated experimentally that the Erk pathway exhibits characteristics of a Negative Feedback Amplifier (NFA) (Sturm et al., 2010). An NFA is an engineering circuit developed in the 1920’s as a solution to the problem of increased noise in telephone lines that resulted when the signal needed to be amplified to travel across long distances. The solution to this problem was to define the magnitude of negative feedback as a function of system output. This form of feedback ensures that an increase in input will result in an increase in feedback, thus constraining the output within a specific range. The three tiered architecture of the MAPK core node, as well as an approximate 3 fold increase in protein abundance from RAF→MEK→ERK endows the cascade with properties of an amplifier (Schilling et al., 2010). Whereas strong negative feedback from ERK to RAF closes the circuit to produce the NFA (Figure 3.17). Assuming a sub-saturation signal:Erk ratio, reduction in the abundance of ERK can be compensated for, on a short timescale, by reducing the degree of negative feedback to, presumably, RAF, or MEK. However, if such a negative feedback loop is active in mES cells it cannot fully compensate for the loss of ERK2. There is a consistent reduction of 20-30% ERK activity in the Erk2 mutant cells, in steady state, and following acute stimulation. This reduction in ERK activity is amplified, possibly by signal loss due to phosphatase activity upon ERK translocation to the nucleus (Kholodenko et al., 2010), and equates to an approximate 60-70% reduction in IEG induction magnitude, which can be rescued by exogenous expression of either ERK1 or ERK2. However, due to the low temporal-resolution in the experiments described in this chapter it cannot be ruled out that IEG peak induction was delayed rather than reduced in ERK2 deficient cells.
One interesting outcome from analysis of IEG induction in mES cells is that stimulation experiments in the presence of MEK inhibition indicates that cFOS is not required to inhibit its own promoter. Work from the lab of Boris Kholodenko has recently shown that neither the duration of ERK signaling, nor DUSP mediated negative feedback can explain the transient nature of *cFos* expression (Nakakuki et al., 2010). From this they postulated that there must be a cFOS induced negative regulator of *cFos* promoter activity and based on evidence from the literature, proposed that this may be cFOS itself. Because the *cFos* promoter is activated in both MEK-dependent and -independent manners in mES cells, and because cFOS protein stabilization is MEK-ERK dependent, if cFOS negatively regulated its own promoter you would expect to see sustained *cFos* expression in samples stimulated in the presence of PD0325901, which was not the case (Figure 3.12-B). Though once again, analysis of the duration of *cFos* promoter activity at a higher temporal resolution that presented in this study would greatly strengthen such an observation.

The reduction in ERK activity seen in *Erk2* mutant cells is also reflected in transcriptional induced negative feedback by DUSP’s. This regulatory mechanism
operates on a longer timescale than the NFA described above, and although not thought to influence ERK activation dynamics to a great extent, DUSP activity regulates nuclear retention of ERK and subsequently its decay kinetics and is therefore a key regulator of the ERK dependent secondary wave of transcription. As depletion of ERK2 leads to an approximate 50% reduction in Dusp6 expression it would be very informative to compare the extent of ERK1 nuclear retention in the presence or absence of ERK2.

3.3.4 Redundancy of ERK isozymes in mES cells

As discussed in section 1.3.3 the issue of redundancy between Erk1/2 is a point of contentious debate. One reason for this may be primarily semantic: what do people define as redundant? If redundancy is defined as having largely overlapping function, then it is safest to say that ERK1/2 are redundant enzymes as all but a few experiments have shown them to be largely equivalent molecules and contribute to the pathway output to an extent that reflects their relative abundance. In the mES cell system the reduction in FGF2 induced, MEK1/2 dependent, transcription observed in Erk2 knockout cells was wholly restored upon overexpression of either ERK1 or ERK2, as well as normal ERK1/2 dependent phosphorylation of p90RSK (data not shown). It has been proposed that ERK1 has a lower affinity for MEK1/2 than ERK2 (Vantaggiato et al., 2006), however, when Erk2 knockout cells were rescued with exogenous ERK1 both endogenous and exogenous isoforms were phosphorylated based on their relative expression, exogenous ERK1 being expressed and phosphorylated to a similar extent as endogenous ERK2 in WT cells, which would suggest that MEK1/2 binding/activation of ERK1/2 isozymes is based on their relative expression, not isoform specificity. However comparative co-IP experiments would be needed to confirm this.

3.4 Future directions

One of the main conclusions from this chapter is that ERK2 alone is not essential for cell fate decisions in mES cells. In the light of this information, the only
evidence in the literature that specifically links ERK1/2 signaling to cell fate decisions in mES cells comes from experiments with MEK inhibitors. Therefore it is important to assess the effect of augmenting ERK-MAPK signaling, at the level of ERK itself, on cell fate decisions and the pluripotent state.

This study has also highlighted mES cells as a potentially useful system to study MAPK signal behavior, specifically robustness. Due to time and financial considerations the regulatory mechanisms that govern ERK activation were inferred from motifs known to be active in other cell types. Therefore our understanding of how this pathway is regulated would benefit greatly from analysis of feedback loops such as those from ERK to RAF. Approaches such as modular response analysis (Santos et al., 2007, Kholodencho et al., 2002), and sensitivity analysis (Schilling et al., 2010, Nakakuki et al., 2010) are proving powerful theoretical tools to determine the connectivity as well as crucial parameters of signaling networks that may operate below the sensitivity of standard biochemical methods.

It is important to note that the evidence presented in this chapter doesn’t discriminate between ERK1 hyperactivation by negative feedback regulation or as a consequence of increased binding to MEK1/2 in the absence of ERK2. Moreover, the issue of why ‘total ERK’ activity was less in the Erk2 mutant cells than WT is not clear. It may be that ERK1 isn’t expressed to a high enough concentration and therefore the ratio of signal to ERK1 is saturating in mutant cells. As very few studies have looked at the ratio of total ERK to ppERK we do not know if there is a linear relationship between concentration and activity, and more so even where the upper and lower limits to such a relationship lie. Therefore it could be informative to determine such a relationship using the FKBP-ERK2 expression system where the ratio of ppERK2:ERK2 can be determined as a function of ERK concentration as well as how this could effect IEG induction.

The experiments in this chapter suggest that the contribution of each ERK1/2 isozyme to the MAPK pathway output is a function of their relative expression levels. Largely unpublished work from both Sylvain Meloche and Philippe Lenormand (alluded to in the Discussion section of Lefloch et al. (2008))
indicate that different phenotypes arise with different Erk1/2 allele doses. Erk2+/+; Erk1+/+ mice are viable, Erk2+/−; Erk1+/+ mice are born at lower than expected Mendelian frequency, and Erk2+/−; Erk1−/− embryos die during development, a phenomenon observed on all genetic backgrounds tested. Gene-trap targeting vectors that allow a cDNA to be expressed from the endogenous Erk2 locus have been generated for this project but could not be implemented due to time constraints. It would be interesting to express an Erk1 cDNA from the Erk2 locus, effectively creating Erk2−/−; Erk1+/+/+ mES cells and assess their abilities to contribute to extra-embryonic structures in morula aggregate experiments thus determining whether the Erk2 knockout phenotype is a consequence of indiscriminate ERK deficiency or due to an absolute requirement for ERK2 function in the extra-embryonic tissue of the early embryo.
Chapter 4 - Functional characterization of ERK2 in mES cells

4.1 Introduction

Several studies have implicated FGF induced MAPK activity as an essential pathway governing normal peri-implantation development (see Chapter 1). Experiments in mES cells as an \textit{ex vivo} model to study early events in lineage specification have underscored this idea. A combination of genetic and pharmaceutical approaches have implicated FGF induced MEK1/2 activity as being essential mediators of this process, however experiments outlined in Chapter 3 of this study have raised questions as to the absolute requisite of ERK2 for lineage segregation in mES cells. Moreover, as second generation MEK inhibitors, as used in 3/2i culture, inhibit several MEK isoforms; MEK 1/2/5, and appear to enhance PI3K activity (Stavridis et al., 2007) the inference that ERK1/2 activity alone is responsible for FGF-induced lineage commitment may not be fully justified.

4.1.1 Aims of this chapter

The primary aim of this chapter was to assess the role of ERK2 in mES cell differentiation and self-renewal. Based on results from these experiments attempts were made to further perturb ERK1/2 function using dominant negative mutants inhibition. Furthermore, as microarray previously generated in our lab indicated that the transcription factor SFR may be a downstream target FGF signaling in mES cells, efforts were made to functionally characterize SRK deficient mES cells.
4.2 Results

4.2.1: ERK2 is dispensable for multi-lineage differentiation

To test how decrease in total ERK activity exhibited by Erk2 knockout mES cells translated into a functional response, their ability to commit to various lineages was tested. Erk2 mutant cells generated βIII tubulin positive neural tissue with similar efficiency to WT and heterozygous lines (Figure 4.1). Q-PCR analysis showed almost complete down-regulation of pluripotency markers, in all cell lines, by day 5 of neural induction. It is notable however that ERK2 deficient cells appeared to express higher levels of the pluripotency markers Nanog and Rex1 under self-renewing conditions (Figure 4.2-Day 0). Moreover, neural induction appeared to progress at a higher efficiency, as judged by expression of neuronal markers NeuroD3, Mash1 and Ngn2 (Figure 4.2), in the mutant cell lines. This phenomenon may be explained by the presence of a more homogenous starting population at day 0 for the mutant lines. If there was a higher abundance of cell in a self-renewing, Nanog-high, lineage unrestricted state it would follow that a higher percentage of cells would be capable of adopting a neural fate when subjected to the appropriate cues.

The ability of ERK2 deficient cells to commit to endodermal and mesodermal lineages was then assessed by embryoid body differentiation. Embryoid bodies derived from mutant lines exhibited similar morphological features to parental cultures (Figure 4.3-A), with cystic embryoid bodies visible at day 6, and rhythmically beating cardiac muscle evident by day 8. A time course of gene expression analysis showed no kinetic difference in the down-regulation of pluripotency markers (Oct4, Nanog), nor up-regulation of either endodermal (Gata6, Foxa2), with similar transient expression of T that marks the onset of mesodermal differentiation (Vigneau et al., 2007), although with a more rapid onset of Acta2 expression for ERK2 deficient lines (Figure 4.3-B).
Figure 4.1: ERK2 deficient cells efficiently neuralize under monolayer conditions. Immunocytochemistry analysis of cultures following 5 days of neural induction for neural marker βIII Tubulin immunoreactivity (red). DNA is counterstained with DAPI in blue. Images are taken at 20x magnification. Genotypes are as shown.
Figure 4.2.2: Gene expression analysis of cultures following 5 and 10 days of neural induction. Markers of pluripotence (Oct4, Nanog, and Rex1) and neural tissue (Mash1, NeuroD3, and Ngn2) are shown. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
Figure 4.3: Multilineage differentiation of mES cells is ERK2 independent. A) Phase contrast images embryoid bodies following 6 days of differentiation, arrowheads indicate presumptive cysts. Images are taken at 20x magnification. B) Time-course analysis of the progression of embryoid body differentiation, showing similar differentiation kinetics in the presence or absence of ERK2. Values are normalized to TBP. Genotypes are as shown.
The increase in *Nanog* and *Rex1* expression observed in Figure 4.2 indicated that mES cells may self-renew more efficiently in the absence of ERK2. However, as the mutant cell lines were subjected to 2 rounds of clonal selection, this heightened self-renewal may be due to a degree of culture adaptation. Therefore ERK2 deficiency was rescued in each independent line by transgene expression as outlined in Section 3.2.1. Clonal rescued lines were picked for both D2N4 and D7N2 lines and the extent of transgene expression was determined by western blotting (Figure 4.4-A). Transgenic expression of ERK2 was capable of reinstating normal levels of both *Nanog* and *Tbx3* to ERK2 deficient cells (Figure 4.4-B). This rescue was less obvious for *Klf4*, whilst no significant difference was observed for *Oct4* levels across all lines tested. ERK2 deficiency had the most significant effect on *Tbx3* levels (a one-tailed t-test giving p values of 0.02 and 0.01 for D7N2 and D2N4, respectively). This is consistent with previous reports that *Tbx3* is a direct target of MAPK signaling in mES cells, and that TBX3 can positively regulate *Nanog* promoter activity (Niwa et al., 2009).

Using *Nanog* promoter activity as a readout, the effect of ERK2 depletion on self-renewal was analysed at a single cell level. An EGFP-ires-Puro cassette (Figure 4.5-A) (Chambers et al., 2007) was targeted to the *Nanog* allele in both D7N2 and D2N4 lines. AdhI linearised pIPC208 (100µg) was electroporated into each ERK2 null line, selected for with puromycin (0.75µg/ml) for 9 days, before colony picking and expansion. Correct targeting was confirmed by Southern blotting (Figure 4.5-B) giving a targeting frequency of 90%. One clone for each parental null line was selected and ERK2 deficiency was transgenically rescued with a blasticidin selectable expression plasmid (pCAG-Erk2-IB). This was cloned by digesting pCAG-3xFLAG-Erk2-IP (Section 3.2.1) with NotI/XhoI to retrieve the Kozak-3xFLAG-Erk2 containing fragment, which was then ligated into pCAG-IB, and verified by diagnostic digestion. ScaI linearised plasmid (25µg) was then electroporated into each D2N4-TNG (Targeted Nanog GFP) and D7N2-TNG lines, and transfectants were selected for with Blasticidin (150µg) for 10 day. Stably transfected cells were collected as a population for further analysis.
Different Nanog expressing (GFP +ve) populations (Chambers et al., 2007) were visualized by flow cytometry analysis of lines under self-renewing conditions, at similar densities (~10^5 cells/cm^2) (Figure 4.6). An increase of 15-20% GFP +ve cells was noted in both high and low populations in the absence of ERK2, whilst the medium expressing population largely remained constant. This is consistent with previous results from Q-PCR population analysis. Although Nanog promoter activity was consistently higher in Erk2 mutant lines, the distribution varied (+/- ~20%) from one day to another.

Due to the long half-life of GFP, ~26h, (Corish et al., 1999), Nanog-GFP reporter lines were not suitable to analyze the effect of ERK2 depletion on the early stages of differentiation. Therefore a novel reporter construct (a gift from Kei Kaji), where an 8kb fragment of the Tcfcp2l1 promoter drives expression of a VENUS-PEST fusion protein (Figure 4.7-A). Addition of the PEST domain reduces the half life of EGFP to approximately 9 hours (Corish et al., 1999). TCFCP2L1 is a pluripotency-associated transcription factor of largely unknown function, whose expression is rapidly down-regulated upon LIF withdrawal (K. Kaji, personal communication) and is hence used as a marker of pluripotency. FspI linearised pTcfcp2l1VP (25µg) was electroporated into each ERK2 null line, selected in puromycin (0.75µg/ml) for 9 day, clones were picked, expanded and assayed for VENUS expression. One clone for each parental null line was selected and ERK2 deficiency was again rescued as outlined above. A time-course of the rate of VENUS down-regulation following LIF withdrawal was determined by flow cytometry (Figure 4.7-B-C). An increase of approximately 15-20% in VENUS +ve cells was observed in ERK2 deficient cells, however no detectable difference in the rate of VENUS down-regulation was noted within the 48 hour window analyzed.

To test if this difference in gene expression was functionally significant the self-renewal ability of ERK2 deficient cells was tested at clonal densities (Chambers et al., 2003), in the presence of LIF and determined by alkaline phosphatase staining after 7 days. Colonies were counted and scored based on the degree of differentiation observed (Figure 4.8). In WT and Erk2+/− a distribution of 10% undifferentiated, 70%
mixed, and 20% differentiated colonies was observed. This was in contrast to ERK2 deficient lines were a distribution of 30% undifferentiated, 65% mixed, and 5% differentiated colonies was observed. This distribution returned to WT upon transgenic expression of ERK2, indicating that ERK2 depletion enhances self-renewal in the presence of LIF. However, upon withdrawal of LIF cultures from all genotypes differentiated to an identical extent (data not shown).
Figure 4.4: ERK2 deficiency results in increased expression of pluripotency markers. A) Western blot analysis of ERK2 expression levels in rescued lines. Upper and lower bands correspond to endogenous ERK1 and ERK2, respectively (lanes 1-2). Lanes 3-8 correspond to the overlapping signals from endogenous ERK1 and exogenous tagged ERK2. Rescued clones highlighted in red were used for further analysis. GAPDH was used as a loading control. B) Gene expression analysis for pluripotence markers. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments. Genotypes are as shown.

Figure 4.5: Nanog-GFP targeting. A) Cartoon depicting the Nanog targeting strategy. Black bars depict positioning of the 5’ external probe used for Southern blotting. B) Southern blot analysis of targeted clones showing correct targeting in 9/10 clones tested. DNA was
digested with SexAI and hybridized with a 5' external probe. WT=10kb, mutated=7kb. D2TNG and D7TNG are derived from D2N4 and D7N2 ERK2 null lines, respectively.

Figure 4.6: Increased Nanog promoter activity in the absence of ERK2. A) Dot-plots of GFP expression in ERK2 deficient and rescued lines, SSC is shown on the Y axis, GFP on the X axis. B) Quantification of data shown in A). Results are representative of 5 separate experiments. Data was analysed using FLOWJO software.

Figure 4.7: ERK2 deficiency does not affect the kinetics of early differentiation. A) Cartoon depicting the reporter construct. B) Depicting dot plots of the change in VENUS
expression following LIF withdrawal in ERK2 null and rescue cell lines. SSC is shown on the Y axis, VENUS on the X axis. C) Quantification of data shown in B). Error bars denote standard error from 2 experiments conducted in parallel.

Figure 4.8: ERK2 depletion enhances self-renewal in the presence of LIF. A) Quantification of clonal counts. B) Representative images of the scoring criteria used to score the extent of differentiation. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments.
4.2.2 Attenuation of ERK signaling by overexpressing ERK2 mutant proteins

During the course of this study several methods were undertaken to gauge the effect of modulating ERK activity on the self-renewal potential of mES cells. RNA interference was employed to deplete ERK2 deficient cells of the remaining ERK1 isozyme, however transient knockdown by siRNA or stable knockdown by transfection with Erk1 shRNA expressing constructs yielded none to very modest knockdown. Therefore, experiments were undertaken to modulate ERK signaling by overexpression of a panel of known, and putative, dominant negative ERK2 isoforms (Table 4.1). ERK2 mutants, K52R, Y185F and T183A/Y185F have all been shown to act in a dominant negative fashion in other cell types (Robinson et al., 1996, Cha et al., 2001, Cha et al., 2001). Mutation of Y261 has been shown to interfere with the ERK:ELK1 and ERK:cFOS interactions by way of their conserved DEF domain (Dimitri et al., 2005). ERK2-EN and ERK2-VP16 mutants were constructed in response to a recent publication that demonstrated that ERK2 was capable of direct DNA binding, and acted as a transcriptional repressor (Hu et al., 2009).

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>K52R</td>
<td>Affects ATP orientation in the ATP-binding pocket. Reduces the $K_{cat}$ by approximately 50%.</td>
<td>Robinson et al., 1996</td>
</tr>
<tr>
<td>Y185F</td>
<td>Single mutation of the TEY activation motif.</td>
<td>Cha et al., 2001</td>
</tr>
<tr>
<td>T183A, Y185F</td>
<td>Double mutation of the TEY activation motif.</td>
<td>Cha et al., 2001</td>
</tr>
<tr>
<td>Y261A</td>
<td>Docking domain mutant.</td>
<td>Dimitri et al., 2005</td>
</tr>
<tr>
<td>Erk2-En</td>
<td>An Erk2 C-terminal fusion with the engrailed repressor domain.</td>
<td>This thesis</td>
</tr>
<tr>
<td>Erk2-Vp16</td>
<td>An Erk2 C-terminal fusion with the Vp16 transactivation domain.</td>
<td>This thesis</td>
</tr>
</tbody>
</table>

Table 4.1: List of Erk2 mutants generated for this study.

ERK2-K52R, ERK2-Y185F and ERK2-Y261A were generated by wraparound mutagenesis with the primers K52RF/R, Y185FF/R and Y261AF/R, respectively. ERK2-T183A/Y185F was constructed by PCR fusion by overlap
extension with the two internal primers: AEF-F/R and external primers: BamHI Erk2 F and Erk2 NotI R. ERK2-EN and ERK2-VP16 were both also constructed by PCR fusion by overlap extension with internal primers EEN-F/R and EVP-F/R, respectively, and external primers XKE-F, and ENN-R (ERK2-EN) and EVPN-R (ERK2-VP16). All mutants were sequence verified and cloned into pIPC37.

Although stable lines were made expressing each mutant on an ERK2 deficient background, there was a degree variation in the number of clones derived for each electroporation. Therefore, to obviate the issue of selection for low transgene expression that may mask a toxic effect of one or more of the mutants, the effect of Erk2 mutant overexpression on self-renewal was determined by transfecting cells with each construct and plating them at clonal density (± LIF) 24 hours later. ERK2 null cells (10^6) were transfected with linearised plasmid (3µg) in the presence of 3µl Lipofectamine 2000. Transfection efficiency was estimated at 30% based on flow cytometry analysis of EGFP transfected cells, conducted in parallel. The rate of plasmid integration was assumed to be ~3% (Kaji et al., 2009). Therefore 24 hours post-transfection, cells were trypsinized and replated at 1.8x10^4 cells/cm^2 directly into selection media (1µg/ml puromycin ±LIF), and selected for 7 days, with media changes every second day, followed by fixation and AP staining. Surprisingly none of the ERK2 mutant tested were capable of sustaining LIF independent self-renewal, even to a modest degree, under these conditions (Figure 4.9). To determine if this was a due to low level transgene expression, a similar experiment was carried out using the episomal E14/T system (Chambers et al., 2003). E14/T cells express the large-T antigen and are capable of maintaining plasmids with a SV40 origin of replication as high copy number episomes (Gassmann et al., 1995, Aubert et al., 2002). E14/T cells were transfected as outlined above, except with circular plasmids, and 24 hours post-transfection cells were plated at a density of 6.6x10^3 cells/cm^2 directly into selection media (1µg/ml puromycin ±Lif). Cultures were selected for 7 days with media changes every second day, followed by fixation and AP staining. Again, no ERK2 mutant was capable of maintaining Lif independent self renewal,
moreover episomal expression of both WT, and every mutants tested, was toxic to mES cells (Figure 4.10).

These data indicated that the mutants do not act in a dominant negative fashion in mES cells, in contrast to published data from other cell types. Therefore, the best described ERK2 mutant, K52R, was selected for further study. ERK2 null clones stably expressing FLAG-tagged ERK2-K52R were generated and the extent of transgene overexpression was determined by western blot analysis (Figure 4.12). Of the six clones tested, all showed a modest degree of overexpression. Analysis of activation kinetics following FGF2 stimulation showed that ERK2-K52R was phosphorylated to a similar extent as WT ERK, if not slightly more so (Figure 4.12-A). Surprisingly, expression of the K52R mutant resulted in an increase in p90RSK phosphorylation (Figure 4.2.11-B), as well as an apparent rescue of cFos expression following FGF stimulation (Figure 4.2.11-C).
Figure 4.9: Clonal analysis of a panel of ERK2 mutant constructs expressed in Erk2<sup>−/−</sup> cells. Cells were plated at 1.8x10<sup>4</sup> cells/cm<sup>2</sup> 24h post transfection and selected in 1μg/ml pruomycin (±LIF). Following 7 days of selection cells were fixed and stained for alkaline phosphatase activity. Mock treated cells were transfected with an empty vector. Nanog transfected cells act as a positive control.
Figure 4.10: Clonal analysis of a panel of ERK2 mutant constructs expressed episomally in E14/T cells. Cells were plated at 6.6x10^3 cells/cm^2 24h post transfection and selected in 1μg/ml puromycin (+Lif). Following 7 days of selection cells were fixed and stained for alkaline phosphatase activity. Mock treated cells were transfected with an empty vector.
Figure 4.11: Western blot analysis of Erk2 mutant cells stably expressing 3xFLAG-tagged ERK2-K52R. A) Top blot ERK1/2, Bottom blot GAPDH. The clone marked in red was used for further studies.

Figure 3.12: Expression of ERK2-K52R positively effects the FGF/ERK pathway output. A) Western blot analysis of both ppERK and p90RSK levels following FGF stimulation. B) Quantification of the p90RSK levels in A. C) Q-PCR analysis of cFos expression in the presence of FGFR inhibitor PD173074 (0h), 1 and 3 hours following stimulation. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
4.2.3 The role of Srf in mES lineage commitment

It was thought that targeting the transcriptional mediators of ERK signaling may help provide valuable information as to the function of ERK1/2 signaling in mES cells. The SRK-ELK1 ternary complex is well described in the literature and can mediate the transcriptional response of mitogen-induced ERK activity (Marias et al., 1993, Latinkic et al., 1996, and reviewed in Papavassiliou, 1996). Srf mutant mES cells have been described previously and are known to exhibit a context-dependent block in mesoderm differentiation, with mesoderm induction impaired under monolayer conditions, but permissive in 3D aggregates (Weinhold et al., 2000), however differentiation into other lineages was not tested. Moreover, as induction of several IEGs in mES cells is SRK dependent in (Schratt et al., 2001), and FGF4 responsive (T. Kunath unpublished), analysis of the differentiation potential of SRK deficient cells could help to ascribe a functional significance to the FGF/ERK→IEG pathway in mES cells.

Srf mutant mES cells (Du et al., 2003) were generated by replacing the first exon of Srf, which codes for a portion of both the 3’ UTR and protein coding region, with a mPGK-Neo cassette to generate heterozygous mutant cells. The heterozygous line was then subjected to high G418 selection (Lefebvre et al., 2001) and clones that survived were screened for loss of the WT and duplication of the neo allele. The published phenotype of the Srf mutant cells was confirmed using the Nishikawa mesoderm monolayer protocol (Nishikawa et al., 1998) (Figure 4.13). By day 4 of the assay mutant cells maintained an mES cell morphology (Figure 4.13-A), failed to downregulate Oct4, nor was their evidence of T upregulation, in comparison to E14tg2a cells (Figure 4.13-B). Srf mutant cells were then assayed for their ability to generate neuronal tissue under monolayer conditions (Figure 4.14). After 7 days in culture, mutant cells maintained OCT4 immunoreactivity and there was scant evidence of up-regulation of the neuronal marker βIII-Tubulin. To test if this phenotype was due to SRF depletion, mutant cells were transgenically rescued by random integration of a 3xFLAG-Srf expressing construct. The Srf coding region was amplified from cDNA derived from E14tg2a total RNA with the primers
‘BSRFF’ and ‘NSRFR’, that contained BamHI and NotI sites, respectively. A polyglycine linker was included at the 3’ end of the forward primer to aid the accessibility of the FLAG epitope. The resulting PCR fragment was TOPO cloned and sequence verified. The NotI site in the Srf coding region was removed by mutating A115 GCG to GCA with the primers DNOTF/R. The insert was again sequenced and cloned into pPy CAG IP (IPC37) as described previously.
Figure 4.13: *Srf*<sup>−/−</sup> mES cells fail to differentiate under mesoderm monolayer conditions. A) Phase contrast images of both *Srf*<sup>−/−</sup> and WT mES cells during monolayer mesoderm differentiation at days 2, 3, and 4. B) Q-PCR analysis of cultures shown in A for *Oct4* (pluripotency) and *T* (mesoderm).
Figure 4.14: *Srf*<sup>−/−</sup> mES cells fail to make neural tissue under monolayer conditions. Cultures were differentiated for 7 days followed by fixation and immunohistochemical analysis for βIII Tubulin and PAX6 (neuronal), and OCT4 (pluripotency). DNA is counterstained with DAPI in blue. Images are taken at 10x magnification.
Srf mutant cells were electroporated, selected, and clones picked and expanded as usual and screened for transgene expression (Figure 4.15-A). All clones tested expressed the transgene at the predicted molecular weight, however expression levels was several fold higher than endogenous. Moreover, there was evidence of protein degradation, which has been reported previously upon SRF overexpression (Drewett et al., 2001). SRF degradation, specifically by Caspases has been shown to generate a stable dominant negative product that can interfere with IEG expression. Therefore to functionally test the extent of transgenic rescue clones were subjected to acute FGF2 stimulation for 1 hour and IEG expression was determined by Q-PCR (Figure 4.15-B). As reported in other Srf−/− lines, IEG induction was Srf dependent in mES cells. However, transgenic expression of SRF was only partially able to rescue this defect, with IEG expression between 10-25% that of observed in WT cells. Moreover this degree of transgenic rescue was unable rescue the neural defect (Figure 4.16). Rescued cultures were largely indistinguishable from the parental mutant line after 6 days of neural induction, with no detectable PAX6 immunoreactivity. Cells maintained an mES cell morphology, as well as OCT4 expression (Figure 4.16-A). Q-PCR analysis showed persistence of Nanog expression in most clones, as well as a failure to up-regulate Ccnd2 (Figure 4.16-B).

Because it was clear from transcriptional analysis that transgenic SRF expression only party restored Srf function, it was not possible to deduce whether the block in neural induction was due to SRF depletion. SRF was recently shown to be involved in epidermal stem cell differentiation (Connelly et al., 2010). Therefore I took advantage of a monolayer protocol developed in our institute that generates AP2α/Keratin 14 positive putative surface ectoderm tissue (Kunath et al., 2007, S. Lowell personal communication) to test whether SRF had a role in this process. Differentiation is carried out according the neural monolayer protocol, except that the culture medium is supplemented with recombinant hBMP4 (10 ng/ml). Following 6 days of differentiation WT mES cells had generated large sheets of flattened OCT4 negative, AP2α positive cells, which were undetectable in Srf mutant lines. There was some evidence of differentiation in the rescued lines, with OCT4 negative, AP2α
positive cells on the periphery of many colonies, however, this was greatly reduced in comparison with WT cells (Figure 4.17). Moreover, AP2α positive cells derived from SRF rescue lines did not exhibit the same morphological features as those from WT cultures, i.e. large flat cells, and thus may not represent actual surface ectoderm cells.

Following the difficulty to rescue the differentiation defect in the Srf mutant cells, and due to concerns as to the genomic integrity of this line, which exhibit an average karyotype of 46, it was decided that they were not a suitable model to test SRF function. Therefore efforts were made to generate new mutant lines.
Figure 4.15: Transgenic Rescue of SRK deficiency. A) Western analysis of Srf expression from right to left: WT E14tg2a, Srf<sup>-/-</sup>, Srf<sup>-/-</sup>-GFP (mock), and 6 rescued clones. B) Q-PCR analysis of cFos and Egr1 induction 1 hour after acute stimulation with 10ng/ml recombinant FGF2. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
Figure 4.16: Neural commitment is not rescued by transgenic add-back. A) Cultures were differentiated for 6 days followed by fixation and immunohistochemical analysis for PAX6 (neuronal), and OCT4 (pluripotency). DNA is counterstained with DAPI in blue. Images are taken at 10x magnification. B) Q-PCR analysis of cultures from A. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
4.2.4 Derivation and analysis of Srf<sup>+/fl</sup> and Srf<sup>fl/fl</sup> lines

Srf<sup>fl/fl</sup> mice are a congenic strain, derived from mice where the targeted allele (Figure 4.17-A) was introduced into 129/Sv mES cells that and correctly targeted cells were injected into C57Bl/6 blastocysts to make chimeras (Miano et al., 2004, Ramanan et al., 2005). The chimeric mice were then backcrossed to C57Bl/6J. Upon arrival males were bred to 129 females and heterozygous mES cells were derived by Jan Ure at the ISCR Transgenics Unit as described in section 2.2.6.2. After the mouse colony was expanded homozygous floxed mES cells were derived. Mice were genotyped at each generation for the presence of the targeted mutation. Western blot analysis showed normal SRF levels indicating the targeted insertion did not negatively affect Srf expression (Figure 4.17-B). Chromosome counts showed variable ploidy for each line derived (Figure 4.17-C), however one of each heterozygous (Fl4) and one homozygous (Sfh11) lines exhibited a near normal karyotype and were selected for further analysis. Both lines proved highly unstable under standard culture conditions (GMEM, 10% FCS and LIF), were difficult to manipulate, and the frequency of spontaneous differentiation decreased rapidly (~10 passages), indicating a degree of culture adaptation. Furthermore, analysis of both lines (~passage 10) showed inefficient neural commitment as well as decreased differentiation in embryoid bodies compared to WT E14th2a mES cells (data not shown). Therefore early passage cells (p3) were thawed to media supplemented with 1µM PD0325901 and 3µM CHIR99021 in the hope of alleviating the selective pressure for cells that have enhanced self-renewal. Under these conditions both lines became more manageable, although there were distinct morphological differences between heterozygous and homozygous lines, most probably reflecting their different genetic backgrounds (data not shown).

Both lines were stably transfected with a plasmid that expressed Cre recombinase C-terminally fused to the FKBP-L106P destabilized domain, and N-terminally ErT2 estrogen binding domain, thus allowing Cre protein stability and nuclear transport to be regulated by Shld1 and doxycycline, respectively (P. McDonnel, subcloned into pIPC37 as a FspI/NotI fragment) (Figure A2.4/4.18-A).
However following 48 hours drug treatment no discernible difference in SRF protein levels could be detected in any clone, from either cell line, in the presence or absence of drugs (Figure 4.18-B).

As the lack of Cre excision may have been due to improper drug induction of the fusion protein the effect of transient expression of a constitutively expressing Cre plasmid was assessed. Both homozygous and heterozygous lines were transfected (Lipofectamine 2000) with either 3µg pCAG-Cre-IP with 1µg pCAG-EGFP-IP, or 3µg pCAG-EGFP-IP alone, or mock transfected with no DNA. The following day cells were trypsinized and replated in selection media (1µg/ml puromycin) and cultured for 48 hours, following which the media was replaced with normal media and cells were cultured for a further 6 days, upon when colonies were fixed and stained for alkaline phosphatase activity. As both cell lines were highly sensitive to puromycin selection 48 hours of antibiotic treatment was sufficient to kill all non-transfected cells. Expression of Cre however appeared highly toxic to both cell lines, with virtually no homozygous mutant clones surviving the experiment (Figure 4.19). As Cre itself can result in a certain degree of cellular toxicity it is not clear if the cell death was due to excision of the Srf locus, or general Cre toxicity. Some heterozygous clones survived the experiment, however as this line was more tolerant of various manipulations (electroporation, transfection, etc.), it is not possible to make any conclusions as to the requirement of Srf for mES cell viability from this.
Figure 4.17: Derivation of $Srf^{+/\text{fl}}$ and $Srf^{+/\text{fl}}$ lines. A) Cartoon depicting the structure of the $Srf$ allele. LoxP sites are represented by triangles. B) Western blot analysis comparing SRF protein levels across 5 mES cell lines. WT; E14tg2a and Ju09, Erk2 genetrap; RRZ307, $Srf^{+/\text{fl}}$; Fl4, $Srf^{+/\text{fl}}$; Sfh11. C) Chromosome counts of both heterozygous and homozygous floxed Srf lines. Lines were karyotyped between passages 3 and 4. Lines highlighted in red were used for further study.

Figure 4.18: Drug regulated Cre mediated excision. A) Cartoon depicting the mechanism of action of dual drug regulatable Cre fusion protein. B) Western blot analysis of the affect of drug induction on heterozygous and homozygous floxed Srf lines. Shld1 was used at 1nM, Dox was used at 1μg/ml.
Figure 4.19: Cre expression is toxic to both heterozygous and homozygous floxed Srf lines. Cells were transfected (lipofectamine 2000) with either Cre (3μg) and EGFP (1μg), or EGFP alone (3μg), or no DNA. Plates were stained for alkaline phosphatase activity after 8 days.
4.3 Discussion

4.3.1 ERK2 deficiency enhances self-renewal under standard culture conditions, but is dispensable for lineage commitment

Following from the observations made in Chapter 3 that the differentiation defect observed in the Erk2 mutant cells, B1 and B3, was not due to the targeted disruption of the Erk2 allele, the role of ERK2 in mES cell biology remained unclear. In fact, the only data that implicated ERK1/2 in attenuation of the pluripotent state was inference from experiments with MEK inhibitors (PD184352 and PD0325901) which inhibit multiple MEK isozymes (Bain et al., 2007). Therefore the contribution of other MAPK’s such as ERK5 cannot be ruled out. As analysis of such mechanisms was beyond the scope of this study, the specific role of ERK2 in mES cells was analysed. Erk2 mutant cells exhibited no apparent defect in commitment to any lineage tested, and appeared to differentiate in a more homogenous fashion that the parental WT line. This data is consistent with recent reports for the requirement of ERK2 in the epiblast of the early embryo (Voisin et al., 2010).

Whilst conducting the experiments into the effect of ERK2 depletion on lineage commitment, an increase in the expression of the pluripotency associated transcription factors, Nanog and Rex1 was noted. Subsequent analysis showed that this was ERK2 dependent. Expression analysis showed a significant decrease in the expression levels of Tbx3 in both ERK2 null lines tested which was rescued upon transgenic add-back. Tbx3 expression has been shown to negatively correlate with MAPK activity in mES cells (Niwa et al., 2009), and that this correlates with an increase in Nanog and Klf4 expression, although the exact mechanism, and transcriptional hierarchy is still not well understood. Moreover, this difference in pluripotent gene expression correlates with the enhanced self-renewal of Erk2 mutant cells at clonal density in the presence of LIF, however there was no observable difference between all genotypes upon LIF withdrawal (data not shown). Experiments with both Tcfcp2I1 and Nanog reporter lines underscored these observations on a single cell level. It is important to note though that Nanog expression is used as a readout of the pluripotent state in these experiments, and no
connection between Nanog heterogeneity and ERK activity is inferred. However, single cell analysis for Nanog-GFP and a cell surface pluripotency marker such as SSEA1 or PECAM would help to distinguish between Nanog low mES cells, and Nanog low differentiating cells and determine the relationship, if any, that ERK activity may have with Nanog heterogeneity.

Since both LIF and FGF signaling activate ERK1/2 in mES cells, the difference in pluripotent gene expression and self-renewal observed in Erk2−/− cells may be due to differential activation kinetics of ERK1/2 by either FGF/RTK or LIF/gp130 signaling, with the former resulting in sustained activation and the latter transient. Experiments were started to compare the activity of ERK isozymes when stimulated with either FGF2 or LIF and preliminary results indicate that LIF induced ERK1 activity is more transient than when stimulated with FGF2 (Figure A12). As ERK1 is known to shuttle to the nucleus at a slower rate than ERK2 in murine fibroblasts (Marchi et al., 2008), ERK1 may be less efficient at relaying transient signals, but equivalent to ERK2 in response to a sustained signal. Therefore in self-renewing cells the cumulative MAPK output (i.e. FGF and LIF) is reduced in the absence of ERK2 as the LIF induced signal is not transmitted as efficiently by ERK1. However, upon LIF withdrawal the output is equivalent between WT and Erk2−/− cells, since all of the activating signal is FGF derived. This may explain why the Erk2 phenotype only manifests under self-renewing conditions. Although preliminary experiments indicated that there may be differences in the duration of ERK1 activity depending on the stimulus, the ERK1/2 activations kinetics are variable and therefore these experiments will have to be repeated several times to determine if this is a reproducible phenomenon. Furthermore, although FGF2 treatment is capable of rescuing the differentiation defect exhibited by FGF4 mutant cells, there may be differences in the kinetics and duration of ERK1/2 activity elicited by either ligand and therefore the physiological relevance of this phenomenon is unclear.
4.3.2 ERK activity cannot be attenuated by overexpressing nonfunctional Erk2 mutants

Although it appears that ERK2 can contribute to the regulation of the pluripotent state, it still remains unclear the extent to which combined ERK1/2 activity acts to transmit the differentiation-inducing signal from FGF4. Erk1−/−:Erk2fl/fl mice have been described previously (Viosin et al., 2010, Fischer et al., 2005), and the effect of ERK1/2 depletion has been described in other systems, however how this relates to the developing embryo, or to mES cells is unclear, and a point of great significance. Therefore experiments were undertaken to knockdown Erk1 in ERK2 deficient cells using siRNA. A pool of siRNA against Erk1 were used (Santa Cruz sc-29308), and although transfection efficiency was over 90% (gauged using FITC-conjugated control oligos), no discernible decrease in ERK1 protein levels could be detected either 24 or 48 hours post-transfection, in either WT, heterozygous, or homozygous mutant lines, indicating that this was not due to selection against low levels of ERK1 in the Erk2 mutant lines. Similar experiments using plasmid based methods (pTRIPz, OpenBiosystems) also failed for technical reasons.

It was then thought that overexpression of a panel of ERK2 mutants may yield valuable information to the role of ERK1/2 in lineage commitment. It was rationalized that blocking ERK1/2 activity downstream of MEK may provide insight to the how pluripotency is maintained by MEK inhibitors, i.e. ERK dependent or independent. Several known, and novel ERK2 mutants failed to maintain mES cell self-renewal in the absence of LIF, which may not have been due to sub-optimal expression levels, as episomal expression proved toxic. A more in-depth analysis of one point mutant, ERK2-K52R, showed that its expression actually enhanced FGF-ERK signaling and resulted in increased p90RSK phosphorylation and cFos induction following FGF stimulation. This may appear counterintuitive, but in light of the hight degree of ERK mediated feedback regulation in mES cells it is possible that the apparent rescue of p90RSK and cFos levels observed upon ERK2-K52R expression may be due to enhanced ERK1 activity. Assuming no preferential activation of either ERK isozyme, as the signal reaches ERK1/2 the degree to which
it is transmitted by either isoform is based on their relative abundance. Hence, as the degree of negative feedback is proportional to the input received by ERK1/2, and as ERK2 is present at a higher abundance, the proportion of negative feedback transmitted through ERK2 is lost, hence resulting in heightened activation of ERK1 (Figure 4.20). Unfortunately, as ERK2-K52R was FLAG-tagged in these experiments, and hence migrated with an identical molecular mass as ERK1, it was not possible to gauge the extent of ERK1 phosphorylation. However it would be of considerable interest to determine if the increase in ERK1/2 activity observed upon ERK2-K52R expression was due to enhanced ERK1 activity, and whether this correlates with an increase in RAF activity, as this may represent a valuable model to determine the feedback thresholds of the FGF-ERK pathway in mES cells.

Figure 4.20: Cartoon depicting three possible feedback scenarios leading to enhanced ERK1/p90RSK/cFos activities.

4.3.2 The role of Srf in mES cell biology

Due to the apparent self stabilizing nature of MAPK signaling in mES cells, it was thought that targeting the downstream transcriptional machinery may prove more tractable. Initially, attempts to target both SRK and ELK1 were made. Over expression of an ELK-EN fusion protein proved toxic to mES cells, however
this cellular toxicity was still observed when the B-box of ELK1, essential for SRF interactions (Vickers et al., 2004), was mutated. This indicates that the toxicity observed was not due to deregulation of SRF targets but most probably a result of inhibition of promoter regulated by other ETS proteins that share a similar DNA binding specificity (Boros et al., 2009). Also the observation that a certain degree of cellular toxicity results from the expression of engrailed fusion proteins in mES cells (H. Niwa personal communication) may also be a contributing factor.

As it was not possible to fully rescue SRF function by transgenic expression it is not possible to discern whether the block in neural and non-neural surface ectoderm commitment could be attributed to the targeted disruption of the Srf locus. Expression of SRF at lower levels using the FKBP-L106P system (Figure A13) resulted in a restoration of differentiation at clonal densities (Figure A14). However this low level expression did not permit neural commitment, as judged by morphology alone (data not shown), therefore efforts were focused on studying Srf function in another cell line.

Initial experiments using shRNA (pSM2c, OpenBiosystems) to ablate Srf expression in mES cells showed that ~50% knockdown was achievable in transient assays, however when the same plasmids were used to generate stable lines no knockdown was observed (3 hairpins, 60 clones). At the time it was thought that this may be due to selection against low SRF levels and therefore raised the issue of selection against high levels of knockdown in long-term culture. Therefore it was decided to take a conditional knockout approach to study Srf function. To do this I took advantage of a mouse line, available from the Jackson Lab mutant mouse repository, where a 2.5kb stretch of the proximal promoter and exon 1 of the Srf allele was flanked with LoxP sites, allowing conditional ablation of Srf expression (Miano et al., 2004). mES lines derived from these mice were very difficult to manipulate, were unstable in culture, had variable ploidy, and eventually work had to be abandoned as it appeared it was not possible to remove the floxed sequence by Cre expression.
At present there is very little evidence as to the function of SRF in cells of the peri-implantation embryo, or in their ex vivo counterparts. Moreover, both Srf null mES cell lines, where differentiation phenotypes have been assayed, have been made by high-G418 selection which places a strong selective pressure on mES cells, which even under standard culture conditions can be prone to culture adaptation. In most cases when assayed, the expression of SRF target genes (skeletal, cardiac, and visceral smooth muscle actin) were used as markers of differentiation, even though mES cells that overexpress SRF ectopically express such actin genes under self-renewing conditions. Furthermore, loss of expression of pluripotency associated genes was not assessed. Therefore it is difficult to conclude whether Srf mutant cells actually exhibit a block in differentiation, as presumably in the absence of Srf promoter activity of these genes will be limited, and upon transgenic ‘rescue’ these genes will again be expressed. SRF itself appears to be expressed at low levels in mES cells, its expression is positively regulated by FGF signaling (data not shown), and increase over the course of differentiation (Weinhold et al., 2001), therefore if SRF has a role to play in lineage commitment it may be secondary to the transition to differentiation competence that is associated for FGF/ERK signaling (reviewed in Nichols and Smith, 2009).

It is clear however that Srf deficient cells can generate mesoderm tissue under certain conditions (EB aggregates, teratomas, DMSO induced monolayer cultures), however the function of mature myocytes and sarcomerogenesis is SRF dependent (Niu et al., 2008). As discussed earlier proper embryonic patterning depends on positional cues derived from the extra-embryonic tissue. During early development the extra-embryonic ectoderm provides BMP signals to the epiblast that are essential for proper streak formation and mesoderm induction (Mishina et al., 1995). Knockout studies into a newly discovered SRK target gene, ELF5 (Sun et al., 2006) show a defect in embryonic patterning, a loss of BMP expression in the extra-embryonic ectoderm, and embryonic lethality between E8.5-E9.5 (Donnison et al., 2005). Moreover, ex vivo analysis of Elf5 mutant mES cells and embryos showed that it was essential from trophoblast specification and maintenance (Ng et al., 2008).
Therefore it would be interesting to determine if the *Srf* mutant phenotype can be rescued by tetraploid complementation with WT 4n cells. Recent evidence that forced SRF expression in TS cells induces differentiation to giant cells (Asanoma et al., 2007) adds further weight to a role for SRF in extra-embryonic function, and indicates that the perceived defect in mesoderm formation noted for the *Srf* knockout embryo may be secondary to defects in the extra-embryonic ectoderm.

4.4 Future directions

As discussed earlier, considering the data presented in Chapters 3 and 4, analysis of a double ERK1/2 knockout mES cells/embryo analysis would be of great relevance to the field. Moreover, as all of the MEK inhibitors used to maintain mES cells in a ‘ground state’ of pluripotency inhibit MEK1/2/5 isoforms (Bain et al., 2007), analysis of the function of ERK5 may also prove informative. ERK5 has been shown to mediate the differentiation-inducing signal of the TNF family cytokine LIGHT (Zou et al., 2006). Although there were several technical problems with the work presented by Zou et al. making it difficult to draw any conclusions, there is increasing evidence to suggest that ERK5 may share similar targets as ERK1/2 and have overlapping functions (Ranganathan et al., 2006, Kato et al., 1997, Kato et al., 1998).

In parallel to attempts to derive *Srf*fl/fl mES cells, strategies were devised to generate SRF deficient cells by targeting the *Srf* locus with a recombination mediated cassette exchange (RMCE) genetrap cassette (Figure 4.21). This vector has been constructed, and when correctly targeted it will generate a trapped allele and provide a tool to express SRF mutants from the endogenous locus. Moreover this vector can easily be modified to insert a different selection cassette to target the second locus, creating a constitutive null line that can be reverted by Flp expression. If the allele functions as predicted it could prove a valuable tool to study the function of SRF in the embryo also, by generating mice where SRF function can be rescued in various embryonic components using tissue-specific Flp mice.
Figure 4.21: RMCE targeting approach to study Srf function. Cartoon depicting the Srf locus targeted with the RMCE allele. In the presence of Cre with positive selection for the incoming plasmid, as well as negative selection against the targeted cassette exchange occurs allowing expression of a host or SRF mutants from the endogenous locus. Flp expression removes the cassette leaving only a single Frt site genomic footprint.
Chapter 5 - Identification and analysis of ERK2 interacting partners in mES cells

5.1 Introduction

ERK enzymes interact with a plethora of cytoplasmic and nuclear targets that fall into many categories with divergent function such as regulators of both transcription and translation, metabolism, apoptosis, and chromatin dynamics. The initial step in understanding ERK function in any given cell type is to identify the pathways in which it acts, what molecules it regulates, and what molecules regulate it. Several methods exist to determine ERK substrates, as well as interaction partners, each with its own strengths and weaknesses (discussed in section 1.6.2).

A quantitate IP-LC-MS/MS approach has recently been employed to address the long standing question as to how differential activation kinetics of ERK by either EGF or NGF affects its binding specificity in the PC12 model system (von Kregesheim et al., 2009) and demonstrates how applicable this approach is to studying kinase function. Von Kregesheim and colleagues immunoprecipitated endogenous ERK1 from SILAC labelled cells stimulated for different times with either EGF or NGF, and then quantified the relative differences in abundance between each interactor in a stimulus dependent context. The initial analysis identified 284 putative ERK1 binding partners. Upon induction of differentiation (NGF stimulation) 60 proteins changed their binding affinities with Erk so as to modulate key aspects of the signal cascade, such as duration, localization and cross talk with other signaling pathways. This large scale study lead to the notion that the functional outcome of ERK signaling in PC12 cell was dependent on control mechanisms distributed across the network, rather than a single bistable switch. This study represents the largest list of potential Erk interacting partners ever compiled, and illustrates the power of such an experimental design.

More recently still, work from the lab of Tony Pawson has employed a FLAG tagging approach to identify GRB2 binding partners in Hek293T cells (Bisson et al., 2011). By affinity purifying exogenously expressed FLAG-GRB2 from cells
either stimulated with various growth factors, or tyrosine phosphatase inhibited, and then subjecting the bead-bound fraction to MS analysis using the novel quantitate method, selective reaction rate monitoring, Bisson et al. were able to identify 108 proteins bound to GRB2 in a stimulus dependent manner. This study detected several membrane bound proteins as well as trans-membrane receptors, usually difficult to isolate with conventional IP/lysis buffers (protein solubilisation was preformed in the presence of 1% non-denaturing detergent (IGEPAL), and 0.5% sodium deoxycholate), underscoring the importance of protein solubilization and IP conditions in such a study.

5.1.1 Aim of this chapter

The aim of this chapter was to generate a ERK2 centered protein interaction network in mES cells and functionally characterize a subset of those interactions. At present there is a correlative connection between ERK1/2 activity and changes in the expression of pluripotence associated markers. Therefore it was hoped that by identifying molecules that interact with ERK2 it may be possible to better understand the relationship between ERK activity and the pluripotent state of mES cells.
5.2 Results

5.2.1 Optimization of the IP protocol

The initial IP parameter optimization was preformed on B1 derived Erk2 null cells ectopically expressing 3xFLAG tagged ERK2. Several lysis buffers were tested, though as there was no discernible difference between any of them in their ability to pull-down potential ERK2 specific interactors, as judged by silver staining SDS-PAGE resolved IP eluates (data not shown), the lysis buffer in Materials and Methods section 2.1.2 was used for all experiments described in this chapter as it was already compatible with downstream MS applications. This technique gave a high level of bait capture (Figure 5.1-A), with moderate background (Figure 5.1-B). However, a substantial amount of \(\alpha\)FLAG IgG came of the beads during acid elution, which could interfere with downstream analysis, therefore several elution methods were tested to see which gave the lowest IgG contamination. A comparison of the IgG:bait ratio from samples eluted with either phosphoric acid, glycine, or ammonia showed that elution under basic conditions, 0.5M NH\(_4\)OH (pH 11.5), 0.5mM EDTA, gave the cleanest results (Figure 5.2), and was hence used for all elutions thereafter, unless otherwise stated.

It was expected that many interactions between ERK2 and other proteins would be transient enzyme-substrate interactions, and as such represent a smaller proportion of the precipitated protein for each IP. Confidence in a protein detected by LC-MS/MS can depend on either a high percentage of peptide coverage in duplicate/triplicate experiments, or a consistent detection of a few peptides per protein in several experiments. Traditionally, samples are resolved by SDS-PAGE, coomassie stained, sliced into bands (usually between 5-8), the protein is digested ‘in gel’ before analysis with each band taking approximately 3 hours to analyze. The advantage to this approach is that the complexity of the IP is greatly reduced before analysis. As peptides will compete for ionization, and as ionization efficiency is a crucial factor in determining the \(m/z\) ratio of a peptide, there is an inverse relationship between sample complexity and sensitivity. Moreover, by separating samples in this manner,
the proportion of peptides that will be selected for a second MS analysis will increase, thus resulting in a higher proportion of peptide sequence determination per experiment. The disadvantages of such an approach is that sample loss and contamination increase with the manual handling steps. As does the possible bias and variation that can result from manual gel processing. Furthermore, the amount of MS time that can be allocated to a single experiment can be a financial consideration. Alternatively, Gel-free analysis consists of protein digestion, either on-bead or in solution, followed by direct LC-MS/MS analysis resulting in reduced processing time, reduced contamination from sample handling, and is highly amenable to high throughput approaches. A disadvantage is that because the total time each IP is subjected to MS/MS analysis, in highly complex samples, such as IPs, highly abundant proteins can dominate the run masking other proteins in the sample.

Therefore the efficiency of in-gel and gel-free analysis was estimated. Figure 5.3 shows a comparison of the proteins detected by either method. Although samples processed gel-free gave a higher sequence coverage of the bait, in-gel 52%, gel-free 69%, the total peptide number, as well as sequence coverage of other proteins identified was greatly reduced by gel-free analysis. This may have been due to technical reasons, as samples analysed gel-free had a large contaminant dominating the end of each run, interfering with peptide analysis which may have come from the basic elution conditions reacting with the bead matrix. Though both experiments were not conducted in parallel and are therefore not wholly comparable, it was felt that in-gel analysis was the most suitable approach as it allowed protein elution by peptide competition, reducing the possibility of bead based contaminant carryover.

Gene ontology analysis of proteins detected in both experiments described above showed a broad distribution across functional categories (Figure 5.4). However the confidence in hits dropped off significantly after known ERK interactors such as p90RSH, HSP90 and DUSP family members which are all know to form relatively stable cytosolic interactions with ERK (Dou et al., 2005, Britson et al., 2009, and reviewed in Anjum et al., 2008). Therefore the stability of ERK
interactions in mES cells was tested by size exclusion chromatography. Whole cell lysate from WT cells was fractionated over a Superdex 200 column at a flow rate of 0.5ml/min. Subsequent western blot analysis showed that ERK1/2 eluted as a monomer with an apparent mass between ~70 and 20kDa (Figure 5.5). ERK1/2 was also detected in the void fraction, however this may represent precipitated ERK containing protein aggregates, rather than actual macromolecular complexes. SEC analysis indicated that most ERK interactions were insufficiently stable to last the 45 minute run time, therefore in the hope of maintaining more transient ERK2 interactions the GFP-Trap® tagging system was tested. This system uses a GFP binding fragment derived from a llama single chain antibody that allows near complete bait depletion from a sample with incubation times between 5 and 30 minutes (Rothbauer et al., 2006, Rothbauer et al., 2008). C-terminal GFP-tagged Erk2 was constructed by PCR by overlapping extension with the internal primers E6GGF/R and the external primers NGFPR and XKE-F. The resulting PCR fragment was then cloned in pIPC37 as an XhoI/NotI fragment, and sequence verified. Erk2 null cells D7N2 and D2N4 were then electroporated with ScAl linearized DNA (20µg), selected in puromycin (1µg/ml) and clones were picked, expanded and ERK2-GFP expression, and localization were confirmed by western blot and fluorescent microscopy (data not shown). Lysates from ERK2-GFP expressing cells were made as usual and incubated with GFP-Trap beads for times ranging from 10 to 120 minutes. Western blot analysis of both precipitated as well as flow through fractions indicated that the efficiency of bait (ERK-GFP) and interactor (p90RSK) increased with incubation time (Figure 5.6 A-C). With a high proportion of the bait being depleted following 60 minutes incubation. A comparison of ERK2 immunoprecipitates using either FLAG or GFP-Trap systems showed little difference in the amount of proteins pulled down between either (Figure 5.6-D). Moreover, it appeared that there were nonspecific proteins pulled down using GFP-trap beads, not seen in the FLAG immunoprecipitates indicating that there are either bead or GFP based background differences between both systems. As previous experiments generated a comprehensive list of FLAG-bead based background contamination in
mES cells, and as the GFP based system failed to yield an obvious increase in protein pull-down efficiency, it was decided to use FLAG-tagged ERK2 for subsequent experiments.
Figure 5.1: Efficiency of Erk2 immunoprecipitation. A) Western blot analysis of IPs from either Erk2 null or FLAG-ERK2 expressing cells. 0.05% of the eluted IP was loaded. B) Silver staining of acid eluted samples. 10% of the eluted IP was loaded. Lysates were made from an 80% confluent 50 cm² dish.

Figure 5.2: A comparison of 4 elution methods. A) 50mM H₃PO₄ (pH1.8), B) 0.5M NH₄OH, 0.5mM EDTA (pH11.5), C) 200mM Glycine, 0.5M NaCl, 0.1% NP40 (pH 3.5), D) 0.2 mg/mL FLAG peptide, 20mM HEPES, 10% Glycerol, 100mM KCl, 1.5mM MgCl₂, 0.2mM EDTA (pH 7.4).
Figure 5.3: Comparison of ‘In-Gel’ versus ‘Gel-free’ methods of immunoprecipitate analysis. A) A list of proteins identified by LC-MS/MS analysis of ERK2 immunoprecipitates. A protein was considered a potential interactor if at least two unique peptides were identified in the positive IP, and no peptides in the negative control. B) Graph of A showing a clear skew along the X axis.
**Figure 4.** Peptides identified in ERK2 IPs by biological process. Proteins were manually curated from gene ontology data provided by MGI (http://www.informatics.jax.org/orthology.shtml).

**Figure 5.5: Gel-filtration analysis of ERK1/2.** Whole cell lysate from WT mES cells (500μg) was size fractionated on a Superdex 200 gel-filtration column. Fractions were collected and subjected to western blot analysis. Top blot: Analysis of every sixth fraction. Bottom blot: analysis of every second fraction between ~70 and 10kDa.
**Figure 5.6: Analysis of GFP-Trap® beads.** A) Western blot analysis of flow through samples from ERK2-GFP immunoprecipitates. B) Western blot analysis of precipitated ERK and p90RSK from ERK2-GFP immunoprecipitates 0.5% total eluate was used per sample. C) Silver staining of samples from A, 10% total eluate was used per sample. All IPs were eluted from the beads in 200mM Glycine (pH 3.5), 0.5M NaCl, 0.1% NP40.
5.2.2 Identification of ERK2 interacting partners in mES cells

The FLAG-ERK2 expressing Erk2 null clone D2N4-WT (Chapter 4, Figure 4.2.4) was used for large scale analysis of ERK2 interacting partners in mES cells. A total of 3 IPs were preformed for each FLAG-ERK2 expressing, and Erk2 null cells. Figure 5.2.6 shows the results of a representative experiment where 1% of each eluted fraction was separated by SDS-PAGE and silver stained. As there was still a considerable amount of bait present in the fourth elution, all eluates for each sample were combined, and processed as outlined in section 2.2.3.8.2. The approximate size of each band is shown in Figure 5.2.6. All IPs were preformed 2-3 days apart, but samples were processed and analysed by LC-MS/MS in parallel to minimize technical variation. MS/MS spectra were searched against the Mascot mouse database, and peek intensities were calculated using MaxQuant software to estimate relative protein abundance. Common contaminants such as keratins and serum proteins were immediately discarded from the list.

As there are still no universally accepted criteria by which to robustly discriminate between specific and non-specific interacting proteins from standard IP-MS/MS studies two separate approaches were used, one based on relative abundance of a protein between positive and negative controls (Malen et al., 2010) (Table 5.1), and one based on sequenced peptide number (Luke-Glaser et al., 2007) (Table 5.2). Analysis by peek intensity ratios/relative abundance is semi-quantitative and sensitive in that it will include peptides hits that were not selected for a second MS run due to random sampling in MS-1. However, this approach is highly susceptible to the inclusion of false positives due to differences in protein concentration/abundance between samples. Moreover it includes proteins identified in single biological replicates. Therefore to obviate the latter issue, proteins were included if the combined peek intensity was 20 fold greater in the positive sample than the negative and proteins detected in a single run were removed from the list. This level of stringency yielded 46 putative ERK2 interacting proteins, many of which contain D-domains (ERK-interacting) motifs, and phospho-serine/threonine sites (individual protein sequences were searched using the Scansite tool [http://scansite.mit.edu/]).
However as it was clear that the probability of a protein possessing one or more ERK binding motifs increased with protein length, this abundance of D-domains may not reflect any real biological significance but the fact that most bead based background is greater than 60kDa. As in previous experiments, proteins identified spanned a wide range of functional categories, however there were much fewer transcription factors identified than in previous experiments using embryo-derived Erk2 null cell lines. Moreover, the high-confidence hits from previous experiments HSP90ab1, HSP90aa1 and PHF3 were absent from this data set. Inclusion of proteins based on Mascot score (>20) and peptide number (at least two unique peptides per sample, presence in at least 2 samples, and no peptides detected in any of the negative control samples) yielded a much smaller list, as expected, mainly populated by known interactors from the p90RSK and DUSP families (Table 5.2). The data set was then searched for the presence of post-translational modifications. Phosphotyrosine, phosphothreonine, and phosphoserine were all specified in Mascot as variable modifications, as well as oxidation of methionine, and N-acetylation of the protein N-terminus. Known phosphorylation sites were detected for all the high-confidence hits (Table 5.3). Interestingly, the presence of both pTEY and TEpY peptides were detected for ERK2, indicating that ERK activation/deactivation proceeds in a processive manner in mES cells (Schilling et al., 2009), however no dual phosphorylated pTEpY peptides were detected in this set of MS experiments, or in the experiments described in section 5.2.1 (data not shown).

The interaction between ERK2 and p90RSK was confirmed by co-IP/western blotting and was specific to 3xFLAG-ERK2 expressing cells (Figure 5.8-A). The binding of ERK to p90RSK proteins has been reported to be stimulus dependent, where following stimulation ERK1/2 and p90RSK rapidly disassociate and reassociate 30-60 minutes post stimulation. This phenomenon has been observed in Hek293T cells stimulated with EGF (Roux et al., 2003), and PC12 cells treated with either EGF or NGF (von Kregsheim et al., 2009). This dissasociation is dependent on phosphorylation of Ser749 in the C-terminal ERK docking domain by the N-terminal kinase domain of p90RSK (Roux et al., 2003). Interestingly, no difference
in the binding of ERK2 to p90RSK was observed following acute stimulation with FGF2 in mES cells (Figure 5.8-B). This may be due to in-solution rebinding of ERK2 and p90RSK during the IP process, or may also indicate that the N-terminal domain of p90RSK may not be fully activated following FGF2 stimulation. However, this may be an artifact of stimulation with FGF2 rather than FGF4 as phosphopeptides corresponding to S227, the PDK1 phosphorylation site in the N-terminal kinase domain of p90RSK, were identified by MS studies (Table 5.3).
Figure 5.7: Analysis of eluted fractions from control and FLAG-ERK2 large scale immunoprecipitates for LC-MS/MS analysis. Samples were eluted by peptide competition in 100μl peptide elution buffer (0.2 mg/ml FLAG peptide, 20mM HEPES (pH 7.4), 10% Glycerol, 100mM KCl, 1.5mM MgCl₂, 0.2mM EDTA ) for 15 minutes per elution with mild agitation at 4°C.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Peek +ve</th>
<th>Peek-ve</th>
<th>Run detection</th>
<th>MAPK site</th>
<th>Phosphopeptide</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tubb2b</td>
<td>158970</td>
<td>0</td>
<td>5</td>
<td>No</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Tuba1a</td>
<td>483887</td>
<td>0</td>
<td>6</td>
<td>No</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Cse1l</td>
<td>160562</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>DNA Replication</td>
</tr>
<tr>
<td>Glud1</td>
<td>1247820</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Metabolism</td>
</tr>
<tr>
<td>Rsk1</td>
<td>407770000</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>Yes</td>
<td>Signaling</td>
</tr>
<tr>
<td>Rsk2</td>
<td>166897000</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>Yes</td>
<td>Signaling</td>
</tr>
<tr>
<td>Rab1</td>
<td>100170</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Hadhb</td>
<td>310760</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Metabolism</td>
</tr>
<tr>
<td>Prkcsh</td>
<td>304500</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Cct4</td>
<td>175274</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>Cct5</td>
<td>439868</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>Cct3</td>
<td>320170</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>Cdc37</td>
<td>466859</td>
<td>0</td>
<td>3</td>
<td>D-Domain</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Dusp6</td>
<td>274090</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>Yes</td>
<td>Signaling</td>
</tr>
<tr>
<td>Erk2</td>
<td>2851980000</td>
<td>298900</td>
<td>4</td>
<td>Bait</td>
<td>Yes</td>
<td>Signaling</td>
</tr>
<tr>
<td>Smarce1</td>
<td>301530</td>
<td>0</td>
<td>2</td>
<td>D-Domain, P-ST Kin</td>
<td>No</td>
<td>Epigenetic</td>
</tr>
<tr>
<td>Eif3k</td>
<td>380185</td>
<td>0</td>
<td>3</td>
<td>D-Domain</td>
<td>No</td>
<td>Translation</td>
</tr>
<tr>
<td>Rbpb7</td>
<td>98998</td>
<td>0</td>
<td>6</td>
<td>No</td>
<td>No</td>
<td>Epigenetic</td>
</tr>
<tr>
<td>Atxn10</td>
<td>503049</td>
<td>0</td>
<td>3</td>
<td>D-Domain</td>
<td>No</td>
<td>Miscellaneous</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Peak +ve</td>
<td>Peak-ve</td>
<td>Run detection</td>
<td>MAPK site</td>
<td>Phosphopeptide</td>
<td>Process</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>---------------</td>
<td>-----------</td>
<td>----------------</td>
<td>---------</td>
</tr>
<tr>
<td>Rsk3</td>
<td>15670000</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Eif3h</td>
<td>342540</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>Igf2bp1</td>
<td>265629</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>mRNA Processing</td>
</tr>
<tr>
<td>Slc2a3</td>
<td>3751390</td>
<td>0</td>
<td>5</td>
<td>D-Domain</td>
<td>No</td>
<td>Metabolism</td>
</tr>
<tr>
<td>Tubb2c</td>
<td>743930</td>
<td>0</td>
<td>5</td>
<td>D-Domain</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Acsl3</td>
<td>814330</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>No</td>
<td>Metabolism</td>
</tr>
<tr>
<td>CItc</td>
<td>66984</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Actb12</td>
<td>69797850</td>
<td>0</td>
<td>6</td>
<td>D-Domain</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Rsk4</td>
<td>11277100</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>Yes</td>
<td>Signaling</td>
</tr>
<tr>
<td>Gatad2a</td>
<td>60216</td>
<td>0</td>
<td>2</td>
<td>D-Domain, P-ST Kin</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Erp</td>
<td>2040200</td>
<td>90708</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Erk1</td>
<td>32682000</td>
<td>0</td>
<td>3</td>
<td>D-Domain</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Eif2s3x</td>
<td>131500</td>
<td>0</td>
<td>2</td>
<td>D-Domain, P-ST Kin</td>
<td>No</td>
<td>Translation</td>
</tr>
<tr>
<td>Psmd1</td>
<td>332490</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>Igh-4</td>
<td>16930600</td>
<td>586066</td>
<td>5</td>
<td>No</td>
<td>No</td>
<td>Miscellaneous</td>
</tr>
<tr>
<td>Cct2</td>
<td>275509</td>
<td>0</td>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>Hspa2</td>
<td>168390</td>
<td>0</td>
<td>6</td>
<td>D-Domain</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Fkbp5</td>
<td>2736740</td>
<td>0</td>
<td>3</td>
<td>No</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Dusp9</td>
<td>21293000</td>
<td>0</td>
<td>3</td>
<td>D-Domain, P-ST Kin</td>
<td>Yes</td>
<td>Signaling</td>
</tr>
<tr>
<td>Lrrfip1</td>
<td>839310</td>
<td>0</td>
<td>3</td>
<td>No</td>
<td>No</td>
<td>Transcription</td>
</tr>
<tr>
<td>Klf20b</td>
<td>248380</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Ccdc88c</td>
<td>132169</td>
<td>0</td>
<td>2</td>
<td>D-Domain, P-ST Kin</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Tbrg4</td>
<td>328915</td>
<td>0</td>
<td>3</td>
<td>D-Domain</td>
<td>No</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>Peg10</td>
<td>663709</td>
<td>0</td>
<td>3</td>
<td>D-Domain</td>
<td>No</td>
<td>Transcription</td>
</tr>
<tr>
<td>Flnb</td>
<td>5622420</td>
<td>0</td>
<td>5</td>
<td>D-Domain</td>
<td>No</td>
<td>Structural</td>
</tr>
<tr>
<td>Tpr</td>
<td>119779</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Signaling</td>
</tr>
<tr>
<td>Rps5</td>
<td>148971</td>
<td>0</td>
<td>2</td>
<td>D-Domain</td>
<td>No</td>
<td>Protein metabolism</td>
</tr>
</tbody>
</table>

Table 5.1: Proteins identified by LC-MS/MS as putative ERK2 interactors judged by peak intensity values. Inclusion in the list was based on the following criteria; 20-fold greater peak intensity between positive and negative samples; a protein must have been identified in at least 2 runs. ERK motif searches were preformed using the Scansite algorithm, set to medium stringency. Where multiple isoforms were identified for a protein the sequence of longest isoform was used. Gene ontology lists were curated manually as in Figure 5.2.3.
Table 5.2: Proteins identified by LC-MS/MS as putative ERK2 interactors judged by peptide abundance. Inclusion in this list was based upon the following; a Mascot score of >20 per peptide; two unique peptides per run; the protein must have been identified in at least 2 of the 3 experiments; no peptides identified in the negative control. Note the presence of 2 ERK2 peptides in the Erk2 null immunoprecipitate (red). This may have resulted from contamination during sample preparation, or peptide carry-over during MS runs. The relatively low standard error for most hits indicates the reproducibility of the experiment.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Probability</th>
<th>Mascot</th>
<th>Residue</th>
<th>Motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erk2</td>
<td>VADFHDHTGFLTEY(ph)</td>
<td>21.5</td>
<td>Y185</td>
<td>Mek</td>
</tr>
<tr>
<td>Erk2</td>
<td>VADFHDHTGFLITEY(ph)</td>
<td>21.5</td>
<td>T183</td>
<td>Mek</td>
</tr>
<tr>
<td>Rsk1</td>
<td>Tph</td>
<td>PRDSPGIPPSAGAHQFLR</td>
<td>25.66</td>
<td>S299</td>
</tr>
<tr>
<td>Rsk2</td>
<td>TPKDS(ph)PPIPSSANGHQLFR</td>
<td>36.95</td>
<td>S369</td>
<td>Mapk</td>
</tr>
<tr>
<td>Rsk4</td>
<td>GFS(ph)</td>
<td>FVSAIAEYK</td>
<td>63.14</td>
<td>T305</td>
</tr>
<tr>
<td>Dusp6</td>
<td>S(ph)</td>
<td>SNIS(ph)PNFMGQLLDFER</td>
<td>40.09</td>
<td>S331/S328</td>
</tr>
<tr>
<td>Dusp9</td>
<td>YILNVT(ph)PNLPNLFEK</td>
<td>50.4</td>
<td>T305</td>
<td>Mapk</td>
</tr>
</tbody>
</table>

Table 5.3: Post-translationally modified peptides identified in this study. The position of the modification is given in the second column. If multiple potential phosphorylation sites are present in a peptide the probability for each amino acid is given in column 3. The kinase corresponding to the phosphorylation site was determined by searching the literature.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Probability</th>
<th>Mascot</th>
<th>Residue</th>
<th>Motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erk2</td>
<td>VADFHDHTGFLTEY(ph)</td>
<td>21.5</td>
<td>Y185</td>
<td>Mek</td>
</tr>
<tr>
<td>Erk2</td>
<td>VADFHDHTGFLITEY(ph)</td>
<td>21.5</td>
<td>T183</td>
<td>Mek</td>
</tr>
<tr>
<td>Rsk1</td>
<td>Tph</td>
<td>PRDSPGIPPSAGAHQFLR</td>
<td>25.66</td>
<td>S299</td>
</tr>
<tr>
<td>Rsk2</td>
<td>TPKDS(ph)PPIPSSANGHQLFR</td>
<td>36.95</td>
<td>S369</td>
<td>Mapk</td>
</tr>
<tr>
<td>Rsk4</td>
<td>GFS(ph)</td>
<td>FVSAIAEYK</td>
<td>63.14</td>
<td>T305</td>
</tr>
<tr>
<td>Dusp6</td>
<td>S(ph)</td>
<td>SNIS(ph)PNFMGQLLDFER</td>
<td>40.09</td>
<td>S331/S328</td>
</tr>
<tr>
<td>Dusp9</td>
<td>YILNVT(ph)PNLPNLFEK</td>
<td>50.4</td>
<td>T305</td>
<td>Mapk</td>
</tr>
</tbody>
</table>

Figure 5.8: ERK2 and p90RSK stably interact following FGF stimulation. A) Confirmation of the ERK2:p90RSK interaction in mES cells. B) Samples were stimulated as usual and
lysates were incubated with 5μl M2 FLAG agarose with end over end rotation for 3 hours at 6°C. Bead-bound complexes were recovered by boiling the beads in Laemmli buffer and eluates were subsequently analysed by western blotting.
5.2.3 The role of p90RSK proteins in MAPK signaling in mES cells

As the analysis of ERK2 IP-MS/MS experiments failed to generate any novel high confidence interactors, nor any interactions between ERK2 and pluripotency-associated molecules, the functional significance of ERK:p90RSK interactions was selected for further analysis. To do this I made use of a mutant mES cell line generated in the lab of Dr. Dario Alessi where exons 3 and 4 of Pdk1 have been flanked by loxP sites and subsequently removed by Cre excision (Williams et al., 2000). This cell line has no detectable p90RSK activity, but has been reported to exhibit no defect in neural specification (Stavridis et al., 2007), however other lineages were not tested. The rational behind this approach was that as p90RSK activity has been implicated in IEG induction (De Cesare et al., 1998), if the IEG response is unaffected in PDK mutant cells, p90RSK proteins may not be essential for mediating the FGF/ERK transcriptional response. However, if IEG induction is abrogated in the absence of p90RSK activity, as Pdk1 mutant cell exhibit no overt differentiation defect, this transcriptional network may not be necessary for lineage specification.

Pdk1 null cells were transgenically rescued with either FLAG-tagged WT PDK1, or PDK1-K114G mutant, where a lysine in the ATP binding domain is replaced with glycine and is reported to act as a dominant negative (Arico et al., 2002). PDK1 was cloned from E14tg2a total RNA with the primers XKFPF/NPDKR, TOPO cloned and sequence verified. The K114G mutation was introduced by site-directed mutagenesis with the primers PKGF/R. Both sequences were cloned as XhoI/NotI fragments into pIPC37, and confirmed by diagnostic digestion. FspI linearized plasmid (25µg) was electroporated into Pdk1 null mES cells, selected in puromycin for 8 days, after which time stable transfectants were collected as a pool for further analysis.

As reported previously, Pdk1 null mES cells exhibit enhanced ERK1/2 phosphorylation. Expression of WT-PDK1, and to a lesser extent PDK1-K114G, was sufficient to reduce ppERK levels, as well as ERK dependent p90RSK phosphorylation (T359/S363) (Figure 5.9-A). This confirmed that exogenous PDK1
was functional and that the enhanced ppERK activity in this cell line was PDK1 dependent. Next the effect of PDK1 abrogation on IEG induction was tested. Cells were stimulated as usual and the extent of *cFos*, *Egr1* and *Egr2* expression at 1 and 3 hours was monitored by Q-PCR. Surprisingly, loss of PDK1 activity, and hence p90RSK activity, actually enhanced IEG induction, compared to WT-PDK1, or PDK1-K114G transfected cells (Figure 5.9 B-D). As the K114G is expressed at lower levels than WT PDK1, the lack of comparable IEG reduction between WT and mutated PDK1 transfected cells does not indicate that this mutant is functioning as a dominant negative, and may just be a consequence of reduced activity due to reduced expression levels. Taken together this indicates that p90RSK activity is not essential for IEG induction in mES cells, however, a p90RSK dependent negative feedback loop, such as the p90RSK-SOS loop in PC12 cells (Douville et al., 1997) may be active.

Next, the differentiation potential of the *Pdk1* mutant cells was tested to determine if PDK1, or p90RSK activity was essential for commitment to any specific lineage. PDK-null, PDK-WT and PDK-K114G cells were cultured in hanging drops for two days (-Lif), and then transferred as aggregates into floating culture for a further 10 days. After approximately 8 days in culture rhythmically beating cardiac muscle was evident in all cultures as well as cystic embryoid bodies (Figure 5.10-A). The extent of differentiation into mesoderm and endoderm lineages was then assessed by Q-PCR analysis. Embryoid bodies from all genotypes downregulated the pluripotence markers *Oct4* and *Nanog* to a similar extent, as well as upregulating markers of endoderm (*Gata4/6*), and mesoderm (*Acta2* and *Pdgfr1a*) (Figure 5.10-B), though *Gata4* expression was greatly decreased in the *Pdk1* mutant derived EBs.

The ability of *Pdk1* null cells to generate Sox1<sup>+</sup> neural precursors has been shown previously (Stavridis et al., 2007), in fact, PDK1 deficient cells differentiate into neural precursors at nearly three times the efficiency of WT cells (Stavridis et al., 2007-Figure 1C). *Pdk1<sup>-/-</sup>* embryos progress through early gastrulation relatively normally, however at E8.5 mutant embryos lack somites, dorsal root ganglia, forebrain and posterior region development, indicating that there
may be a requisite for PDK1 activity for the generation of certain neural tissue (Lawlor et al., 2002). Therefore the ability of Pdk1 null mES cells to generate forebrain type neurons was investigated in vitro. Monolayer culture in N2B27 alone resulted in a high portion of β-III Tubulin, Ngn2, and Mash1 positive cells in both PDK-WT and PKD-K114G expressing cell lines, which were undetectable in PDK1 knockout cultures (Figure 5.11). Ngn2 and Mash1 have been shown to mark cells of the forebrain in the developing embryo (Fode et al., 1999), indicating neuronal differentiation is also PDK1 dependent in vitro, however downregulation of the pluripotency genes Oct4 and Nanog indicate that there is no general block in differentiation under these conditions (Figure 5.11-B). Furthermore differentiation in the presence of a novel p90RSK inhibitor BI-D1870 (Sapkota et al., 2007), at a concentration reported to inhibit p90RSK activity in other cell types (Sapkota et al., 2007, Chen et al., 2009) indicated that the block in neuronal differentiation is not p90RSK dependent (Figure 5.12).
Figure 5.9: PDK acts to restrain ERK activity in mES cells. A) Western blot analysis of PDK<sup>-/-</sup> (lanes 1-3), PDK<sup>-/-</sup>(WT PDK1) (lanes 4-6), and PDK<sup>-/-</sup>(PDK1K114G) (lanes 7-9) cell lines. Cells were analysed under normal culture conditions (LIF+), under FGFR inhibition (Lif+ PD17+), or following 30 minutes acute FGF2 stimulation (LIF+ FGF2+). The membrane was stripped and sequentially blotted with the antibodies as indicated. The low total ERK1/2 signal in lane 3 is due to band bleaching by the previous ppERK blot. B-D) Q-PCR analysis of IEG induction following acute FGF2 stimulation for the indicated times. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
Figure 5.10: PDK1 is dispensable for mesoderm and endoderm specification in vitro.

A) Phase contrast images embryoid bodies following 12 days of differentiation. Images are taken at 20x magnification. B) Q-PCR analysis of the embryoid bodies in A for markers of pluripotence (Oct4, Nanog), endoderm (Gata4/6), and mesoderm (Acta2, PdgfR1a). Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.
**Figure 5.11: Neural differentiation is PDK1 dependent.** A) Immunohistochemistry analysis of \( PDK^+ \), \( PDK^-(WT\ PDK1) \), and \( PDK^-(PDK1K114G) \) cells following 7 days neural differentiation. Cultures are stained for β-III Tubulin (yellow), and DNA is counterstained with DAPI in blue. Results are representative of three separate experiments conducted at two different starting densities. B) Q-PCR analysis of cultures from A. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.

**Figure 5.12: Neuronal differentiation is p90RSK independent.** Immunohistochemistry analysis of \( PDK^+(WT\ PDK1) \) cells following 7 days differentiation in the presence or absence of 2μM BI-D1870. Cultures are stained for β-III Tubulin (red), Nestin (green) and DNA is counterstained with DAPI in blue.
5.2.4 The effect of the p90RSK inhibitor BI-D1870 in MAPK activity and mES cell differentiation

Data from Pdk1 mutant mES cells indicated that p90RSK activity may be involved in a negative feedback loop that regulates ERK activity, possibly through hyperactivation of the SOS-RAS-RAF-MEK core. If p90RSK activity fed back to the top of the signaling cascade then this may represent a point where the FGF-MEK pathway could be chemically hyperactivated and as such may prove a valuable tool in studying MAPK feedback regulation, as well as a potential role in increasing the efficiency of various mES cell differentiation assays. Therefore, using ppERK and IEG induction as readouts, the effect on p90RSK specific inhibition on the FGF-MEK pathway was assessed. Addition of BI-D1870 directly to WT mES cells resulted in a rapid, but highly transient induction of ppERK (Figure 5.13-A). A similar rapid but transient ppERK1/2 induction was observed using Fgf4-/- cell (Figure 5.13-B) however, ERK1/2 phosphorylation was more sustained having declined to basal levels by 120 minutes. However, acute FGF2 stimulation WT mES cells in the presence of BI-D1870 resulted in a modest increase in ppERK1/2 levels, if any, compared to treatment with FGF2 alone (Figure 5.14-A), with a difference in cumulative ppERK1/2 levels of 15% across the 120 minute window analysed (Figure 5.14-B). Analysis of the effect of BI-D1870 treatment on FGF2 induced IEG expression showed a dramatic increase in IEG induction with induction nearly ten fold that seen in cells stimulated in the absence of BI-D1870 (Figure 5.15). This level of induction is far greater that that observed in Pdk1-/- cells, indicating that there may be factors other than p90RSK inhibition contributing to the heightened IEG response in the presence of BI-D1870.

Next it was tested if BI-D1870 treatment affected mES cells self-renewal. Cells were plated at medium density, 3 x 10^4/cm^2, in complete medium, and the following day LIF was withdrawn in the presence or absence of BI-D1870. Gene expression was monitored at over a 28 hour window. LIF withdrawal in the presence of BI-D1870 resulted in a heightened loss of Nanog expression, however expression of other pluripotency genes such as Tbx3 and Stella decreased in a similar fashion
regardless of BI-D1870 treatment (Figure 5.16-A-C). It was then asked if BI-D1870 treatment effected the self-renewal capacity of mES cells at clonal density. After 7 days in culture, no AP positive cells were detectable in cultures treated with BI-D1870, regardless of the presence of LIF (Figure 5.16-D). However, substantially fewer colonies survived BI-D1870 treatment, indicating a high degree of toxicity associated with longterm culture in this drug at clonal density. Therefore the presence of only differentiated colonies following BI-D1870 treatment may be a result of selection for a cell type that can tolerate treatment at this concentration, and not enhanced differentiation due to activation of the MAPK pathway.
Figure 5.13: BI-D1870 treatment activates ERK1/2 in mES cells. A) Western blot analysis of BI-D1870 treated (3 μM) mES cells. WT cells were plated in duplicated overnight in complete medium. The following day BI-D1870 was diluted in conditioned media from one plate and cells were treated for the indicated times. Membranes were stripped and sequentially reprobed as indicated. B) As in A except Fgf4⁻/⁻ cell were used. As a positive control Fgf4⁻/⁻ cells were incubated overnight with complete medium supplemented with recombinant FGF2 (10ng/ml) (lane 1).

Figure 5.14: BI-D8170 treatment only slightly increases Erk phosphorylation following Fgf stimulation. A) Western blot analysis of WT mES cells following acute FGF stimulation for the indicated times in the presence (bottom) or absence (top) of 3μm BI-D1870. Cells were pretreated ± BI-D1870 for 24 hours before stimulation, then stimulated ±BI-D1870 and Fgf2 (10ng/ml). B) Calculation of the cumulative ppERK response from the western blots shown in A. ppERK1/2 levels were quantified, normalized to total ERK1/2 and plotted as a curve. The time-integrated response was calculated as the area under each curve.
Figure 5.15: BI-D1870 treatment facilitates increased induction of IEGs following Fgf2 stimulation. Q-PCR analysis of IEG induction in the presence of BI-D1870. Cells were incubated overnight in complete medium in the presence of 100nM PD173074 ± 3 μM BI-D1870. The following day, the inhibitors were washed away and cells were treated with FGF2 (10ng/ml) ± 3 μM BI-D1870 for the indicated times. The x-axis values are graphed to a log10 scale. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel.

Figure 5.16: BI-D1870 attenuates the self-renewal of mES cells. A-C) Q-PCR analysis of pluripotency-associated gene expression following Lif withdrawal ± 3μM BI-D1870 for the indicated times. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. D) Analysis of the self renewal efficiency of mES cells ± Lif, ± 3μM BI-D1870. Cells were plated as indicated and analysed for alkaline phosphatase activity after 7 days. Error bars denote standard error from 2 experiments conducted in parallel.
However, as BI-D1870 treatment did not appear to affect the survival of mES cells at higher densities (Figure 5.12), possibly due to reduction in the effective concentration if free drug at higher cellular densities, the effect of BI-D1870 treatment on neural differentiation was tested. It was rationalized that if enhanced FGF/MEK activity would positively regulate neuronal differentiation, that would happen at an early stage, and hence short-term BI-D1870 treatment, ~24 hours should be sufficient to test this, as FGF activity has been reported to be required for the first 17-24 hours of differentiation (Stavridis et al., 2007). To test this 46C Sox1-GFP reporter cells (Ying et al., 2003) were treated ± 3μM BI-D1870 for 24 hours and the extent of neural induction was monitored by flow cytometry over the course of 7 days (Figure 5.17). BI-D1870 is a naturally fluorescent compound which accounts for the fluorescent peaks seen at day 1 and day 2. However, by day 3 the drug has washed out and an obvious increase in GFP expression was detected in BI-D1870 treated cells compared to untreated. As Sox1-GFP expression is transient during neural differentiation under monolayer conditions (Chung et al., 2004), it is difficult to determine at time-points later that day 3 the effect of BI-D1870 as there will be a mixture of cells in the population transiting from Sox1\(^{+ve}\) to Sox1\(^{-ve}\) and vice versa, therefore the expression of both Oct4 and Ngn2 were monitored over the 7 days of differentiation. Q-PCR analysis showed no significant difference in the levels of either Ngn2 or Oct4 expression by day 7, though as Ngn2 expression was only evident by day 7 (Figure 5.18-A), analysis of later time points may prove more informative. However immunohistochemistry analysis showed little difference between both cultures ± BI-D1870 (Figure 5.18-B). As neural induction proceeds at a high efficiency under these conditions (Ying et al., 2003), the lack of any obvious effect of BI-D1870 treatment on the later stages of differentiation may be in part due to a requirement of other cell types in the culture required to create a niche suitable for the survival of neural cells, and thus a maximal ratio of neural to non-neural cells may already be reached under standard monolayer differentiation conditions. Furthermore, the action of other signaling pathways such as Notch known to affect the heterogeneity of neural commitment may be a contributing factor (Lowell et al.,
2006). It would be therefore interesting to assess the effect of BI-D1870 treatment of other, less efficient differentiation protocols such as mesoderm (Tada et al., 2005) or endoderm (Morrison et al., 2008). Though the toxicity that is associated with BI-D1870 treatment is a concern, this study highlights how targeting the mediators of MAPK negative feedback can enhance this pathway and play a positive role in lineage specific differentiation protocols. Analysis of other specific p90RSK inhibitors such as Sl1010 (Smith et al., 2005), or FMK (Cohen et al., 2005) would help to confirm that the results presented in this chapter are due to p90Rsk inhibition, or a consequence of off-target interactions.
Figure 5.17: BI-D1870 enhances Sox1-GFP promoter activity. A time course of flow cytometry analysis of 46C Sox1-GFP reporter cells undergoing neural induction. Red histogram: WT mES cells, blue histogram: 46C mES cells, green histogram: 46C cells plated in 3µM BI-D1870 for 24 hours (day1) then replenished with N2B27 for the remainder of the assay.

Figure 5.18: BI-D1870 treatment has little effect on the efficiency of mES cell neuronal differentiation A) Q-PCR analysis of a time course of neural induction ± 3µM BI-D1870 (for 24 hours) for the neural marker Ngn3, and the pluripotency marker Oct4. Values are normalized to TBP. Error bars denote standard error from 2 experiments conducted in parallel. B) Immunohistochemical analysis of cultures following 7 days of neural induction 3µM BI-D1870 (for 24 hours). Cultures are stained for β-III Tubulin (red), NESTIN (green) and DNA is counterstained in blue.
5.3 Discussion

5.3.1 The use of IP-MS/MS as a tool to study ERK2 interactions in mES cells

Work presented in this study represents the first attempt to identify the ERK2 interactome in mES cells. In doing this I employed an affinity tagging approach in a cell line where all the ERK2 protein was exogenous and FLAG-tagged. This allowed FLAG immunoprecipitates from ERK2-deficient cells to be used as a negative reference to identify non-specific binding proteins and aid the robust discrimination between bead/FLAG based background, and ERK2 specific interacting proteins. A partial list of common FLAG-agarose bead based background proteins is presented in appendix Table A1 and Figure A15 that may prove useful to others in analyzing similar experiments. Because of the transient nature of ERK2 interactions, several aspects of the IP protocol were optimized including lysis conditions, tagging strategies, elution conditions, as well as parameters to minimize bead/antibody based background. However, without subsequent MS analysis it was difficult to judge by silver staining alone the effect such optimization had on bait/interactor capture.

Although IP-MS approaches have been used successfully to determine the interaction networks of various pluripotency associated transcription factors in mES cells such as NANOG (Wang et al., 2006) and OCT4 (van den Berg et al., 2010), its application to the study of kinase-substrate interactions was found to be limited by several technical factors. The transient nature of these interactions is a primary concern, and although theoretically such interactions may form in-solution during the IP procedure, because many of the dynamics of MAPK signaling are based on the subcellular localization and the involvement of various scaffolding complexes (Kolch et al., 2005), this may not be the case. As all the high confidence hits identified in this study are known to interact stably with ERK2, approaches that focus on identifying bona fide substrates should be considered for further experiments.

The use of analogue-substrate mutants is an elegant solution to this problem (Eblen et al., 2003, Allen et al., 2007). This approach is based on the ability
of various ERK molecules, harboring mutations in their ATP-binding domain, to selectively bind ATP analogues. The ATP analogue is phospho-labelled, either isotopically or chemically, and thus as ERK phosphorylates its substrate it also labels it, which can be detected by a combination of 2D-gel electrophoresis and autoradiography (Ebeln et al., 2003), or immunoprecipitation with an antibody specific to that modified phospho moiety (Allen et al., 2007). This allows direct identification of substrates from a complex mixture of proteins thus circumventing the issue of transient substrate interactions as well as substoichiometric enzyme:substrate ratios. Moreover it prevents the identification of non-physiological interactions between proteins and kinases that are located in various cellular compartments that although can interact, would never interact in the context of the cell type being studied. Such approaches are compatible with pre-purification chromatographic steps allowing the researcher an extra degree of resolution. Although this technology is still in its infancy and there are some concerns as to the quality of the thiophosphate antibody used, the method developed by Allen and colleagues would currently be the most applicable to the mES system developed in this study as:

• It requires a test cell line that is deficient of the kinase of interest.
• IP-MS conditions are already optimized.
• All the reagents are now commercially available.
• Does not involve the optimizing of 2D separation techniques.
• Genetrap targeting constructs to express Erk2 from the endogenous locus are already constructed allowing endogenous expression of the analogue-specific ERK2 protein.

Another such approach would be to use phosphopeptide enrichment methods to identify shifts in the mES cell phosphoproteome in response to various stimulations and inhibitions. As illustrated by Kozako et al. (2009), when phosphopeptide enrichment is combined with a tool such as DiGE that can discriminate between differentially phosphorylated proteins in the control and test samples, it is possible to identify many bona fide ERK substrates, and minimize the
issue of abundant phosphopeptides dominating the MS analysis. Although this method may be technically more involved than an IP-MS/MS approach, and would need to be developed largely from scratch, it does allow a dynamic experimental approach that benefits from some very important conceptual considerations. At present the majority of our knowledge as to the function of MAPK signaling in mES cells comes from inhibitor data and the proof of principal experiments involved in 2i culture (Ying et al., 2008). Approaches that focus specifically on ERK2 are biased by the fact that we still have no categorical proof that the pro-differentiation FGF signal is transmitted exclusively through ERK1/2. Focusing on a top-down, phosphoproteome profiling method would allows the effect of 2i culture to be assessed in a largely unbiased manner, and from there candidate effectors/targets of the FGF signal can be determined and functionally assessed. As our understanding of the molecular mechanisms by which 2i culture maintains mES cells in a naive state is still lacking, such a study would be beneficial to the field at large.
5.3.2 Molecules that interact with ERK2 in mES cells

When this project was initially conceived there were few large scale datasets publicly available on ERK2 interacting partners. Recent work from the lab of Walter Kolch identified a staggering 284 potential binding ERK1 binding partners in PC12 cells many of which change their binding affinities in a stimulus dependent manner (Von Kregsheim et al., 2009). A comparison between both datasets from this study (ERK2 IPs from either embryo-derived or targeted parental lines), and the complete list of ERK1 interactors compiled by Von Kregsheim et al. (i.e. results from starved, EGF and NGF stimulated cells), showed a 20% overlap between putative ERK interacting proteins in mES cells and PC12 cells (Table 3.1). Moreover, most high confidence hits from this study were also identified in PC12 cells. Experiments were undertaken to analyze ERK2 interaction proteins following acute stimulation though due to time and technical concerns it was not possible to complete this experiment during this study.

<table>
<thead>
<tr>
<th>Erk2</th>
<th>Erk1</th>
<th>Rsk1</th>
<th>Rsk2</th>
<th>Rsk3</th>
<th>Rsk4</th>
<th>Hsp90ab1</th>
<th>Hsp90aa1</th>
<th>Cdc37</th>
<th>Dusp9</th>
<th>Er</th>
<th>Tpr</th>
<th>Flap5</th>
<th>Hspa2</th>
<th>Cct1</th>
<th>Cct2</th>
<th>Cct3</th>
<th>Cct4</th>
<th>Cct5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embryo Derived</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Targeted</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

Table 5.4: A comparison of Erk2 interactors identified in this study, with proteins identified by Von Kregsheim et al., (2009). Tick marks indicate the presence of the protein in the PC12 data set.

5.3.2.1 ERK1/2

It has been proposed that ERK1 and ERK2 could interact as heterodimers that may regulate their activity (Pelech, 2006). Though gel-filtration and sedimentation experiments have shown that ERK2 forms a homodimer whose dissasociation constant is dependent on the phosphorylation status of the monomers (Khokhlatchev et al., 1998, and reviewed in Cobb et al., 2000), similar experiments have yet to show interactions between ERK1 and ERK2. Using GFP-ERK2, that
migrates at ~60kDa, it was not possible to confirm the ERK1:ERK2 interaction (data not shown). Furthermore, size exclusion chromatography indicated that ERK2 eluted as a monomeric species, therefore indicating that the ERK2 homo-(ERK1) heretodimerisation is either unstable, or tethered to an insoluble cellular component in mES cells.

Interestingly, although mono-phosphorylated pTEY and TEpY peptides were identified in every MS experiment, dual-phosphorylated pTEpY peptides were never detected. It is not clear if this is due to a processive mode of activation or high DUSP activity in mES cells (phosphorylated DUSP6 and DUSP9 peptides were readily detected in every experiment), though this observation does have implications for analysis of ERL1/2 activity in mES cells. Table 5.2 shows a partial list of the phospho-ERK1/2 antibodies used in mES cell studies, that vast majority of which recognize both mono- and dual-phosphorylated forms of ERK1/2. It has been reported that mono-phosphorylated ERK2 possesses and intermediate kinase activity, 2-3 orders of magnitude higher than un-phosphorylated, but 1-2 orders of magnitude lower than dual-phosphorylated ERK2 (Zhou et al., 2002). Structural analysis of mono-phosphorylated ERK1 (TEpY) shows that it represents a molecule conformationally distinct from either un- or dual-phosphorylated ERK1 (Kinoshita et al., 2008), and may therefore may also be functionally distinct. This being said, this MS study indicated that ERK2 activity is low in self-renewing mES cells, and that use of antibodies that detect multiple phosphorylated forms of ERK1/2 may not accurately reflect relative activity of ERK isozymes.
<table>
<thead>
<tr>
<th>Antibody</th>
<th>Epitope</th>
<th>Vendor</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>9101</td>
<td>pTEY, TEpY, pTEpY</td>
<td>Cell signaling</td>
<td>Ying et al., 2008, Chan et al., 2003, Lu et al., 2008.</td>
</tr>
<tr>
<td>4377</td>
<td>TEpY, pTEpY</td>
<td>Cell signaling</td>
<td>Stavridis et al., 2007, Stavridis et al., 2010, Kunath et al., 2007, Watanabe et al., 2007, Yang et al., 2006, Li et al., 2006, Lee et al., 2006.</td>
</tr>
<tr>
<td>4376</td>
<td>TEpY, pTEpY</td>
<td>Cell signaling</td>
<td></td>
</tr>
<tr>
<td>4370</td>
<td>TEpY, pTEpY</td>
<td>Cell signaling</td>
<td></td>
</tr>
<tr>
<td>D13.14.4E</td>
<td>pTEY, pTEpY</td>
<td>Cell signaling</td>
<td>This study</td>
</tr>
</tbody>
</table>

Table 5.5: Antibodies used to detect phospho-ERK1/2 in mES cell studies.

### 5.3.2.2 TPR

Translocated promoter region protein (TPR) has been identified previously as both a substrate for, and a protein that stably interacts with ERK2 (Eblen et al., 2003, Vomastek et al., 2008, Von Kregsheim et al., 2009). TPR was originally identified as a transcript fused to the common proto-oncogenes MET and RAF (King et al., 1988). Subsequent work identified it as a component of nuclear pore complexes involved in export of proteins containing a leucine-rich nuclear export signal (Fosst et al., 2002), and then found to exert its function through CRM1 (Bene-Efraim et al., 2009). Peptide mapping identified four sites (T2102, T2123, S2141, and T2200) located at the C-terminal portion of TPR that may be ERK phosphorylated, and that phosphorylation of these sites stabilized the ERK2/TPR interaction. This positive feedback as well as the data that suggests that depletion of TPR as well as over expression result in an attenuation of ERK2 nuclear location suggests that TPR acts as a scaffold for ERK2 at the nuclear pore where it can be tethered so as to phosphorylate other nuclear pore components such as NUP50, NUP155 and NUP214 to regulate their activity (Kosake et al., 2009), and also to selectively phosphorylate cargo transiting to and from the nucleus.

Although TPR is expressed ubiquitously during development (Table 5.2), NUP proteins show a more dynamic expression pattern. Knockout studies show an
essential role for NUP214 for cellular viability with embryonic lethality at E4.0-E4.5 once maternal stores NUP214 transcripts are depleted. Moreover, it was not possible to generate Nup214−/− mES cells (van Deursen et al., 1996). The accumulation of polyadenylated RNA in the nucleus as well as a general impairment of nuclear transport that coincided with G2 arrest and cell death implies a generic role for NUP214 in nuclear transport. However the complex phenotypes presented by Nup50 (Smitherman et al., 2000) and Nup155 (Zhang et al., 2008) mutants indicate more specialized roles for ERK2 targeted nuclear pore components. Interestingly mES cells depleted for the putative ERK target NUP153 exhibit various differentiation defects including a dramatic reduction in the ability to generate post-mitotic neurons, to downregulate pluripotency-associated transcription factors during in vitro differentiation, and highly variable chimera contribution all indicating a role for nuclear pore function during lineage specification (Porter et al., 2010, Lupu et al. 2008).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Tpr</th>
<th>Nup50</th>
<th>Nup153</th>
<th>Nup214</th>
<th>Nup155</th>
</tr>
</thead>
<tbody>
<tr>
<td>oocyte</td>
<td>154</td>
<td>154</td>
<td>411</td>
<td>823</td>
<td>0</td>
</tr>
<tr>
<td>unfertilized ovum</td>
<td>963</td>
<td>253</td>
<td>0</td>
<td>253</td>
<td>50</td>
</tr>
<tr>
<td>zygote</td>
<td>73</td>
<td>0</td>
<td>146</td>
<td>330</td>
<td>0</td>
</tr>
<tr>
<td>cleavage</td>
<td>443</td>
<td>258</td>
<td>36</td>
<td>406</td>
<td>36</td>
</tr>
<tr>
<td>morula</td>
<td>2789</td>
<td>159</td>
<td>53</td>
<td>0</td>
<td>1142</td>
</tr>
<tr>
<td>blastocyst</td>
<td>432</td>
<td>43</td>
<td>86</td>
<td>86</td>
<td>101</td>
</tr>
<tr>
<td>egg cylinder</td>
<td>165</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>82</td>
</tr>
<tr>
<td>gastrula</td>
<td>72</td>
<td>181</td>
<td>36</td>
<td>108</td>
<td>0</td>
</tr>
<tr>
<td>organogenesis</td>
<td>805</td>
<td>446</td>
<td>303</td>
<td>143</td>
<td>223</td>
</tr>
<tr>
<td>fetus</td>
<td>170</td>
<td>185</td>
<td>52</td>
<td>82</td>
<td>61</td>
</tr>
<tr>
<td>neonate</td>
<td>125</td>
<td>77</td>
<td>87</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>juvenile</td>
<td>175</td>
<td>147</td>
<td>58</td>
<td>57</td>
<td>34</td>
</tr>
<tr>
<td>adult</td>
<td>120</td>
<td>86</td>
<td>72</td>
<td>30</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 5.6: Abundance of Tpr, Nup50, Nup155, Nup153 and Nup214 transcripts during murine development based on EST data. Numbers are the transcript per million (TPM) in each pool. Spot intensities are based on TPM. Source: NCBI Unigene database http://www.ncbi.nlm.nih.gov/unigene.

5.3.2.3 The HSP90-CDC37-FKBP5 complex

HSP90AA1 and HSP90AB1 belong to a large, ubiquitously expressed family of chaperone proteins that account for approximately 1% of all proteins in a cell
under non-stressed conditions (reviewed in Chen et al., 2005). HSP90AA1/AB1 proteins act as chaperones for an ever increasing number of clients (>300) playing a central role in the regulation of client function, subcellular localization, enzymatic activity, degradation and binding partner specificity (reviewed in Picard, 2002). In general both isoforms are considered functionally equivalent although Hsp90ab1 mutant mice exhibit embryonic lethality owing to defects in placental development (Voss et al., 2000), whilst HSP90AB1 depletion appears not to impinge on embryonic development and instead plays a role in spermatocyte growth and differentiation (Grad et al., 2010). However, the different knockout phenotypes do not necessarily indicate different isoform functions and may be a consequence of genetic background penetrance or differences in tissue-specific expression levels of the other isoform.

HSP90AA1/AB1 proteins have been shown to co-IP with various kinases including AKT (Sato et al., 2000), p38 (Ota et al., 2010), ERK1/2 (Piatelli et al., 2002, Setalo et al., 2002), and possibly ERK5 (Truman et al., 2006), as well as the immunophilin FKBP5 also identified in this study. Inhibition of HSP90 function by geldanamycin treatment results in a decrease in RAF-RAS association, RAF half-life, as well as the membrane trafficking of newly synthesized RAF (Schulte et al., 1995) providing a mechanism by which HSP90 function directly regulates RAS-RAF-ERK signaling.

A common binding partner of cytosolic HSP90 isoforms is the kinase CDC37 (homologue of S. cerevisiae cell devision cycle 37) (reviewed in Pearl, 2005). Whilst there are a growing list of HSP90-independent functions for CDC37, it is best categorized as a HSP90 cochaperone linking HSP90 to various cellular kinases. In yeast, Cdc37 plays an essential role in kinase post-translational maturation (Mandal et al., 2007), whilst in mammalian systems it can behave as an oncogene and is becoming an increasingly popular target in the treatment of various cancers. (Gray et al., 2007, and reviewed in Gray et al., 2008). Cdc37 has proven to be essential for viability not only in yeast (Gerber et al., 1995) but also in Drosophila (Cutforth and
Interactions between CDC37 and RAF isoforms have been described previously (Grbovic et al., 2006, Silverstein et al., 1998), with overexpression of drosophila Cdc37 resulting in an increase in Raf-1 dependent signaling in a HSP90 dependent fashion (Grammatikakis et al., 1998). This process is geldanamycin sensitive and is therefore most probably mediated through a CDC37:HSP90 heterocomplex. The failure to detect any RAF peptides by MS/MS analysis is most likely due to the choice of lysis conditions used (0.02% non-ionic detergent) during sample preparation which is not likely to solublize membrane bound proteins therefore proteins that are also in complex with membrane bound proteins will be depleted from the sample. Hence the putative HSP90:CDC37:FKBP5:ERK2 complex could represent protein trafficking mechanism that may or may not function as a preformed signaling module.

5.3.2.4 DUSP6 and DUSP9

DUSP6 and DUSP9 belong to a large family of induced dual specificity phosphatases that regulate MAPK activity in both the cytoplasm and nucleus. ERK activity regulates both gene expression and protein stability of some dUSP proteins (Brondello et al., 1995, Brondello et al., 1997, Brondello et al., 1999). Moreover it appears a general mechanism that ERK:DUSP binding results in an increase in the catalytic activity of the phosphatase (Christie et al., 2005 and references therein) thus forming a positive feedforward loop for ERK inhibition. Gene expression analysis following acute stimulation of Fgf4−/− mES cells with FGF4 (T. Kunath unpublished) showed that 5 of the 23 Dusp family members are induced upon stimulation (Figure 5.19). Both DUSP6 and DUSP9 were identified in this data set, however Dusp9 appeared to be negatively regulated following FGF stimulation in WT mES cells (Figure 3.13-A). As Dusp9 expression is highly transient the conflicting data may be due to differences in rate of transcript production/decay between both cell lines, or as a consequence of the different stimulations used, and therefore Dusp9 expression
may have peaked and already have declined below basal levels before 1h and thus appearing to be negatively regulated. Moreover identification of DUSP9 as a high confidence hit in both data sets (embryo and mES derived) would imply that that DUSP9 is a bona fide ERK2 substrate in mES cells. Moreover this data indicates that ERK induced negative feedback mediated by DUSP proteins may operate on multiple timelines underscoring the regulatory complexity of this network.

Figure 5.19: FGF4 induced Dusp expression in mES cells. A) Analysis of Dusp family expression in mES cells. Fgf4^-/- mES cells with recombinant FGF4 and gene expression was monitored at 0, 1, 6, 12, and 24 hours by microarray analysis. B) Dusp genes from A whose expression was positively regulated by FGF stimulation. Genes that showed a 1.5 fold induction were determined as positively regulated.

5.3.2.5 p90RSK proteins

p90RSK proteins are canonical mediators for ERK signaling and play essential roles in both transcriptional and translational regulation. They were initially identified in *Xenopus* eggs as the kinase responsible for phosphorylating 40S
ribosomal subunit protein S6 (rpS6) (Nielsen et al., 1982). It was subsequently shown that the *Xenopus* p90RSK homologue S6KII was activated by ERK2 in response to insulin stimulation (Sturgill et al., 1988). Molecular cloning for p90RSK homologues from various species showed that the unique dual kinase domain architecture was conserved across evolution (reviewed in Anjum et al., 2008). Since then 4 p90RSK family members have been identified and show a high level of sequence homology and overlapping expression patterns (Zeniou et al., 2002), and are in general thought to be functionally equivalent, however there are some key regulatory differences between the isoforms such as the growth factor independent p90RSK4 activity (Dummler et al., 2005) or direct phosphorylation of p90RSK2 by FGFR3 (Kang et al., 2007).

Inactive p90RSK proteins are predominantly localized to the cytoplasm, however nuclear inactive p90RSK has also been detected, where it is complexed with ERK1/2 (Chen et al., 2002). The mechanism of p90RSK activity is a multiple step process involving both ERK and PDK1 enzymatic activities following which a portion of active p90RSK remains bound to ERK and translocates to the nucleus, however the majority of ERK and p90RSK disassociate and shuttle to their individual sites of action (Roux et al., 2003). This disassociation is an active mechanism resulting from the phosphorylation of Ser749 by the N-terminal kinase domain, and should be therefore PDK1 dependent. Interestingly this disassociation was not detected in mES cell stimulated with FGF2, indicating that p90RSK proteins may not be fully activated upon acute FGF2 stimulation in mES cells. Moreover, no difference in the level of Histone H3 S9, reported to be activated by p90RSK2 (Sassone-Corsi et al., 1999) could be detected in either *Pdk1* mutant mES cells, or BI-D1870 treated cells (data not shown), arguing for non-canonical p90RSK function in mES cells. Therefore it would be informative to compare the binding of PDK1 to p90RSK proteins following either FGF4 or FGF2 stimulation by co-immunoprecipitation and how this relates to ERK1/2 binding and disassociation.

Analysis of the gene induction regulated by the FGF/ERK pathway in *Pdk1* null mES cells showed that p90RSK activity is dispensable for IEG induction,
contrary to other cell types (De Cesare et al., 1998). It has been previously shown that mitogen and stress activated protein kinase 1 (MSK1) can bind to IEG promoters, phosphorylate Histone H3 and initiate transcription in response to mitogenic stimuli (Perez-Cadahia et al., 2010), and thus may share overlapping functions with p90RSK proteins in mES cells. As MSK1 activity is PDK1 independent, but ERK responsive (Hague et al., 2006 and references therein) this could explain the IEG response in Pdk1 null mES cells. Small molecule inhibitors to MSK1 have recently been developed (Bamford et al., 2005), it would be therefore interesting to determine the IEG response in Pdk1 null cells in the presence of such an inhibitor however the viability of MSK1/MSK2 double knockout mice indicate that MSK activity is not essential for the pluripotent state, in vivo.

Analysis of ERK1/2 phosphorylation in Pdk1 mutant cells showed that ERK2 1/2 was hyperphosphorylated, in agreement with previous studies (Williams et al., 2000) and that this correlated with enhanced IEG expression in response to FGF stimulation. Moreover transgenic addback experiments showed that the enhanced ERK activity was PDK1 dependent. Although previous studies had shown that PDK1 activity was dispensable for induction of neural progenitors from mES cells (Stavridis et al., 2007), this conclusion was based only on the expression of Sox1, whilst the ability of Pdk1 null cells to generate post-mitotic neurons, nor other lineages in vitro was not known. Analysis of mutant and rescued cell lines following EB differentiation showed no obvious defect in mesoderm nor endoderm differentiation, however there was a severe impairment in neuronal differentiation, consistent with embryological data. Recent evidence also indicates that PDK1 activity is essential for neuronal differentiation and subtype specification during forebrain development (Oishi et al., 2009). The recapitulation of aspects of the forebrain phenotype in Pdk1 mutant cells would imply that mES cells may represent a tractable cell type to model forebrain differentiation, in vivo.
5.3.3 BI-D1870 as a tool to modulate ERK function in mES cells

Evidence in the literature pointed to a p90RSK dependent negative feedback loop that inhibited SOS upstream of the RAF-MEK-ERK cascade that may be responsible for the increase in ERK activity that was noted in Pdk1 mutant cells. However, as PDK1 is essential for p70S6K as well as PKB in mES cells it was not possible to conclude if ERK activity was negatively regulated by p90RSK. Analysis of ERK activation dynamics, by monitoring steady state ppERK levels, or IEG induction in the presence of the BI-D1870 showed that p90Rsk inhibition coincided with hyperactivation of the ERK pathway. Interestingly, BI-D1870 treatment resulted in a rapid, but transient increase in ppERK levels, with values peaking at 10 minutes, and returning to basal levels by 30 minutes. This is in contrast to similar experiments in Rat-2 cells where ERK activity was sustained for over an hour after BI-D1870 treatment. This transient expression could be explained by the ERK dependent expression of DUSP proteins. As ERK activation has previously been assessed following PD173074 treatment followed by growth factor stimulation, upon stimulation the level of DUSP proteins will be quite low and therefore will not affect the early stages of ERK activation, only after ERK dependent expression and protein stabilization (>45 minutes) will DUSP activity affect ppERK levels. When cells were treated with BI-D1870, Erk activity was already in steady state, therefore there will be a proportion of Dusp proteins present and potentially activated by Erk thus resulting in rapid dephosphorylation of ERK following BI-D1870 treatment. In keeping with this notion, addition of BI-D1870 to Fgf4-/- cells, where ppERK levels are low, results in again a peak of ppERK activity at 10 minutes, but is largely sustained over a 90 minute period, where upon ERK activity drops to pre-stimulated levels. The levels of IEG induction however, ranging from five to ten fold higher in the presence of BI-D1870, do not correlate with the moderate increase in ppERK levels seen when cells were subjected to acute stimulation with FGF2. Analysis of the cumulative (time-integrated) ppERK response following stimulation in the presence of BI-D1870 showed only a 15% increase in the ppERK response. As BI-D1870 also inhibits both Aurora B kinase and Polo-like kinase 1 (PLK1) with similar
efficacy as p90RSK proteins, it is possible that the increase in IEG expression observed was not necessarily Erk dependent. It has been shown previously that the Aurora B kinase inhibitor ZM447439 can enhance the cFos response to norepinephrine stimulation (Price et al., 2009), and that Aurora B inhibition leads to apoptosis in some cell types (Li et al., 2010), which can be accompanied by enhanced Egr1 (Catania et al., 1999) and cFos (Abiko et al., 1997) expression. Therefore it would be informative to gauge the levels of IEG induction in the presence of both BI-D8170 and PD0325901 to determine the actual contribution of ERK activity to this phenomenon. It may be possible that off-target inhibition of Aurora B indirectly activates some IEG promotors but that the epigenetic effect of mitogen stimulation, such as Histone H3 phosphorylation seen at the cFos promoter (Li et al., 2003, and reviewed in Thompson et al., 1999), exaggerates this and thus leads to hyperactivation upon FGF stimulation.

The apparent reduction in self-renewal and Nanog expression seen upon BI-D1870 could also be equally attributed to Aurora B inhibition as to ERK activation. Aurora B is known to bind p53 in a complex also involving novel INHAT repressor (NIR), a repressor of p53 transcriptional activity (Wu et al., 2011). Wu and colleagues showed that overexpression of Aurora B resulted in inhibition of p53 transcriptional activity, and that Aurora B could phosphorylate p53 in its DNA binding domain to interfere with p53 function. There is evidence to suggest that p53 binds directly to the Nanog promoter and recruits mSIN3a to silence the locus during differentiation (Lin et al., 2004, Han et al., 2008), and therefore the decrease in Nanog expression could be due to enhanced p53 activity resulting from Aurora B inhibition. This could explain why no difference in the levels of Tbx3, a putative MAPK target (Niwa et al., 2009), was noted upon BI-D8170 treatment.

The effect of BI-D1870 treatment has been studied in several cell types to date. Fonseca and colleagues noted that BI-D1870 inhibited proliferation of HEK293e cells, in a p90RSK independent fashion (Fonseca et al., 2011). In HeLa S3 cells, although BI-D1870 inhibited cell motility, no documented increase in apoptosis or decrease in proliferation was observed (Tanimura et al., 2011), as was the case in
3T3-L1 adipocytes (Chen et al., 2009), and in SW1353 cells (Petrella et al., 2011). However, at least in the case of experiments in 3T3-L1 adipocytes, BI-D1870 treatment was very short (30 minutes) and therefore other physiological affects would not have manifested. Patrella and colleagues provided no information as to the duration of BI-D1870 treatment in their experiments. Although, the possible off-target effects of BI-D1870 treatment indicate that it is not suitable to specifically modulate p90RSK function in mES cells, the fact that it potentially also inhibits Aurora B as well as p90RSK would imply that it may be useful in the treatment of some kinds of cancers (Girdler et al., 2006, and reviewed in Anjum et al., 2008).

Although efforts to affect ERK activity through p90RSK inhibition failed, genetic evidence would still indicate that targeting PDK1 activity may prove fruitful. Several small molecule inhibitors of PDK1 are currently available (reviewed in Peifer et al., 2008), although the specificity of some of these inhibitors are a concern (Bain et al., 2007). Recent work from the lab of Dario Alessi has identified GSK2334470 as a novel and highly specific inhibitor of PDK1 (Najafov et al., 2011). Tested against a panel of nearly 100 related and unrelated kinases, GSK2334470 failed to inhibit any kinase tested other than PDK1 at a concentration (0.1µM) that reduced PDK1 activity by ~90%, although at higher concentrations maternal embryonic leucine zipper kinase (MELK) was also significantly inhibited. In vivo PDK1 inhibition by GSK2334470 was incomplete however, and dependent of the subcellular localization of PDK1. This being said, the cytosolic activity of PDK1 was dramatically reduced by treatment with GSK2334470, and this resulted in inhibition of p90Rsk2 activity in HEK293, U87, and MEF cells and thus may be useful in studying the relative contribution of p90RSK activity to ERK regulation.
Chapter 6 - The pERKs and ERKs of MAPK signaling in mES cells

In the 30 years since the initial isolation and characterization of the first mES lines our understanding of the basic biological pathways that maintain mES cell identity in culture has greatly expanded. Significant progress has been made in developing defined culture conditions for the maintenance and expansion of mES cells that allow the precise molecular requirements for mES cell maintenance to be determined. However, our understanding of how mES cells make the decision to commit to a specific lineage has lagged behind in comparison.

There is a growing body of work that implicates FGF4 activation of the Erk-MAPK kinase as a critical determinant for mES cells to exit the self-renewal cycle and gain competence to differentiate, however the mechanism by which this is achieved is largely unknown. The initial aim of this study was to mechanistically characterize ERK2 by generating a protein interaction network in the hope of identifying molecules targeted by FGF-ERK signaling. In doing so I discovered that the differentiation defect exhibited by Erk2 mutant cells could not be attributed to the target disruption of the Erk2 allele and therefore the aim of the study shifted slightly to attempt to understand the relative contribution of ERK2 to FGF-MAPK signaling and how this affects mES cell self-renewal. To do this I deleted a critical exon of the Erk2 allele by two rounds of homologous recombination in 129/Ola mES cells. This mutation produced a nonfunctional allele with no detectable ERK2 protein and dramatically reduced mRNA expression. Initial analysis of ERK2 deficient cells showed that ERK1 was hyperphosphorylated, however a reduction in FGF2 induced phosphorylation and gene expression was observed indicating that the extent to which ERK1 was hyperactivated was insufficient to completely compensate for loss of ERK2. To address the issue of functional redundancy between ERK1/2 isoymes, I rescued ERK2 deficiency by expressing either exogenous ERK2 or ERK1 and noted that either isoform was capable of rescuing the defect in signal transduction at the level of IEG induction.
Analysis of the self-renewal efficiency of Erk2 mutant cells showed that when maintained in standard culture conditions ERK2 deficient cells exhibited a slight increase in pluripotency-associated gene expression as well as enhanced self-renewal at clonal density. However, when mutant cells were induced to differentiate, they exhibited no block in lineage commitment, and appeared to exit self-renewal as efficiently as WT cells.

Due to the high level of interconnectivity and crosstalk between several pathways downstream of FGFR it is difficult to draw any firm conclusions from the functional experiments described in this study. If ERK1 hyperphosphorylation was due to decreased negative feedback to SOS then presumably this will result in hyperactivation of the PI3K pathway, and thus the increase in self-renewal observed in Erk2 mutant cells may simply be due to increased PI3K dependent Nanog expression (Storm et al., 2007). Indeed, MEK inhibition in mES cells results in increased PKB phosphorylation (Stavridis et al., 2007), indicating that attenuating the ERK-MAPK signaling has knock-on effects on other connected pathways. Similarly, experiments where ERK2-K52R expression appeared to ‘rescue’ the WT ERK2 defect may be due to again decreased negative feedback to up-stream components of the pathway resulting in the increased activation of ERK5. ERK5 is known to be expressed and active in mES cells (Zou et al., 2006). ERK5 can contribute to RAS-dependent signaling (English et al., 1999), and can also phosphorylate p90RSK and promote cFos expression (Ranganathan et al., 2006, Terasawa et al., 2003).

6.1 Technical limitations of this study

A major obstacle in interrogating signaling pathways is that critical parameters may operate below the detection limit of conventional quantitate western blotting (Schilling et al., 2010). Furthermore, analysis of phosphorylation events on a population level may average across subtle cell-to-cell variations, whilst further still, information as to the subcellular localisation and positional activation of proteins is lost by this method of analysis. An exciting new solution to this problem is the use of
various FRET-based biosensors that can report on the subcellular localisation, phosphorylation, or protein:protein interaction of a target molecule at a single cell level (reviewed in Lalonde et al., 2009). Although still an emerging technology, biosensors represent a powerful tool to probe signal transduction pathways.

It is hoped that this study will go some way to highlighting some common pitfalls when interpreting differentiation defect phenotypes in mES cells. Because mES cell proliferate up to twice as fast as some differentiated cell types, cells that develop spontaneous mutations that decrease their ability to differentiate, or enhance their self-renewal, can rapidly dominate a culture making culture adaptation a constant concern. If treated appropriately mES cells can maintain normal ploidy and retain the ability to contribute to the mouse germ-line after prolonged periods of culture, however certain in vitro manipulations can place a selective pressure on cells. Clonal expansion, as is routine in all gene targeting experiments as well as procedures such as high G418 selection for transgene homozygosity can place selective pressure on cells. The chromosomal aberrations observed in Srf mutant cells was initially assumed to be a consequence of high G418 selection, however subsequent analysis of the aneuploid of the parental heterozygous line showed that they were also aneuploid. When generating Erk2 knockout cells I tried to control for clonal variation and culture adaptation by individually targeting two independent Erk2 heterozygous lines and then analyzing the phenotype in individual lines derived from both targeting events, with or without transgenic ERK2 add-back. Furthermore, cultures were routinely karyotyped following each stage of clonal selection, and without exception exhibited normal ploidy even after four rounds of clonal selection.

In contrast to the Ju09 derived lines used for Erk2 studies, mES cell lines derived from a C57Bl/6j/129 mixed background were highly unstable in culture and distinct morphological and behavioral differences were observed, depending on the contribution of either genetic background: Srf<sup>+/fl</sup> lines derived from C57Bl/6j/129 x 129/Ola intercrosses displayed an atypical mES cell morphology, underwent necrotic death at medium densities and showed little responsiveness to 2i culture. Srf<sup>fl/fl</sup> lines derived from C57Bl/6j/129 x C57Bl/6j/129 matings more closely resembled E14tg2a
mES cells, however they were very difficult to manipulate and the frequency of spontaneous differentiation rapidly decreased following prolonged culture. However, Srffl/fl lines responded well to 2i culture and became far more robust to both electroporation and lipofection making it possible to generate stable transfectants under these conditions. Although 2i culture allows derivation of mES lines from backgrounds other than 129, without the use of feeders or immunosergical depletion of TE cells, it may not ensure that WT lines derived under these conditions will be stable in prolonged culture. Indeed, mES cell line derived from NOD blastocysts exhibit a near normal karyotype (~80-90%) at p14, however information as to the karyotype at higher passages was not provided (Nichols et al., 2009).

6.2 Naïve and Primed states of pluripotency

The recent derivation of pluripotent cells from the post-implantation epiblast has helped to further our understanding of the nature of pluripotency (Tesar et al., 2007, Brons et al., 2007). Initially derived from rodent E5.5 epiblasts, epiblast stem cells (EpiSCs) exhibit some features similar to their ICM derived counterparts such as teratoma formation. They express Oct4, Sox2 and Nanog, though they lack the expression of other mES associated transcripts such as Klf4, Stella and Tbx3. Moreover they express the lineage associate transcripts Sox17 and Acta2, amongst others (Tesar et al., 2007). However they contribute poorly to chimeras indicating that they represent a pluripotent cell type distinct from mES cells derived from ICM cells. This being said, it is possible to reprogram EpiSCs to an mES cell like state by forced expression of Klf4 (Guo et al., 2009), or inhibition of multiple signaling pathways (Zhou et al., 2010), however, transition from the mES cell state to the EpiSC state requires only exposure to EpiSC medium (N2B27 supplemented with Activin A (20 ng/ml) and FGF2 (10 ng/ml). Indeed, the observation that mES cells exhibit the heterocogenous and reversible expression of mES cell markers Rex1 (Toyoooka et al., 2008) and Stella (Dppa3) (Hayashi et al., 2008), which are not expressed in EpiSCs indicates that mES cells in culture may be transiting between
ES and EpiSC states, either state can be stabilized in the appropriate culture condition, with 3i conditions stabilizing the naïve, ES cells state and EpiSC culture conditions stabilizing the primed EpiSC state. FGF-ERK signaling acts to maintain the identity of EpiSCs by blocking differentiation towards neuroectodermal fates (Greber et al., 2010), furthermore, chemical inhibition of FGFR activity following 24 hours of neural induction increased neural marker gene expression (Stravidis et al., 2010). Therefore it may be possible that mES cells need to transit along a path of restricted potential similar to that followed by the developing embryo: with FGF-ERK activity required for the separation of PE and epiblast lineages, but then the subsequent inhibition of FGF activity is required for epiblast-like cells to exit their self-renewal and commit to other fates (Sterneckert et al., 2010). The observation that present in any mES cell culture is a mix of potentially ICM-, epiblast- and PE-like populations (Toyooka et al., 2008, Canham et al., 2008), each with potentially different responses to FGF-ERK signaling underscores the limitations of conventional biochemical approaches to understanding signaling dynamics.
Appendix

A1: Steady-state behavior ERK behavior in mES cells

The dynamics of MAPK signaling are generally studied following acute stimulation with ligand/drug of choice. Classical cell lines for studying these dynamics, such as PC12, an MCF7 cell line, do not express ERK-activating ligands, therefore under standard culture conditions synchronization by serum withdrawal, followed by stimulation with the desired ligand is sufficient to elicit a near 100% ppErk response and all cells (Sasagawa et al., 2005, Santos et al., 2007, Nakakuki et al., 2010), with results indicating that as much as 50% of all Erk molecules in a cell will be phosphorylated following acute stimulation (Fujioka et al., 2006). mES cells however express high levels of autocrine Fgf4 (Luster et al., 2003) and LIF (Chambers et al., 2003) resulting in a heterogenous ERK activation (Figure A2-B). This behavior has been mathematically predicted and modeled for networks that exhibit multiple layers of feed-forward and feed-back control (Kholodenko et al., 2000, and reviewed in Kholodenko et al., 2010). Alternatively heterogenous Erk activity could be cell-cycle linked as seen in yeast systems where activity of the ERK orthologue Fus3 is actively restricted to G1 (Strickfaden et al., 2007).

If a high dose of FGF2 (10-20ng/ml) is added to mES cells without pretreatment with the FGFR inhibitor PD173074 no detectible induction of IEG’s is observed (Figure A2-A). This may be because only a small population of cells are FGF responsive due to the steady state dynamics described above, with IEG expression being masked because of population analysis. Negative feedback acting on a time scale consistent with the maturation of DUSP proteins has been known to result in pathway desensitization (reviewed in Birthwistle and Kolch, 2011). Conversely, the lack of detectable IEG induction may be due to an ERK induced transcriptional repressor, as predicted to act on the cFos promoter (Nakakuki et al., 2010). In this scenario alternative waves of ERK activity, due to either cell-cycle position or steady-state dynamics, would generate oscillations of such a transcriptional repressor. Therefore only cells with coinciding levels of high ERK
activity and low repressor expression will result in a change in gene expression profile.

Although this method allows bulk cell analysis of ERK activation dynamics there are several considerations to be taken into account when interpreting the resulting data; the kinetics (i.e. time-to-peak and decay) may not be representative of ‘actual’ signaling kinetics as the dissociation rate of receptor:drug complexes must be taken into account; the amplitude of Erk activity may also vary depending on the efficiency of receptor inhibition.

Figure A1: Map of pIPC37 used for general purpose transgene expression.
Figure A2: Behavior of steady state ERK signaling in mES cells. A) Q-PCR analysis of IEG induction with and without PD173074. Error bars denote standard error from 2 experiments conducted in parallel. Results are representative of 2 separate experiments. B) Immunostaining for ppERK1/2 and total ERK1/2 for mES cells cultured under standard mES conditions. ppERK1/2 shown in green, ERK1/2 in red and DNA counterstained with DAPI in blue.
A2: Analysis of the Erk2 genetrap cell line RRZ307

Due to time considerations, and a lack on information as to the targeting efficiency of the Erk2 locus, it was decided to take advantage of one of the public cell line collections of gene-trap mutated mES cells. BayGenomics, whose repository is now managed by The Mutant Mouse Regional Resource Center (MMRRC) at UC Davis, has generated a wide collection of germ-line competent gene-trap cell lines on a 129 background. A search of the MMRRC catalogue (http://mmrrc.compmed.ucdavis.edu/) yielded five cell lines reported to have trapping events within the Erk2 locus (Figure A1.3-A). The cell line RRZ307, trapped with pGT2lxf, was selected as in-house 5’ RACE analysis showed it to be trapped in intron 1 and thus would most likely produce a nonfunctional allele. RT-PCR analysis, with a forward primer against Erk2 exon 1 and reverse primers against the β-Geo cassette was performed on RRZ307 derived cDNA to confirm the presence of the a Erk2 exon 1/Engrailed exon splice junction, which was subsequently sequence verified (Figure A3 B-C). Amplification of the Erk2:En fusion transcript was very inefficient (requiring >30 PCR cycles), however indicating that exon trapping may not be occurring faithfully. Moreover, western blot analysis showed no difference in the level of ERK2 protein, compared with 3 separate WT mES cell lines (Figure A3-D). Since there is no evidence in the literature to indicating ERK2 can regulate its own promoter, this datum, along with the difficulty in amplifying the trapped exon by RT-PCR, indicated that the Erk2 exon trapping event may be more complicated than the RACE data suggests. Moreover, although RRZ307 cells were G418 resistance, the expression of β-Geo could not be detected by X-Gal staining. This phenomenon has been reported before and may be due to the expression level of the β-Geo cassette (Friedrich et al., 1991).

Electroporated plasmid DNA can insert as concatemers at very high frequencies in mES cells (Davis et al., 1992). Therefore Southern blot analysis was employed to determine gene-trap cassette copy number, as well as to assess the presence of multiple integrations (Figure A4-A). Genomic DNA from RRZ307 cell was digested with several enzymes that would cut the gene-trap cassette on either
side of the Southern probe sequence, as well as cutting multiple times within the 30kb stretch of intron 1 (Figure A4-B-C). A BamHI digest was also included as there were restriction sites on either side of the probe sequence, therefore serving as a reference for copy number. A single, non-concatemer, integration should yield a single fragment for each restriction enzyme corresponding to the pGT2lxr sequence and a portion of intronic sequence. Southern blot analysis showed multiple bands for each restriction enzyme, except BamHI as expected. The prevalence of a dominant band of approximately 8.5kb; full length plasmid, indicates the plasmid inserted as concatemer. However, the presence of up to 4 bands for some restriction enzymes implies multiple insertion events.

A3: Comparison of ERK1 activity by either FGF2 or LIF

To determine if there may be different kinetics of ERK1 activation by either FGF2 or LIF, Erk2 null cells were incubated overnight in the presence of both LIF antagonist; hLIF05 (section 2.2.4.6), and the FGFR; inhibitor PD173074 (100nM). The following day the media was washed away as usual and samples were stimulated for the indicated times with either LIF (100U) and PD173074 (100nM), or FGF2 (10ng/ml) and hLIF05. Western blot analysis reveals ERK1 is stably phosphorylated by FGF2 (within the 120 minute timeframe of the assay), whilst LIF stimulated ppERK1 peaks at 30 minutes and steadily decreases thereafter (Figure A12).
**Figure A3: Selection and validation of Erk2 genetrap cell line.**  
A) RACE analysis performed at Baygenomics indicated the cell line RRZ307 was trapped in the first intron.  
B) DNA Sequence trace chromatogram showing proper alignment of a PCR fragment amplified from cDNA from RRZ307 with the trapped exon:exon junction.  
C) PCR confirmation of the trapped locus, the location of the primers used is depicted in B.  
D) Western blot analysis of ERK2 expression between RRZ307 and 3 WT mES cell lines, E14tg2a, Ju09, and Fl4.
Figure A4: Southern blot analysis of pGT2lxf integration in RRZ307. A) Southern blotting with a probe targeted against a 600bp region of the β-Geo cassette in pGT2xf: lane 1 WT BamHI digested E14tg2a gDNA, lanes 2-9 RRZ307 gDNA digested with enzymes as indicated. B) Theoretical band sizes produced for each restriction enzyme used if the genetrap cassette inserted as a concatemer only. C) Cartoon depicting the location of the restriction enzymes used, relative to the Southern probe. HindIII was used to linearize the plasmid before electroporation.
**Figure A5: pETV.** Plasmid from which the Erk2 targeting arms were cloned.

**Figure A6: AGS 703.** Plasmid from which the Hygromycin-Thymidine Kinase cassette was cloned.
Figure A7: Rosa26.1. Plasmid from which mPGK-DT-A cassette was cloned.
Figure A8: pETVH. A) A map of the final targeting construct. B) Restriction fragment analysis confirming correct cloning. C) Expected band sizes for the enzymes used in B.
Figure A9: pL451. Plasmid from which the mPGK-EM7 Neomycin cassette was cloned.
**Figure A10: pETVN**  A) A map of the final targeting construct. B) Restriction fragment analysis confirming correct cloning, clone A selected for use. C) Expected band sizes for the enzymes used in B.
Figure A11: pGT2lx. Genetrap plasmid used to generate RRZ307.

Figure A12: LIF and FGF activate ERK1 for different durations. Western blot analysis of lysates from Erk2−/− cells acutely stimulated with either FGF2 (10ng/ml), or LIF (100U) for the indicated times. FGF stimulation was performed in the presence of the LIF agonist hLIF05, whilst LIF stimulation was performed in the presence of the FGFR inhibitor PD173074.
A4: Low level expression of FKBP-SRF in SRF deficient cells.

As there was evidence from the literature that SRF is cleaved in a concentration dependent manner, and that the resulting fragments can act as semi-dominant negative proteins to interfere with normal SRF function, the FKBP-L106P system (Banaszynski et al., 2006) was employed to empirically tune SRF levels. The Srf coding region was amplified from E14tg2a total RNA with the primers RVSGL and NSRFR that contained EcoRV and NotI sites, respectively. A poly-glycine linker was included at the 3’ end of the forward primer to aid the accessibility of the FKBP epitope. The resulting PCR fragment was TOPO cloned, sequence verified, and then cloned into pCAG-FKBP-IH (Kei Kaji) (Figure A13). FspI linearised plasmid (25µg) was electroporated into Srf null cells and selected for 8 days with Hygromycin B (150µg/ml). Figure A14-A shows representative images of clones derived from either mock electroporated (pIPC37-stuffer-IP), pIPC37-Srf-IP or pCAG-FKBP-Srf-IH. The majority of clones electroporated with pCAG-FKBP-Srf-IH exhibited a skirt of differentiated cells at the periphery of the colony, which was absent in both mock and pIPC37-Srf-IP electroporated clones. Western blot analysis of two clones expressing pCAG-FKBP-Srf-IH showed that in the absence of the stabilizing drug Shld1, SRF expression levels are similar to endogenous (Figure A14-B). Upon drug treatment there is a dose dependent response in SRF protein levels. Moreover, at high SRF expression levels, the appearance of a truncated protein was observed, however it is not possible to tell if SRF degradation was a consequence of its expression level, or whether at low levels the truncated protein was out of the detection limits of the assay. Though this method yielded near endogenous expression levels of SRF, it was not capable of restoring neural commitment in the Srf mutant line. Due to issues of abnormal ploidy it was decided that this was not a suitable cell line to determine Srf function.
**Figure A13**: pCAG-FKBP-L106P Srf-IH. Plasmid used to express FKBP fusion proteins.

**Figure A14**: Low-level expression of SRF allows differentiation at clonal destiny. Comparison of the affect of high and low SRF expression at clonal density. MTV=pIPC37-Stuffer. B) Western blot analysis comparing the levels of SRF expression between FKBP-SRF (with and without Shld1) and endogenous. Note the appearance of a degradation product upon high expression (lanes 3-4).
Figure A15: Plasmid used to express FKBP-Cre-ErT2 fusion protein.

<table>
<thead>
<tr>
<th>Protein Family</th>
<th>Symbol</th>
<th>Peptide number in negative sample</th>
<th>Peptide number in positive sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hnnp</td>
<td>HnnpA3</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>HnnpL</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>HnnpU</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>HnnpA0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>HnnpA2b1</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Hnnpab</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>HnnpA1</td>
<td>19</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>HnnpC</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>HnnpM</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>HnnpH1</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>HnnpK</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>HnnpF</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Eif5a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Eif4a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Eif3k</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Eif3f</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Protein Family</td>
<td>Symbol</td>
<td>Peptide number in negative sample</td>
<td>Peptide number in positive sample</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------</td>
<td>-----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Eukaryotic initiation factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif3h</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Eif3a</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Eif3e</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Eif4b</td>
<td>23</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Eif3b</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Eif2s3y</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Eif3c</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Eif2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Ribosomal proteins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rps27a</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Rps19</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Rps25</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Rps3</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Rps7</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Rps8</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Rps18</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Rps4x</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Histone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2B-1</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>H2A-1</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>H3.3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>H4</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Baf proteins (Smarc)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baf57</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Baf155</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Baf53a</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Baf60a</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Psm proteasome proteins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psmd2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Psmb5</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Psmd11</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Psma7</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Psmd1</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Psmb1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Psmc2</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Protein Family</td>
<td>Symbol</td>
<td>Peptide number in negative sample</td>
<td>Peptide number in positive sample</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------</td>
<td>----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>NuRD components</td>
<td>Psmc6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NuRD components</td>
<td>Gatad2a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NuRD components</td>
<td>Hdac1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>NuRD components</td>
<td>Rbbp4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>NuRD components</td>
<td>Rbbp7</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>NuRD components</td>
<td>SalI4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Ddx RNA Helicases</td>
<td>Ddx9</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Ddx RNA Helicases</td>
<td>Ddx39</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Ddx RNA Helicases</td>
<td>Ddx5</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Ddx RNA Helicases</td>
<td>Ddx3x</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Ddx RNA Helicases</td>
<td>Ddx48</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Tubb2c</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Tuba1b</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Tubb2b</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Tuba1a</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Tubb5</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>Small nuclear riboproteins</td>
<td>Prp31</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Small nuclear riboproteins</td>
<td>SnrnpB</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Small nuclear riboproteins</td>
<td>SnrnpA</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table A1**: A list of common FLAG-agarose binding proteins in mES cells. Peptide numbers are the cumulative peptide identified in the triplicate MS experiment described in chapter 5.
Figure A16: Analysis of the distribution of +ve and -ve background. Values are taken from Table A1.
References


Sabio, G., Arthur, J. S., Kuma, Y., Peggie, M., Carr, J., Murray-Tait, V., Centeno, F.,
Goedert, M., Morrice, N. A. and Cuenda, A. (2005). p38gamma regulates the localisation of
SAP97 in the cytoskeleton by modulating its interaction with GKAP. *Embo J* 24, 1134-45.

Sajan, M. P., Standaert, M. L., Bandyopadhyay, G., Quon, M. J., Burke, T. R., Jr. and
Farese, R. V. (1999). Protein kinase C-zeta and phosphoinositide-dependent protein kinase-1 are

capture of genes specifically active in differentiated embryonic stem cells and in mouse embryos. *Dev Dyn* 212, 326-33.

Santos, S. D., Verveer, P. J. and Bastiaens, P. I. (2007). Growth factor-induced MAPK

Sapkota, G. P., Cummings, L., Newell, F. S., Armstrong, C., Bain, J., Frodin, M.,
Grauert, M., Hoffmann, M., Schnapp, G., Steegmaier, M. et al. (2007). BI-D1870 is a specific


Satoh, Y., Saitoh, D., Takeuchi, A., Ojima, K., Kouzu, K., Kawakami, S., Ito, M.,


Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type
cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* 12, 309-22.

Saxton, T. M., Henkemeyer, M., Gasca, S., Shen, R., Rossi, D. J., Shalaby, F., Feng, G.

Schilling, M., Maiwald, T., Hengl, S., Winter, D., Kreutz, C., Koleh, W., Lehmann, W.
D., Timmer, J. and Klingmüller, U. (2009). Theoretical and experimental analysis links isoform-

Schlake, T. and Bode, J. (1994). Use of mutated FLP recognition target (FRT) sites for the
exchange of expression cassettes at defined chromosomal loci. *Biochemistry* 33, 12746-51.

Schmidt, M., Goebeler, M., Posern, G., Feller, S. M., Seitz, C. S., Brocker, E. B., Rapp,

family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-

Mek1/2 gene dosage determines tissue response to oncogenic Ras signaling in the skin. *Oncogene* 28, 1485-95.


