GENE-SPECIFIC REPORTERS AS POTENTIAL TOOLS FOR THE ENRICHMENT OF ESC-DERIVED OSTEOGENIC PROGENITOR CELLS

A Dissertation in
Cellular and Molecular Biology
by
Jean E. DeMarco

© 2009 Jean E. DeMarco

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2009
The dissertation of Jean E. DeMarco was reviewed and approved* by the following:

Sarah K. Bronson  
Associate Professor of Cellular and Molecular Physiology  
Dissertation Advisor  
Chair of Committee

Rebecca Craven  
Associate Professor of Microbiology and Immunology

Edward Gunther  
Assistant Professor of Medicine

Christopher Niyibizi  
Associate Professor of Orthopaedics and Rehabilitation

Timothy Ritty  
Assistant Professor of Orthopaedics and Rehabilitation

Henry Donahue  
Michael and Myrtle Baker Professor  
Director of the Cell and Molecular Biology Graduate Program

* Signatures are on file in the Graduate School
ABSTRACT

Regenerative medicine has become a contender for the treatment of a variety of diseases and disorders. Transplantation of stem cell derived progenitors, tissues, or organs, has the potential to repair and restore function to the host at a level unachievable by any mechanical device or drug treatment. Substantial proliferative capacity in a stem or progenitor cell could allow for a single transfer of therapeutic cells, versus the chronic, and sometimes toxic, nature of many pharmaceutical based therapies. Osteogenic progenitors can be used for the therapy of multiple musculoskeletal afflictions including severe fractures, resection of cancerous bone, and osteogenesis imperfecta.

Embryonic stem cells (ESCs) are an ideal candidate for stem cell therapy given their pluripotentiality and a nearly infinite self renewal capacity, permitting growth of large numbers of therapeutic cells. Differentiated ESC progeny for transplantation can be engineered to be histocompatible with the recipient by somatic cell nuclear transfer and to have very precise genetic alterations by homologous recombination.

To utilize ESC-derived progenitors for clinical transplantation, methods to identify and enrich for specific lineages from heterogeneous differentiation cultures must be developed, and stages of differentiation harboring therapeutic capacity must be determined. To this end, we have engineered murine ESCs to contain fluorescent reporters, permitting enrichment of progenitors from osteogenic differentiation cultures using fluorescence activated cell sorting (FACS). The overall hypothesis is that by using published in vivo and in vitro data regarding gene expression in osteogenesis, as well as our data on gene expression from the ESC-derived osteogenic cultures, we can identify specific genes with expression characteristic of various stages in osteogenic differentiation, generate transcriptional reporters in ESCs, and use reporter expression to identify specific osteogenic intermediates. These different cell populations will be tested by subsequent in vitro assays and ultimately the specific intermediates could be tested in vivo in the context of a therapeutic paradigm.
Three bacterial artificial chromosome (BAC) based reporters have been engineered to place GFP under the transcriptional control of specific genes expressed at various times during osteogenic differentiation (Twist2, Sox9, or Osterix). Subsequently, these BAC-based reporters were targeted in a single copy to a chosen site in an ESC. An additional reporter, containing 7.6 kb of regulatory sequence from the rat Nestin gene linked to GFP, was targeted to the same genomic site. A fifth reporter, generated by Fehling et al. (Brachyury-GFP) was also tested for enrichment ability. A constitutively expressed YFP transgene has been integrated and tested for ubiquitous expression to allow for tracking of the cells at all times upon transplantation in a murine host model, and a one of the osteogenic-specific reporter ESC lines also contained this tracking reporter.

The Twist2-EGFP BAC-based reporter demonstrated the most promising results: (i) Twist2-EGFP fluorescent expression paralleled the endogenous Twist2 mRNA expression pattern, (ii) transgenic Twist2-EGFP mRNA expression recapitulated endogenous Twist2 mRNA expression, (iii) QRT-PCR analysis of mRNA from Twist2-EGFP positive cells sorted from the osteogenic cultures showed higher expression of osteogenic progenitor markers as compared to the EGFP negative cells, demonstrating enrichment for osteogenic progenitors, and (iv) sorted populations of Twist2-EGFP cells survived and were able to undergo osteogenic differentiation.

This thesis provides proof of concept for FACS-based osteogenic enrichment of ESC-derived differentiation cultures, while simultaneously revealing which stages of the differentiation culture are most accessible, and the possible limitations of interrupting the differentiation process. Ultimately this approach has the potential to enrich for therapeutic osteogenic cells for a variety of applications from cell transplant, to loading in artificial supports for subsequent surgical placement, to in depth in vitro studies that can improve and inform in vivo therapeutic modalities.
TABLE OF CONTENTS

List of Tables ........................................................................................................... ix
List of Figures .......................................................................................................... x
List of Abbreviations .............................................................................................. xiv

1. Chapter 1. INTRODUCTION ................................................................................ 1
  1.1 Development of the Skeleton ............................................................................ 1
      1.1.1 Bone Formation ...................................................................................... 1
      1.1.2 Transcriptional Regulation of Bone Formation .................................. 4
  1.2 Osteoblastic Differentiation of Stem Cells Ex Vivo .................................... 9
      1.2.1 Mesenchymal Stem Cells ...................................................................... 9
      1.2.2 Embryonic Stem Cells .......................................................................... 13
  1.3 Transcriptional Reporters of Osteogenic Gene Expression ......................... 22
      1.3.1 Random Integration ................................................................................ 22
      1.3.2 Homologous Recombination .................................................................. 23
      1.3.3 Enrichment of Subpopulations ............................................................... 29
      1.3.4 Existing Osteogenic Reporters ............................................................... 31
          1.3.4.1 Reporters based on Sub-cellular Localization ............................ 31
          1.3.4.2 Reporters based on Gene Expression .................................. 32
          1.3.4.3 Expression Reporter Systems ....................................................... 36
  1.4 Statement of Thesis ......................................................................................... 39

2. Chapter 2. MATERIALS AND METHODS ............................................................ 42
  2.1 Cell Culture and General Techniques ............................................................. 42
      2.1.1 Cell lines and culture ............................................................................ 42
      2.1.2 Osteogenic differentiation of ESC clones ............................................. 42
      2.1.3 Fixation and staining of osteogenic cultures ...................................... 43
      2.1.4 ESC genomic DNA purification ......................................................... 43
      2.1.5 PCR analysis of genomic ESC clone DNA for integration of transgene . 44
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.6 RT-PCR analysis</td>
<td>44</td>
</tr>
<tr>
<td>2.1.7 SYBR Green real-time quantitative RT-PCR analysis</td>
<td>45</td>
</tr>
<tr>
<td>2.2 Creation and Evaluation of the EYFP ESC Reporter</td>
<td>46</td>
</tr>
<tr>
<td>2.2.1 Construction of CAGGS-EYFP-PNTlox2 vector</td>
<td>46</td>
</tr>
<tr>
<td>2.2.2 Electroporation of CAGGS-EYFP-PNTlox2 into ESCs</td>
<td>46</td>
</tr>
<tr>
<td>2.2.3 Metaphase counts of EYFP ESC clones</td>
<td>47</td>
</tr>
<tr>
<td>2.2.4 Teratoma formation from EYFP ESC clone</td>
<td>47</td>
</tr>
<tr>
<td>2.2.5 Fluorescence Microscopy of EYFP ESC clones and teratoma tissue sections</td>
<td>47</td>
</tr>
<tr>
<td>2.3 Creation and Evaluation of Gene-Specific EGFP Reporters</td>
<td>48</td>
</tr>
<tr>
<td>2.3.1 BAC clones</td>
<td>48</td>
</tr>
<tr>
<td>2.3.2 Preparation of BAC DNA</td>
<td>49</td>
</tr>
<tr>
<td>2.3.3 Preparation of BAC clone glycerol stocks</td>
<td>49</td>
</tr>
<tr>
<td>2.3.4 In vitro Cre/loxP modification of BAC DNA</td>
<td>50</td>
</tr>
<tr>
<td>2.3.5 Construction of recombineering vectors</td>
<td>50</td>
</tr>
<tr>
<td>2.3.6 Recombineering of BAC DNA</td>
<td>51</td>
</tr>
<tr>
<td>2.3.7 Construction of Nestin-EGFP plasmids</td>
<td>53</td>
</tr>
<tr>
<td>2.3.8 Transgene targeting in ESCs</td>
<td>53</td>
</tr>
<tr>
<td>2.3.9 HAT selection</td>
<td>54</td>
</tr>
<tr>
<td>2.3.10 Pulse Field Gel Electrophoresis of BAC DNA</td>
<td>54</td>
</tr>
<tr>
<td>2.3.11 Flow Cytometry</td>
<td>55</td>
</tr>
<tr>
<td>2.3.12 Fluorescent Activated Cell Sorting (FACS) of EBs and osteogenic cultures</td>
<td>55</td>
</tr>
<tr>
<td>3. Chapter 3. CREATION OF AN ESC LINE CONTAINING A CONSTITUTIVE EYFP ESC REPORTER</td>
<td>61</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>61</td>
</tr>
<tr>
<td>3.2 Results</td>
<td>65</td>
</tr>
<tr>
<td>3.2.1 Genomic Analysis of the EYFP Constitutive Reporter in ESCs</td>
<td>65</td>
</tr>
<tr>
<td>3.2.2 Functional Analysis of the EYFP Constitutive Reporter in ESCs</td>
<td>71</td>
</tr>
<tr>
<td>3.3 Discussion</td>
<td>79</td>
</tr>
</tbody>
</table>
4. Chapter 4. CREATION OF GENE-SPECIFIC EGFP ESC REPORTERS 81

4.1 Introduction................................................................. 81

4.2 Results................................................................. 86

4.2.1 BAC Clones................................................................. 86
4.2.2 Modification of BACs to contain Hprt homology .... 89
4.2.3 EGFP Recombineering Vectors........................................ 91
4.2.4 Recombineering of BACs to contain EGFP........... 95
4.2.5 Targeting the Hprt-EGFP-BAC constructs in ESCs... 98
4.2.6 Creation of Nestin-EGFP ESC........................................ 103

4.3 Discussion................................................................. 109

5. Chapter 5. EGFP REPORTERS AND ENRICHMENT POTENTIAL... 112

5.1 Introduction................................................................. 112

5.2 Results................................................................. 115

5.2.1 Brachyury-GFP ESC Reporter........................................ 115
5.2.2 Nestin-EGFP ESC Reporter........................................ 122
5.2.3 Hprt-Osx-EGFP ESC Reporter........................................ 127
5.2.4 Hprt-Twist2-EGFP ESC Reporter..................................... 132
  5.2.4.1 Initial characterization........................................ 132
  5.2.4.2 Analysis of D2 and D8 sort cultures................. 142
  5.2.4.3 Analysis of D4 sort cultures........................... 146
  5.2.4.4 Analysis of D6 sort cultures........................... 152

5.3 Discussion................................................................. 160

6. Chapter 6. GENERAL DISCUSSION................................ 169

6.1 Overview................................................................. 169

6.2 Methodological Considerations...................................... 171

6.2.1 Targeting Efficiency............................................. 171
6.2.2 Dual Reporter System........................................... 174

6.3 Enhancement of the Osteoculture................................. 176
6.3.1 Bioadhesive Surfaces ........................................ 176
6.3.2 Trophic Factors ............................................. 177

6.4 Alternative Avenues ............................................ 178
6.4.1 Lineage Selection ........................................... 178
6.4.2 Existing Constructs .......................................... 179

6.5 Importance of Lineage Specificity and Stem Cell Therapy ...... 183

REFERENCES .................................................................................. 185
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Genetic reporter constructs targeted to the <em>Hprt</em> locus of ESCs</td>
</tr>
<tr>
<td>2.1</td>
<td>Primers for PCR and RT-PCR analysis of transgene integration and Alu fingerprinting</td>
</tr>
<tr>
<td>2.2</td>
<td>Primers for RT-PCR and QRT-PCR analysis of osteogenic cultures and sort populations</td>
</tr>
<tr>
<td>2.3</td>
<td>Primer sequences for Homology Arms of recombineering fragments</td>
</tr>
<tr>
<td>3.1</td>
<td>Euploid CAGGS-EYFP clones as determined by metaphase spread counts</td>
</tr>
<tr>
<td>4.1A</td>
<td>BAC clones</td>
</tr>
<tr>
<td>4.1B</td>
<td>BAC clone sizes</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary of electroporations of BAC Reporters into ESCs</td>
</tr>
<tr>
<td>5.1</td>
<td>Summary of QRT-PCR timepoints performed on D2-D8 Sort Populations</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Intramembranous and endochondral ossification</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Multilineage potential of mesenchymal stem cells</td>
<td>10</td>
</tr>
<tr>
<td>1.3</td>
<td>Derivation of embryonic stem cells</td>
<td>14</td>
</tr>
<tr>
<td>1.4</td>
<td>Culture systems for the osteogenic differentiation of ESCs</td>
<td>18</td>
</tr>
<tr>
<td>1.5</td>
<td>Targeted insertion of transgenic sequences</td>
<td>24</td>
</tr>
<tr>
<td>1.6</td>
<td>Modification and targeting of BACs to the Hprt locus</td>
<td>27</td>
</tr>
<tr>
<td>1.7</td>
<td>General schematic of temporal expression of osteogenic-GFP reporter transgenic mouse lines</td>
<td>33</td>
</tr>
<tr>
<td>1.8</td>
<td>Specific transcriptional reporters used to identify intermediate populations within the osteogenic culture</td>
<td>41</td>
</tr>
<tr>
<td>2.1</td>
<td>Flow chart of construction and characterization of BAC reporters</td>
<td>57</td>
</tr>
<tr>
<td>3.1</td>
<td>pCAGGS-EYFP-PNTlox2 construct</td>
<td>66</td>
</tr>
<tr>
<td>3.2</td>
<td>Integration of pCAGGS-EYFP into ESC clones</td>
<td>67</td>
</tr>
<tr>
<td>3.3</td>
<td>QRT-PCR analysis of CAGGS-EYFP ESC clones</td>
<td>69</td>
</tr>
<tr>
<td>3.4</td>
<td>Sequence of Cre/EYFP junction in pCAGGS-EYFP-PNTlox2...</td>
<td>70</td>
</tr>
<tr>
<td>3.5</td>
<td>Fluorescent microscopy of CAGGS-EYFP ESCs</td>
<td>73</td>
</tr>
<tr>
<td>3.6</td>
<td>Osteogenic differentiation of CAGGS-EYFP ESC clones</td>
<td>75</td>
</tr>
<tr>
<td>3.7</td>
<td>Analysis of EYFP expression in CAGGS-EYFP clones throughout the osteogenic culture</td>
<td>76</td>
</tr>
<tr>
<td>3.8</td>
<td>Expression of EYFP in a teratoma derived from CAGGS-EYFP ESCs</td>
<td>78</td>
</tr>
<tr>
<td>4.1</td>
<td>Hprt modification of BAC clones</td>
<td>85</td>
</tr>
<tr>
<td>4.2</td>
<td>Physical maps of the genomic inserts of BACs to be recombined to create EGFP reporters</td>
<td>87</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>Identification of properly modified <em>Hprt</em> BACS</td>
<td>90</td>
</tr>
<tr>
<td>4.4</td>
<td>Modification of BACs to express EGFP recombineering fragment</td>
<td>92</td>
</tr>
<tr>
<td>4.5</td>
<td>Restriction enzyme digests of properly cloned recombineering vectors</td>
<td>94</td>
</tr>
<tr>
<td>4.6</td>
<td>Proper EGFP recombineering of Twist2 BAC</td>
<td>96</td>
</tr>
<tr>
<td>4.7</td>
<td>Proper EGFP recombineering of Sox9 and Osterix BACS</td>
<td>97</td>
</tr>
<tr>
<td>4.8</td>
<td>Proper modification and kanamycin removal of BACS</td>
<td>99</td>
</tr>
<tr>
<td>4.9</td>
<td>Verification of recombination of BACs into the <em>Hprt</em> locus of HAT-resistant ESC clones</td>
<td>102</td>
</tr>
<tr>
<td>4.10</td>
<td><em>Hprt</em>-Nestin-EGFP targeting fragment</td>
<td>105</td>
</tr>
<tr>
<td>4.11</td>
<td>Verification of targeted <em>Hprt</em>-Nestin-EGFP in ESC clones</td>
<td>106</td>
</tr>
<tr>
<td>4.12</td>
<td>pNestin-EGFP-Neo random integration construct</td>
<td>107</td>
</tr>
<tr>
<td>5.1</td>
<td>Osteoblast differentiation from ESCs</td>
<td>114</td>
</tr>
<tr>
<td>5.2</td>
<td>Brachyury GFP knock-in reporter</td>
<td>116</td>
</tr>
<tr>
<td>5.3</td>
<td>Flow cytometric analysis of Bry-GFP fluorescence throughout the osteogenic culture</td>
<td>118</td>
</tr>
<tr>
<td>5.4</td>
<td>FACS of Bry-GFP cells at D4 and D6 OC</td>
<td>120</td>
</tr>
<tr>
<td>5.5</td>
<td>Colony formation of Bry-GFP sort populations</td>
<td>121</td>
</tr>
<tr>
<td>5.6</td>
<td>Analysis of <em>Hprt</em>-Nestin-EGFP ESC reporter clones for EGFP expression</td>
<td>123</td>
</tr>
<tr>
<td>5.7</td>
<td>Analysis of randomly integrated Nestin-EGFP ESC reporter clones for EGFP expression</td>
<td>125</td>
</tr>
<tr>
<td>5.8</td>
<td>RNA analysis of Nestin-EGFP clones</td>
<td>126</td>
</tr>
<tr>
<td>5.9</td>
<td>Analysis of Osx-EGFP expression throughout the osteoculture</td>
<td>128</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>5.10</td>
<td>Detection of Osx-EGFP mRNA in Hprt-Osx-EGFP ESC derived osteocultures</td>
<td>130</td>
</tr>
<tr>
<td>5.11</td>
<td>Analysis of the Hprt-Osx-EGFP ESC derived osteoculture for a small population of EGFP expressing cells</td>
<td>131</td>
</tr>
<tr>
<td>5.12</td>
<td>Initial verification of Twist2-EGFP fluorescence</td>
<td>134</td>
</tr>
<tr>
<td>5.13</td>
<td>Flow cytometric analysis of Twist2-EGFP fluorescence throughout the osteogenic culture</td>
<td>135</td>
</tr>
<tr>
<td>5.14</td>
<td>FACS of Twist2-EGFP cells throughout the osteoculture</td>
<td>137</td>
</tr>
<tr>
<td>5.15</td>
<td>QRT-PCR analysis of Twist2 endogenous and transgenic expression in FACS positive sort populations throughout the osteogenic culture</td>
<td>138</td>
</tr>
<tr>
<td>5.16</td>
<td>QRT-PCR analysis of Twist2 expression in FACS sort populations</td>
<td>140</td>
</tr>
<tr>
<td>5.17</td>
<td>Flow cytometry of Twist2-EGFP negative sort population</td>
<td>141</td>
</tr>
<tr>
<td>5.18</td>
<td>Colony formation of Twist2-EGFP sort populations</td>
<td>143</td>
</tr>
<tr>
<td>5.19</td>
<td>QRT-PCR analysis of Day 2 and Day 8 FACS sorted and replated Twist2-EGFP populations and Day 21 replate-OC</td>
<td>145</td>
</tr>
<tr>
<td>5.20</td>
<td>QRT-PCR analysis of Day 4 FACS sorted and replated Twist2-EGFP populations at D0, D14, and D21 replate-OC</td>
<td>147</td>
</tr>
<tr>
<td>5.21</td>
<td>FACS of Day 4 Twist2-EGFP OC to include a larger population of Twist2-EGFP fluorescing cells</td>
<td>150</td>
</tr>
<tr>
<td>5.22</td>
<td>QRT-PCR analysis of Day 4 larger-gate sort at D0</td>
<td>151</td>
</tr>
<tr>
<td>5.23</td>
<td>QRT-PCR analysis of Day 6 FACS sorted and replated Twist2-EGFP populations at D0, D14, and D21 replate-OC</td>
<td>153</td>
</tr>
<tr>
<td>5.24</td>
<td>D6 Twist2-EGFP positive sort colony morphology</td>
<td>154</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>5.25</td>
<td>QRT-PCR analysis of Day 6 FACS sorted and densely replated Twist2-EGFP populations</td>
<td>156</td>
</tr>
<tr>
<td>5.26</td>
<td>QRT-PCR analysis of Day 4 FACS sorted and densely replated Twist2-EGFP populations</td>
<td>158</td>
</tr>
<tr>
<td>5.27</td>
<td>Schematic to explain similar Twist2 mRNA levels in the positive and negative FACS sort populations</td>
<td>165</td>
</tr>
<tr>
<td>6.1</td>
<td>BAC trimming</td>
<td>173</td>
</tr>
<tr>
<td>6.2</td>
<td>Gene trapping</td>
<td>180</td>
</tr>
</tbody>
</table>
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3’ RACE</td>
<td>3’ rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>AFP</td>
<td>alpha-fetoprotein</td>
</tr>
<tr>
<td>AG</td>
<td>chicken beta-actin promoter</td>
</tr>
<tr>
<td>BAC</td>
<td>bacterial artificial chromosome</td>
</tr>
<tr>
<td>βGP</td>
<td>β-glycerophosphate</td>
</tr>
<tr>
<td>BK4</td>
<td>subclone of E14TG2a Hprt deficient ESC line</td>
</tr>
<tr>
<td>BMP</td>
<td>bone morphogenic protein</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BrdU</td>
<td>bromodeoxyuridine</td>
</tr>
<tr>
<td>Bry</td>
<td>brachyury</td>
</tr>
<tr>
<td>CFU</td>
<td>colony forming unit</td>
</tr>
<tr>
<td>CIP</td>
<td>calf intestinal alkaline phosphatase</td>
</tr>
<tr>
<td>CMV</td>
<td>cytomegalovirus</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>Col1a1</td>
<td>type 1 collagen (type I, alpha 1)</td>
</tr>
<tr>
<td>Col10a1</td>
<td>type 10 collagen (type X, alpha 1)</td>
</tr>
<tr>
<td>D</td>
<td>day (of EB or OC culture)</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle Medium</td>
</tr>
<tr>
<td>DMP-1</td>
<td>dentin matrix protein 1</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>dpc</td>
<td>days post coitum</td>
</tr>
</tbody>
</table>
ds                  double stranded
DTT                 dithiothreitol
EB                  embryoid body
*E. coli*           *Escherichia coli*
EDTA                ethylenediaminetetraacetic acid
ECFP                enhanced cyan fluorescent protein
EGFP                enhanced green fluorescent protein
ESC                  embryonic stem cell
EYFP                enhanced yellow fluorescent protein
FACS                fluorescence activated cell sorting
FBS                 fetal bovine serum
F-factor            fertility factor
floxed              *loxP* flanked
FRET                fluorescence resonance energy transfer
FRT                 *flpe* recognition target
G418                geneticin, an aminoglycoside antibiotic
GENSAT              gene expression nervous system atlas
GFP                 green fluorescent protein
HAT                 hypoxanthine-aminopterin-thymidine
HepG2               human hepatocellular liver carcinoma cell line
hESC                human embryonic stem cell
HLA                 human leukocyte antigen
Hprt                hypoxanthine phosphoribosyltransferase
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSC</td>
<td>hematopoietic stem cell</td>
</tr>
<tr>
<td>ICM</td>
<td>inner cell mass</td>
</tr>
<tr>
<td>IL-6</td>
<td>interleukin-6</td>
</tr>
<tr>
<td>IMEM</td>
<td>Iscove’s Modified Eagle Medium</td>
</tr>
<tr>
<td>IRES</td>
<td>internal ribosome entry site</td>
</tr>
<tr>
<td>kan</td>
<td>kanamycin</td>
</tr>
<tr>
<td>kb</td>
<td>kilo base pair</td>
</tr>
<tr>
<td>lacZ</td>
<td>β-galactosidase</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertania</td>
</tr>
<tr>
<td>LIF</td>
<td>leukemia inhibitory factor</td>
</tr>
<tr>
<td>Lpl</td>
<td>lipoprotein lipase</td>
</tr>
<tr>
<td>M-CSF</td>
<td>macrophage colony-stimulating factor</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MSC</td>
<td>mesenchymal stem cell</td>
</tr>
<tr>
<td>Myh11</td>
<td>smooth muscle myosin 11</td>
</tr>
<tr>
<td>Neo</td>
<td>neomycin (aminoglycosidase 3′-phosphorylase)</td>
</tr>
<tr>
<td>OC</td>
<td>osteogenic culture</td>
</tr>
<tr>
<td>Ocn</td>
<td>Osteocalcin</td>
</tr>
<tr>
<td>OI</td>
<td>osteogenesis imperfecta</td>
</tr>
<tr>
<td>Osx</td>
<td>osterix</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PFGE</td>
<td>pulse field gel electrophoresis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>poly A</td>
<td>polyadenylation signal sequence</td>
</tr>
<tr>
<td>QRT-PCR</td>
<td>quantitative real time – polymerase chain reaction</td>
</tr>
<tr>
<td>RANKL</td>
<td>receptor for activation of nuclear factor kappa B ligand</td>
</tr>
<tr>
<td>Rpl7</td>
<td>ribosomal protein L7</td>
</tr>
<tr>
<td>RSV</td>
<td><em>Rous sarcoma</em> virus</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase – polymerase chain reaction</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SOC</td>
<td>super optimal broth with catabolite repression</td>
</tr>
<tr>
<td>STAT</td>
<td>signal transducer and activator of transcription</td>
</tr>
<tr>
<td>SV40</td>
<td><em>Simian virus 40</em></td>
</tr>
<tr>
<td>TE</td>
<td>Tris/EDTA buffer</td>
</tr>
<tr>
<td>TGF-β</td>
<td>transforming growth factor beta</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>wt</td>
<td>wild type</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

1.1 Development of the Skeleton

1.1.1 Bone Formation

Bone formation can occur via two different processes: intramembranous ossification, where mesenchymal cell condensations differentiate directly into osteoblasts that subsequently secrete bone matrix, or endochondral ossification, whereby mesenchymal cell condensations differentiate into chondrocytes that secrete cartilage matrix, forming a cartilage template that, after vascular invasion is resorbed and replaced by osteoblasts and bone matrix (Figure 1.1). Craniofacial bone, the dermal skull roof, cranial sutures, and parts of the mandible and clavicle are derived from intramembranous ossification, whereas bones that participate in joints and bear weight such as long bones of the limbs, vertebrae, and ribs, as well viserocranial, neurocranial and occipital bones, are formed through endochondral ossification (1).

Endochondral ossification relies on the formation of cartilage template, where chondrocytes express cartilaginous matrix proteins including Type II collagen, Type IX collagen, Type XI collagen, and Aggrecan. After a period of rapid proliferation during which the template elongates to create the shape of the skeletal elements it prefigures, the cells in the center of the template stop proliferating and differentiate into hypertrophic chondrocytes, now expressing Type X collagen. Concurrently, thin layers of cells on the periphery of the template differentiate into osteoblasts, forming the bone collar. Hyper-
Figure 1.1: Intramembranous ossification (a-d) and endochondral ossification (f-j). (a) Intramembranous ossification, in this case of frontal bones, occurs directly from mesenchymal condensations. (b) From the mesenchymal condensations, osteoblasts lay bone matrix toward the top of the skull. (c) The frontal bones meet at the midline, forming a suture. (d) Within the suture, mesenchymal condensations differentiate into osteoblasts. (e) A frontal view of a mouse skull (18.5 dpc) stained with alizarin red (mineral) and alcian blue (cartilage). By contrast, endochondral ossification is dependent on a cartilage template intermediate. (f) Again, formation of a mesenchymal condensation is the initial step in long bone formation. (g) The mesenchymal condensation gives rise to chondrocytes, forming the cartilage template; cells on the periphery form the perichondrium. (h) The proliferating chondrocytes cause the cartilage template to elongate, the chondrocytes then differentiate into hypertrophic chondrocytes. (i) The perichondrium differentiates into osteoblasts which invade the cartilage template; blood vessels also begin to invade at this time. Osteoclasts degrade the cartilage template while osteoblasts deposit bone matrix. (j) A mouse humerus stained with alizarin red and alcian blue (Figure and legend adapted from source (2)).
trophic chondrocytes begin to undergo apoptosis upon invasion of the region by vascular
tissue and osteoblasts from the bone collar. The cartilaginous matrix is replaced as
osteoblasts deposit bone extracellular matrix. Type I collagen constitutes 95% of the
bone extracellular matrix while non-collagenous proteoglycans and glycoproteins such as
Osteocalcin, Osteopontin, Fibronectin, and Bone sialoprotein constitute the remaining
5%. Osteoblasts deposit crystalline salts of calcium and phosphate in the form of
hydroxyapatite in the organic matrix, which provides both enhanced structural stability
and a physiological depot of these minerals. Both endochondral and intramembranous
ossification depend on matrix deposition and mineralization by osteoblasts.

The dynamic of bone remodeling is dependent on osteoblasts not only to
synthesize new bone, but to regulate the activity and differentiation of osteoclasts, the
bone resorbing cells. Osteoclasts are specialized macrophage polykaryons of the
monocyte/macrophage family, which demineralize bone through acidification of the
matrix by vacuolar H+-adenosine triphosphatase, and degrade the collagenous matrix by
the lysosomal protease, cathepsin K (3-7). Macrophage colony-stimulating factor (M-
CSF) and Receptor for activation of nuclear factor kappa B ligand (RANKL), both made
by osteoblasts, have been shown to be necessary for the proliferation and differentiation
of osteoclasts (8-13). While M-CSF is a secreted product, RANKL is located on the
surface of the osteoblast, requiring contact between osteoclast progenitors and
osteoblasts, or marrow stromal cells, for the initiation of osteoclast differentiation (9;10).
Osteoblasts also secrete Osteoprotegerin, a competitor of RANKL which attenuates
osteocalstogenesis, thereby regulating the quantity of bone resorbed (12;14;15). Ongoing
resorption by osteoclasts and rebuilding by osteoblasts serves to regulate bone mass
throughout adult life. If the amount of bone deposited by osteoblasts is insufficient to replace the amount resorbed by osteoclasts, this negative balance results in osteoporosis.

1.1.2. Transcriptional Regulation of Bone Formation

In the process of bone formation, many transcription factors are required to direct osteoblastic differentiation. Runx2, a member of the Runt-domain family of transcription factors, has been shown to be an essential player in bone development. Also known as Cbfa1 or Osf2, Runx2 contains a conserved 128 amino acid runt domain which binds to the DNA consensus sequence PyGPyGGtPy. Runx2 was originally discovered by binding affinity for DNA sequences in the promoter of Osteocalcin and was subsequently found to have recognition sites in the promoters of multiple genes indicative of osteoblast function such as Type I collagen, Bone sialoprotein, and Osteopontin (16;17). Osteocalcin is the most abundant non-collagenous matrix protein secreted by osteoblasts and has been shown to be largely osteoblast specific (18-20). It has a high affinity for calcium and phosphate ions and hydroxyapatite crystals, suggesting a role in bone matrix mineralization (21-23); a role has also been implicated for Osteocalcin as a matrix signal in the recruitment and differentiation of osteoclasts (24). Ducy et al. showed that Runx2 mRNA levels peak in the mouse embryo at 12.5 days post coitum (dpc) where expression is restricted to mesenchymal condensations of the developing skeleton, and cells of the osteoblast lineage (17). Runx2 mRNA is also detected in hypertrophic chondrocytes and is thought to induce differentiation, regulate Type X collagen expression, and promote endochondral ossification. (25-30). It has also been shown that transfection of Runx2 into non-osteoblastic cells leads to the expression of osteogenic genes – including
Osteocalcin, Bone sialoprotein, and Type I collagen (17). Runx2 has been demonstrated to be necessary for bone formation, as Runx2 deficient mice display a lack of mature osteoblasts and recognizable bone matrix (27;31).

Runx2 expression has been detected as early as 10.5 dpc, yet mature osteoblasts are not typically seen until ~13-14 dpc, and replacement of the cartilage template by bone is not observed until 15 dpc (17;32). This seeming inconsistency was ultimately explained by the discovery that Runx2 function is negatively regulated by the Twist proteins, which are expressed prior to and in concert with Runx2 in the earliest stages of skeletal development (32). Twist1 and Twist2 are basic helix-loop-helix transcription factors. They contain a 20 amino acid domain, called the Twist box, which binds the DNA binding domain of Runx2 to inhibit its function. Only upon decreased expression of the Twists can Runx2 fully promote osteoblast differentiation. Twist1 is expressed predominantly in the skull; Twist2 is expressed predominantly in the lower skeleton. Twist1 +/− mice, with insufficient Twist expression to deter Runx2 function, present a craniosynostosis phenotype, including premature osteoblast differentiation, larger intraparietal bones, and early fusion of the coronal sutures. Conversely, Runx2 +/- mice have delayed ossification of the cranial bones with wide cranial sutures. Compound heterozygotes, Runx2+/− Twist1+/−, have skulls of normal shape with average size intraparietal bones and no premature fusion of the coronal sutures, demonstrating that Runx2 haploinsufficiency is rescued by Twist1 +/- heterozygosity, and visa versa. Twist2 null embryos display premature osteogenic development of the ribs and clavicles, and only a complete deletion of Twist2 is able to rescue the clavicle hypoplasia of Runx2 haploinsufficient mice (32).
Another transcription factor has been shown to be essential for bone development. Osterix (Osx) is a zinc finger-containing transcription factor that is specifically expressed in endochondral and intramembranous bones (33). Mice heterozygous for a null allele of Osx are normal, while homozygous Osx-mutants die immediately after birth. Analysis shows that despite normal cartilage formation, the null embryos have no bone. In the appendages there is an invasion of the cartilage template by dense mesenchyme, however these cells are arrested in differentiation and no ossification occurs. These Osx null mutants have severely reduced expression of Type I collagen, and no expression of Osteocalcin. Runx2 expression in these mutants, however, is similar to that of preosteoblasts in wild type mice, indicating that Osx is not necessary for Runx2 expression. Runx2 null mice show no Osx transcription, placing Osx downstream of Runx2 (33;34). These findings allowed the de Crombrugghe group to suggest a model for Runx2 and Osx in osteoblast development, where osteoprogenitors in both endochondral and intramembranous bone require Runx2 expression to differentiate into preosteoblasts. The preosteoblasts express low levels of Type I collagen yet do not express any mature markers of osteogenesis. The preosteoblasts require Osx expression to differentiate into mature osteoblasts, expressing Type I collagen as well as characteristic osteoblast marker genes such as Bone sialoprotein, and Osteocalcin. Thus, in the Osx null mutants, preosteoblasts were missing the second signal and osteogenesis was arrested. The Osx null preosteoblasts however, did express several chondrocyte markers. From this they postulate that preosteoblasts are bipotential, with the ability to differentiate into either osteoblasts or chondrocytes (33).
Further experimentation by de Crombrugghe’s group supported the concept of a common osteochondro-precursor cell. To follow cell fate, transgenic mice were created to carry the Cre recombinase gene inserted into the 3’-UTR region of the Sox9 gene (Sox9-Cre); these mice were subsequently crossed with the Rosa26 LacZ reporter strain. Cells that had expressed Sox9 at any time during development were identified histologically after incubation with the X-gal substrate. Sox9 expressing limb bud cells gave rise to both chondrocytes and osteoblasts (35).

Sox9 is a high-mobility-group domain transcription factor expressed in chondrocytes and required for cartilage formation (36;37). Sox9 directly regulates transcription of Type II collagen (Col2a1) and Type XI collagen (Col11a2), essential collagens of cartilage formation (38–40). The following series of experiments have shown that Sox9 is essential for chondrocyte and osteoblast differentiation: First, in chimeric mice generated by injection of Sox9-/- ESCs into Sox9 +/- blastocysts, the -/- cells fail to contribute to any cartilage (36). Second, mice with tissue specific Sox9 genomic deletions, obtained by breeding a loxP flanked (floxed) Sox9 allele to homozygosity in the presence of a Prx1-Cre transgene, displayed a lack of mesenchymal condensations at 12.5 dpc and a resultant absence of cartilage or bone thereafter, [Prx1 is expressed in the limb bud prior to the mesenchymal condensation.] Runx2 expression was also absent in these mutants, placing Sox9 upstream of Runx2 (41). Third, mice with Osterix-deletion in Sox9-expressing cells, obtained by breeding mice with one Osterix lacZ knock-in allele and one floxed Osterix allele in the presence of the Sox9-Cre transgene, resulted in arrested osteoblast differentiation of limb mesenchymal cells and therefore skeletal elements that were hypoplastic, bent, and deformed. No mineralization was observed in
facial and skull templates generated by intramembranous bone formation, while parts of the endochondral-derived maxilla, mandible and parietal bones were calcified. Due to the lack of membranous bone formation, it was concluded that Sox9-expressing cells are the precursors of osteoblasts in membranous bone (35).

A number of other transcription factors have roles in bone formation including Msx1, Msx2, Dlx5, Dlx6, Krox-20 and Sp3. However several of these perform homeotic functions and act largely as patterning signals in the developing skeleton. While mutations in these genes result in an array of bone defects, none are recognized as being essential for bone development.

Genetically altered mice have provided key insights into the role of, and requirement for, transcription factors in osteogenic differentiation. However, through the study of in vitro systems, including ex vivo differentiation of bone marrow derived mesenchymal stem cells and embryo-derived stem cells, one can begin to more closely dissect the molecular signals required to induce differentiation, as well as gain the ability to generate osteogenic cells for therapeutic applications.
1.2 Osteoblastic Differentiation of Stem Cells Ex Vivo

1.2.1 Mesenchymal Stem Cells

Mesenchymal Stem Cells (MSC), also known as marrow stromal cells, are adult stem cells capable of multi-lineage differentiation into several mesenchymal cell types such as osteoblasts, chondroblasts, myoblasts, fibroblasts and adipocytes (Figure 1.2). The initial identification and characterization of MSCs was performed by Friedenstein in the 1960’s and 70’s based on bone marrow derived cells (42-44). By culturing whole bone marrow on plastic dishes with repeated washes to remove the non-adherent cells, he observed the appearance of colonies with fibroblast morphology, each derived from a single precursor cell. These cells, thought to be from the bone marrow stroma, displayed high osteogenic potential and their multipotential capabilities were soon demonstrated by differentiation into other mesenchymal cell types (44-48). Other adult and fetal tissues such as blood, adipose tissue, muscle, dermis, placenta, synovial tissue, and pancreas have also been shown to contain MSC populations (49-61).

MSC-derived osteogenic precursors make up a very small percentage of the bone marrow population. Estimates by limiting dilution plating of cells from digested fetal rat calvaria show that these osteogenic colony forming cells constitute less than 1% of the population, or from rat and mouse bone marrow stroma ~0.5-1x10^{-5} of the population (62-64). When cultured in vitro, in medium containing ascorbic acid, β-glycerophosphate, and fetal calf serum, MSCs tend to have a default differentiation pathway towards osteogenesis. Glucocorticoids such as dexamethasone and cytokines such as bone morphogenetic proteins (BMPs) can further promote osteogenesis and increase the number
Figure 1.2: Multilineage potential of mesenchymal stem cells (65).
of bone colonies formed (66-68). *In vitro* osteogenic differentiation is identified morphologically by the deposition and mineralization of matrix, as well as by characteristic gene expression.

The MSC population within the bone marrow stroma is quite small, and heterogeneous with regard to size, morphology, and differentiation potential. Thus attempts have been made to identify cell surface markers to permit identification and purification MSCs (69-74). At this time however, most immunophenotyping approaches only permit separation of MSCs from hematopoietic stem cells and do not distinguish multipotent MSCs from other precursor cells within the MSC population (reviewed in (75;76)).

The multipotentiality of MSCs has made them promising candidates for cell therapy. In several animal studies, transplanted MSCs from culture have been able to regenerate tissue in trauma-damage or disease models, including disc cartilage, cardiomyocytes, neurons, and epithelia (77-80). In a model of osteogenic regeneration, introduction of calcium phosphate ceramics loaded with marrow cells allowed for complete bone union in 8 of 12 implants, whereas unloaded ceramics introduced alone into a femoral defect resulted mainly in the formation of fibrous tissue (81).

MSCs also have the potential to be used therapeutically for inherited osteogenic disorders. In a mouse model of the brittle bone disease osteogenesis imperfecta (OI), preliminary experiments suggested that transplanted MSCs produced progeny that resided within the bone of recipient mice, and significant increases in collagen and mineral content were reported (82). OI is a disease resulting from mutations within the Type I collagen helix presenting a range of clinical manifestations, from the mildest form
with mild bone fragility and normal stature, to the most severe form resulting in intra-
uterine fractures, severe bone malformations, and perinatal death. Transplant of MSCs
with normal Type I collagen may engraft the bone marrow, produce progeny capable of
secrating the correct form of Type I collagen, assemble normal matrix, and correct the
defect.

In bone marrow transplant trials in OI patients, Horowitz et al. reported the
engraftment of donor cells as well as increases in body length and bone mineral content
(83;84). However, the reported level of osteoblast engraftment was quite low, with 2%
of cells within trabecular bone of donor origin, suggesting that the results of increased
length and mineral content should be interpreted with caution. Encouraged by these
results though, LeBlanc et al. has reported engraftment in bone (0.3% - 7.4%
engraftment) after in utero transplantation of fetal MSCs in a patient with OI (85). A
large majority of OI cases are diagnosed in utero or in early infancy, an ideal time for
stem cell therapy, as the immature immune system may be more permissive to the
induction of immune tolerance from non-HLA matched donor cells and levels of
engraftment are higher in periods of rapid growth.

Recent experiments suggest that MSCs with a preference for engraftment within
bone and cartilage after intravenous injection can be elicited (86;87). In this paradigm,
MSCs marked with GFP are transplanted into neonatal mice via the superficial temporal
vein, retrieved from recipient femurs after 25 days, introduced again to neonates
intravenously, retrieved from femurs at day 35, expanded in culture, and transplanted into
normal or OI neonatal mice. Evaluation 2 to 4 weeks after the final transplantation
revealed GFP positive cells within the bone of recipient mice, including active areas of
bone formation; when using early passages of the MSCs, no GFP positive cells were detected in the lungs or liver.

While experimentation with MSCs has provided valuable information on osteogenic differentiation, there are also many limitations. For instance, because the MSC is a progenitor or a pre-engaged cell from an adult organism, it allows the study of only the later stages of differentiation (88). Also MSCs are rare in adult tissues and no universal cell surface marker has been identified to aid in enrichment. Finally, MSCs have a finite capacity for self renewal, leading quickly to senescence, decreased differentiation capacity, and loss of stem cell characteristics, limiting in vitro expansion (reviewed in (88)). In light of these limitations, embryonic stem cells have emerged as an alternative in vitro system for the study of osteogenic differentiation.

1.2.2 Embryonic Stem Cells

Embryonic stem cells (ESCs) are pluripotent cells derived from the inner cell mass (ICM) of a pre-implantation blastocyst. The outer layer of the blastocyst is composed of trophoblast cells which are required for uterine implantation and will contribute to the placenta. The ICM has the ability to form the three somatic germ layers (endoderm, ectoderm, and mesoderm), primordial germ cells, and two extraembryonic lineages of the yolk sac (reviewed in (89-91)) (Figure 1.3). The ICM of an embryo at the expanded blastocyst stage can be separated from the trophoblast by immunosurgery. Alternatively the embryo is kept intact, plated onto a mitotically inactivated mouse embryonic fibroblast feeder layer, where the trophoblast layer will split open permitting
Figure 1.3: Derivation of embryonic stem cells. ESCs are derived from the inner cell mass of a pre-implantation embryo and have the ability to form the three somatic germ layers. The outer layer of the blastocyst is composed of trophoblast cells which will give rise to the placenta (source (89)).
the outgrowth of the ICM. The cells are expanded, replated onto new feeder layers, and if secondary colonies of undifferentiated cells arise, further expansion allows the creation of ESC lines (89;91).

Derivation of ESCs was first demonstrated in 1981 by Evans and Kaufman, and Martin (92;93). Prior, early differentiation studies had focused on pluripotent embryonic carcinoma cells derived from teratocarcinomas. However, teratocarcinoma cells are aneuploid, contribute poorly to chimeras, show poor differentiation potential, and are unable to produce mature gametes (91). Early work with embryo-derived ESCs guided the establishment of a standard definition for ESCs, including: 1) derivation from the pre-or peri-implantation embryo, 2) prolonged proliferation in the absence of, or with minimal, differentiation, 3) potential to form derivatives of all three embryonic germ layers even after prolonged culture, and 4) for murine ESCs, the ability to contribute to the germ line in chimeras (94).

Maintenance of an undifferentiated state was at first dependent on co-culture with feeder cells. Later it was discovered that feeder-conditioned medium was able to support ESC self renewal and that direct contact of ESCs with the feeder layer was not essential to maintain the undifferentiated state (95). Leukemia inhibitory factor (LIF) was identified as the feeder-cell-derived molecule that could sustain murine ESC self renewal in the absence of feeders (96;97). LIF is a member of the IL-6 cytokine family which functions through the gp130 receptor to activate STAT3. Activated STAT3 is a transcription factor which suppresses differentiation by controlling the expression of genes essential to ESC self renewal (98-100). When cultured in medium containing LIF, murine ESCs have the ability to self renew for prolonged periods without undergoing
senescence. Interestingly, LIF is insufficient to sustain human ESCs (hESCs) in an undifferentiated state as human ESC self-renewal is STAT3 independent (94;101;102). It was also determined that BMPs, within the fetal bovine serum of culture medium, played a major role in maintaining a pluripotent state. BMP4 was shown to act synergistically with LIF to maintain pluripotency in the absence of serum (103;104).

Upon removal of LIF, murine ESCs begin to differentiate, and under the appropriate conditions, cultures can produce cells from all of the embryonic germ layers. Murine ESCs are referred to as pluripotent as they are capable of generating the entire fetus, but cannot produce trophectoderm. Human ESCs however, are capable of generating tissues of all the embryonic germ layers, yet have also been shown to be capable of differentiating into trophectoderm (94;105;106). In contrast, adult stem cells have the ability to give rise to only a limited number of tissue specific cell types.

To initiate ESC differentiation in vitro, three approaches have been established. One approach depends upon aggregation of ESCs to form embryoid bodies (EB), a second approach relies on co-culture of ESCs with stromal cells, and the last approach involves culturing ESCs in a monolayer on extracellular matrix proteins (reviewed in (90)). Using these approaches many differentiation protocols have been established for an array of specific lineages, further demonstrating the pluripotency of ESCs. Differentiation methodology has been developed for ectodermal lineages such as neurons, astrocytes, oligodendrocytes, and epithelial cells; for endodermal lineages such as pancreatic beta cells and hepatocytes; and for mesodermal lineages such as chondrocytes, adipocytes, smooth muscle, and of particular interest and further discussion – osteoblasts.
Multiple culture protocols have been developed for the osteogenic differentiation of murine embryonic stem cells (Figure 1.4). Most protocols rely on the preliminary step of EB formation, first allowing ESC aggregation. In one, Buttery et al. allowed murine EBs to form in suspension for 5 days in culture, then dissociated the EBs and plated the cells at a density of $5 \times 10^4$ cells/10 cm$^2$, allowing adherence to tissue culture plastic (Figure 1.4A Top). The cultures were maintained for 21 days and medium was tested with the addition of 1) ascorbic acid and β-glycerophosphate; 2) ascorbic acid, β-glycerophosphate, and dexamethasone; and 3) ascorbic acid, β-glycerophosphate, and retinoic acid. Confluent plates containing nodular colonies were identified as osteogenic by immunostaining for Osteocalcin and Type I collagen, as well as the calcium stain Alizarin Red. Cultures supplemented with ascorbic acid, β-glycerophosphate, and dexamethasone after 14 days of culture showed the greatest number of mineralized bone nodules. Retinoic acid had no additive effect (107).

In an alternate protocol, EBs were maintained in suspension through a ‘hanging drop’ method (108;109) (Figure 1.4 B). In this system, cells were seeded onto the inner side of the lid of a Petri dish and allowed to aggregate at the base of the drop for 2-3 days to form EBs. The EBs were transferred to bacterial grade plastic to discourage adherence, with or without supplementation with retinoic acid. At day 5, intact EBs were plated onto tissue culture plastic and allowed to adhere and differentiate for approximately 20 days. The medium was supplemented with ascorbic acid and β-glycerophosphate, and in some cultures the effect of supplementation with 1α,25-OH vitamin D3, compactin, or BMP-2 was tested. Successful differentiation of osteoblasts was demonstrated by immunostaining, Alizarin Red staining, and gene expression.
Figure 1.4: Culture systems for the osteogenic differentiation of ESCs. (A) ESCs are cultured on feeders; cells are removed from the feeders and either (Top) suspended as EBs before being plated or (Bottom) placed directly on Petri dishes for 1 day followed by plating as a single cell suspension (figure and legend adapted from (110)). (B) In the hanging drop method, cells are seeded on the inner side of the lid of a Petri dish and allowed to condense for 2 days. After the 2 days, the EBs aggregate in suspension for 3 days before being plated for the osteogenic culture (figure and legend adapted from (88)).
characteristic of osteogenesis. Ascorbic acid and β-glycerophosphate were determined to be necessary for deposition and mineralization of the extracellular matrix (109).

While early osteogenic differentiation experiments were performed using murine ESCs, experiments on human ESCs followed closely afterwards. Sottile et al. and Bielby et al. both used the 5 day EB suspension protocol to culture hESCs; it was determined that the addition of the glucocorticoid, dexamethasone, to the medium was able to greatly enhance osteogenic differentiation (111;112). Karp et al. compared this protocol to direct plating of the hESCs, omitting the EB step (110) (Figure 1.4A Bottom). By omitting the EB step, they observed a 7.6 fold greater number of bone nodules formed, as well as earlier formation of the nodules (at 10-12 days as compared to 4 weeks with the EB step). However, both Alkaline phosphatase and Osteocalcin levels in these cultures remained low. The mineral-to-matrix ratio of these cultures was also lower than when the EB step was included. Interestingly, the mineralized nodules from the cultures without the EB step formed regardless of whether ascorbic acid, β-glycerophosphate or dexamethasone was added to the medium, all suggesting that the nodules may be representative of another non-osteogenic lineage.

Mixed results have also been reported in attempts to culture murine ESCs without the EB step. While early and increased mineralization is seen in these cultures, it is thought that an unconventional osteogenic process is occurring as Type I collagen and Osteocalcin secretions are delayed and the matrix is disorganized (113;114). Treatment of the ESCs with conditioned medium from a human hepatocarcinoma cell line (HepG2) however, resulted in mineralized nodules with strong expression of Type I collagen as well as characteristic osteogenic gene expression (115). It is suggested that these
contrasting reports may be influenced by the use of different ESC lines (R1, D3, and E14Tg2a cells have all been used).

Our method of initiating osteogenesis, where day 2 EBs are dispersed and plated at low density such that the potential and differentiation of a single cell can be observed, “is the only described method that allows for the observation and manipulation of the commitment stage of osteogenesis from single, embryonic progenitors” (116). In our system, a confluent plate of ESCs is trypsinized to release the colonies, brought to a nearly single-cell suspension, and plated densely back onto the original plate with no additional feeders. After 24 hours the colonies are flushed from the plate with medium and transferred to bacterial grade Petri dishes where the aggregates stay in suspension and form EBs for 2 days. EBs are dispersed to single cells and plated at 2x10^3 cells/10 cm dish and allowed to adhere. Cultures are maintained for 21 days in the presence of ascorbic acid and supplemented with β-glycerophosphate after day 7. Evidence for osteogenesis in these cultures includes cell and colony morphology, von Kossa staining of mineralization, quantitative RT-PCR from entire plates as well as individual colonies revealing expression of genes characteristic of osteogenesis, confocal microscopy demonstrating appropriate expression of an Osteocalcin-GFP reporter, and Type I collagen immunostaining (117-120). While other methods rely on plating an entire EB or a high density of cells from a dispersed EB, our method allows for the observation of the commitment stage from single progenitors. This permits analysis of individual colonies in a heterogeneous culture, with and without genetic or environmental manipulation (116;119). This idea of using a fluorescent reporter (Osteocalcin-GFP) to visualize individual colonies and cells in the osteogenic culture can be further implemented to
identify different stages of differentiation through the use of multiple gene-specific reporters and will serve as an essential tool in the understanding and application of osteogenic progenitor populations.
1.3 Transcriptional Reporters of Osteogenic Gene Expression

1.3.1 Random Integration

Most methods for the introduction of exogenous DNA into a genome involve a random integration event. Random integration however, can yield variable expression of the transgenic sequences resulting in the necessity to screen multiple lines for the desired transgene expression. Transgene expression is affected by the site of integration of the introduced sequences as well as the number of copies of the transgene that were integrated.

There are several ways the site of integration can affect transgene expression. Regulatory elements adjacent to the site of integration can override the minimal promoter sequences of the transgene to positively or negatively alter expression (121). The transgene may also be integrated in heterochromatin, or adjacent to heterochromatin, yielding position effect variegation where gene expression is silenced in a clonal subpopulation of cells due to the spreading of the heterochromatin (122). Finally, a transgene can insert into the coding sequence or regulatory element of an endogenous gene, resulting in aberrant or disrupted expression of that gene. Such insertional mutagenesis can cause phenotypic changes in the transgenic cell or organism that are mistakenly attributed to the transgene.

Transgene expression can also vary based on copy number. Transgenes mainly integrate at a single locus as head to tail concatamers consisting of a few to several hundred copies; rarely more than one integration event is observed. High copy number does not always correlate with high levels of transgene expression and can even exert a
repressive influence on expression due to increased methylation and chromatin compaction at the transgene locus (123). Silencing methylation of transgenes can also occur if prokaryotic sequences, such as the gene for bacterial beta-galactosidase (LacZ), are part of the integrated DNA (123;124).

### 1.3.2 Homologous Recombination

Genetic alteration through homologous recombination was first demonstrated by Smithies and Capecchi (125-128). Homologous recombination allows for the targeted insertion, or replacement, of transgenic sequences at a chosen locus (Figure 1.5 A). Mansour et al. produced one of the first targeted genetic reporters by placing the *Escherichia coli* (*E. coli*) LacZ gene into the murine Int-2 locus (129). This method avoids the pitfalls related to copy number, serves as a control for position effect upon targeting a construct within different cell lines (or different constructs within the same line), and does not allow unplanned insertional mutagenesis to occur. Targeting reporter sequences also increases the fidelity of reporter gene expression as expression is controlled by the endogenous promoter and regulatory elements. A consequence of this strategy is that in most instances the integrated transgene is expressed in replacement of the endogenous gene, creating the possibility for a haploinsufficient phenotype.

In 1996 a method was introduced by Bronson et al. to target transgenes in a single copy to a chosen integration site, the murine *Hypoxanthine phosphoribosyltransferase* (*Hprt*) locus (130) (Figure 1.5 B). The *Hprt* gene is a housekeeping gene embedded in a region of euchromatin, and placement of a transgene at this site minimizes position
Figure 1.5: Targeted insertion of transgenic sequences. (A) General schematic for knock-in targeting where a transgenic gene replaces the endogenous gene (source (131)). (B) Targeting transgenes to the Hprt locus of ESCs. Recombination restores the Hprt locus as well as inserts a single copy of the transgene (source (118)).
effects (132-134). By using the E14Tg2a ESC line containing a partial deletion of the
*Hprt* locus, the targeting vector is designed to rescue the deletion allowing for direct
forward selection of functional HPRT with hypoxanthine-aminopterin-thymidine (HAT)
medium (125;135). Selection in HAT medium is ideal as it identifies the
complementation of a functionally selectable gene, as opposed to selection schemes that
rely on the simple disruption of the *Hprt* gene. If the integration of transgenic sequences
was into a functional *Hprt* gene, the event could be selected for using 6-thioguanine,
which is toxic to Hprt expressing cells. However, 6-thioguanine resistance could be
achieved by several other events ranging from point mutations to loss of the entire X-
chromosome. Complementation of a partially deleted *Hprt* locus by homologous
recombination is the only event that will allow survival in HAT medium. Also, selection
through *Hprt* complementation minimizes the use of prokaryotic sequences for selectable
markers within the construct and the potential for methylation of the transgene or
surrounding genes. This method therefore, eliminates the complications of copy number,
position effect, insertional mutagenesis, and provides a means of direct forward selection.

Multiple small recombinant transcriptional reporters have been targeted to the
*Hprt* locus to create transgenic ESC lines and mice (136-159) (Table 1.1). Many of these
reporters were generated to characterize promoter fragments, binding motifs, and
conserved regulatory or enhancer elements. The size limitation of genomic fragments
carried in conventional targeting plasmids (~10-25 kb) however, constrains the
assessment of reporter gene expression under the control of more than the most proximal
regulatory elements. To overcome this obstacle, Heaney et al. introduced a method of
targeting bacterial artificial chromosomes (BAC) to the *Hprt* locus (118) (Figure 1.6).
<table>
<thead>
<tr>
<th>Construct</th>
<th>Tissue Expression</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tie2-LacZ</td>
<td>Endothelial cells of the vasculature</td>
<td>Evans et al. Physiol Genomics. 2000</td>
</tr>
<tr>
<td>α-MHC-LacZ</td>
<td>Cardiomyocytes</td>
<td>Misra et al. BMC.Biotechnol. 2001</td>
</tr>
<tr>
<td>Flt-1-LacZ</td>
<td>All vascular beds except liver</td>
<td>Minami et al. Blood. 2002</td>
</tr>
<tr>
<td>Cis-NF-kappa B-EGFP</td>
<td>LPS and anti-CD3 inflamed tissue</td>
<td>Mangess et al. J. Immunol. 2004</td>
</tr>
<tr>
<td>α-MHC-LacZ</td>
<td>Cardiomyocytes</td>
<td>Rudy-Reil et al. Circ Res. 2004</td>
</tr>
<tr>
<td>MBP Mod4-LacZ</td>
<td>Schwann Cells</td>
<td>Denarier et al. J Neurosci. 2005</td>
</tr>
<tr>
<td>Gdf9-GFP</td>
<td>Oocytes</td>
<td>Yan et al. Biol Reprod. 2006</td>
</tr>
<tr>
<td>Robo4-LacZ</td>
<td>Endothelium</td>
<td>Okada et al. Circ Res. 2007</td>
</tr>
<tr>
<td>pUbi-mRFP-1</td>
<td>Ubiquitous</td>
<td>Yurchenko et al. Transgenic Res. 2007</td>
</tr>
<tr>
<td>ARR2PB-CreER</td>
<td>Prostate</td>
<td>Luchman et al. Genesis. 2008</td>
</tr>
<tr>
<td>TAL1-LacZ</td>
<td>Hematopoietic stem cells</td>
<td>Smith et al. Genome Res. 2008</td>
</tr>
<tr>
<td>α-MHC-CFP</td>
<td>Cardiomyocytes</td>
<td>Pandya et al. Proc Natl Acad Sci. 2008</td>
</tr>
<tr>
<td>CAG-EGFP</td>
<td>All brain regions analyzed</td>
<td>Yang et al. Genomics. 2009</td>
</tr>
<tr>
<td>GR-tetO-LacZ</td>
<td>Cardiomyocytes, forebrain, liver and kidney*</td>
<td>Palais et al. Physiol Genomics. 2009</td>
</tr>
</tbody>
</table>

The promoters and expressed genes are shown. eNOS = endothelial nitric oxide synthase, MHC = myosin heavy chain, vWF = von Willebrand factor, MBP = myelin basic protein, SRF = serum response factor, Gdf9 = growth differentiation factor 9, pUbi = human ubiquitin C promoter, SM22 = smooth muscle 22 alpha, ARR2PB = androgen response element-probasin promoter, Plp/DM20 = proteolipid protein DM20, CAG = cytomegalovirus enhancer/chicken beta-actin promoter, GR-tetO = glucocorticoid receptor/tetracycline operator. *GR-tetO-LacZ mice were crossed with MHC-tetOFF, CamKII-tetOFF, and LAP-tetON2 mice for tissue specific expression.
Figure 1.6: Modification and targeting of BACs to the Hprt locus. (A) *In vitro* Cre-mediated recombination of the BAC-modifying fragment into a BAC at the *loxP* site. (B) Targeting of an *I-SceI*-linearized BAC transgene to the *Hprt* locus. The gray box represents the ~36 kb deletion at the mutant *Hprt* locus. (Figure and legend modified from (118)). Note - Figure B is a BAC specific schematic of Figure 1.5 B.
BACs can carry genomic inserts as large as 300 kb allowing for the transgenic insertion of very large genes as well as distant regulatory sequences associated with smaller genes. For BAC Hprt targeting, the advantage of direct forward selection at the Hprt locus allowing identification of rare recombination events was indispensable, as the frequency of targeting BACs to the Hprt was shown to be lower than 1 per 5,000,000 cells electroporated (118).

BACs are circular DNA molecules propagated by a vector carrying the fertility (F) factor in E. coli. To create a BAC, partially digested genomic DNA is ligated into a vector containing three F factor genes (parA, parB, parC), two genes that mediate plasmid replication (oriS, repE), and a wild-type loxP site (160;161). The BAC construct is electroporated into E. coli where the DNA is stable and accessible to manipulation.

The modification of BAC DNA within E. coli by homologous recombination has been termed ‘recombineering.’ The relative efficiency of homologous recombination in bacteria makes it a desirable alternative to the manipulation of recombinant DNA by restriction endonucleases and ligases, approaches that are largely impossible when working with large constructs due to the absence of unique restriction sites and the fragility associated with large DNA fragments. Woll et al. used this method of recombineering to create a BAC-based Osteocalcin-GFP reporter which was then targeted to the Hprt locus (119). This reporter was created by integrating an IRES-EGFP cassette under control of the Osteocalcin promoter, allowing for gene-specific expression of the reporter under a near-endogenous context. The Osteocalcin-GFP reporter BAC was targeted in a single copy to the Hprt locus of an ESC line. Subsequent confocal fluorescent microscopy of in vitro differentiated nodules from the osteogenic culture
system revealed the presence of GFP positive cells on the raised surface of the nodules as well as in the proliferative single cell layer on the periphery of the nodule – exactly where osteoblasts would be expected to be observed (paralleling the presence of mature osteoblasts on the bone surface), confirming both osteogenesis in these cultures and demonstrating the utility of the Osteocalcin reporter. While gene-specific reporters allow the identification of a specific cell type within a mixed population, these reporters also can facilitate the isolation or enrichment of a specific cell type for further characterization or application.

1.3.3 Enrichment of Subpopulations

Embryonic stem cells can differentiate into multiple lineages in vitro, suggesting the possibility of using ESC derived tissue for transplantation and therapy. Challenges that must be overcome are the heterogeneity of ESC differentiation cultures and the likelihood of forming teratomas in vivo. To utilize ESC-derived progenitors for clinical transplantation, methods to produce more homogenous lineage-specific cultures must be employed. Genetic reporters can mark specific lineages or cell types by ‘reporting’ on the expression of a gene specific to that lineage or cell type. GFP reporters have been used in conjunction with fluorescence activated cell sorting (FACS) to purify or enrich for specific populations (162-166). For example, ESC neural differentiation cultures are widely in use, however these cultures produce a mixed population that includes neurons, astrocytes, and oligodendrocytes. In 2002, Xian et al. created a GFP reporter for Olig2 gene expression to identify cells of the oligodendrocyte lineage within ESC neural differentiation cultures (164). FACS was used to physically separate the subset of GFP
positive cells from the culture and it was shown that the GFP positive population survived the sort procedure, and subsequently developed morphological characteristics of oligodendrocytes when reintroduced to culture.

The transplantation of ESC-derived cells has also been assessed using GFP reporters and FACS-methodology. Yin et al. demonstrated that liver progenitor cells can be purified from ESC-differentiation cultures when the ESCs are transgenic for an alpha-fetoprotein (AFP) GFP reporter (165). While undifferentiated ESCs injected intrasplenically formed teratomas, injection of differentiated AFP-GFP positive cells resulted in engraftment and differentiation into mature functional hepatocytes in vivo.

The ability to identify and isolate subpopulations within ESC cultures has also been used to understand the establishment of the primary germ cell layers. A Brachyury-GFP reporter ESC line was used to study the induction and development of the mesodermal and endodermal lineages. Using FACS for GFP from a Brachyury-GFP ESC differentiation culture, Fehling et al. were able to separate cells with mesodermal potential from those with neuroectoderm potential, and to study their specification to the hematopoietic lineage (162). Sorting based on the expression of Brachyury (Bry) and a fluorescent antibody to receptor tyrosine kinase Flk1 not only allowed the identification and enrichment of the hematopoietic lineage, but identified three distinct subpopulations ranging from pre-mesoderm to hemangioblast (GFP-Flk1-, GFP+Flk1-, GFP+Flk1+). It was further reported that ESC-derived Brachyury-GFP positive cells constitute a mes-endoderm progenitor and can be used not only for the induction and development of the mesodermal lineage, but the endodermal lineage as well (163).
1.3.4 Existing Osteogenic Reporters

1.3.4.1 Reporters Based on Sub-Cellular Localization

Tai et al. have created an Osterix-GFP reporter system to track Osterix activation and translocation within the cell (167). A 1.5 kb fragment of the full-length mouse Osx cDNA was used to create an Osx-GFP fusion protein and linked to a constitutive promoter. MSCs and ESCs were transiently transfected with Osx-GFP on days 4, 7, and 21 of the osteogenic differentiation culture; specifically, cells at the different days of osteogenic differentiation were primed for lipofection in reduced serum medium for 1 hour, incubated with Osx-GFP/lipofectamine plasmid cocktail for 3 hours, replaced with osteogenic medium and incubated overnight, and subsequently assessed by fluorescent microscopy. Uniform cytoplasmic distribution of GFP was observed in day 4 transfected cultures, GFP was visualized in the nucleus as well as the cytoplasm of day 7 transfected cultures, and in day 21 transfected cultures, most of the GFP was located in the nucleus. The authors contend that the translocation of Osterix to the nucleus at day 7 provides visual evidence of the commitment of precursor cells to the osteoblast lineage. While this reporter provides useful information on the intracellular location of Osterix, the use of a constitutive promoter does not permit the identification of Osx expressing cells.

Similarly, an EGFP-RUNX2 fusion protein has been created to visualize the nuclear localization of Runx2 (168). A reporter plasmid was created containing an EGFP cDNA fused to the Runx2 cDNA, under control of a constitutive promoter. Transfection of the reporter into human osteosarcoma SaOS-2 cells revealed a punctate nuclear pattern, demonstrating localization to subnuclear domains. These sites were shown to be areas of active transcription as they were also positive for BrdU labeling. Again, this
construct allows for visualization of the intracellular localization of Runx2 within the cells, rather than cells expressing endogenous Runx2.

1.3.4.2 Reporters Based on Gene Expression

Type I Collagen:

Rowe et al. created a series of osteogenic-GFP reporter transgenic mice using different fragments of the rat Type I collagen (Col1a1) promoter. Type I collagen is the predominant structural protein of bone, constituting approximately 95% of the bone matrix (1). However, Type I collagen is also found in tissues such as skin, tendons, ligaments, teeth, and gives structural integrity to major organs and blood vessels. By using a 3.6 kb or 2.3 kb Col1a1 promoter fragment (Col3.6GFP or Col2.3GFP), specific random integration events allowed different stages of osteoblast differentiation to be observed (Figure 1.7) (169;170). The Col3.6GFP construct appeared to identify a broad range of osteoblast differentiation, including a preosteoblastic stage when GFP expression was concomitant with alkaline phosphatase staining and Type I collagen mRNA expression (day 7 of MSC culture), as well as in mature nodules (day 21 MSC culture). Col3.6 GFP transgenic cells within primary MSC cultures displayed different levels of GFP expression, with cells surrounding the mineralized portion of the osteogenic nodules having weak expression, and cells located within the central mineralized areas having strong expression. Strong GFP expression was observed in vivo in osteoblasts lining the periosteal (outer surface of bone), endosteal (bordering the bone marrow cavity), and trabecular surfaces, as well as in cortical bone at the collar region of
Figure 1.7: General schematic of temporal expression of osteogenic-GFP reporter transgenic mouse lines. The Col3.6 GFP reporter identified a preosteoblastic stage of differentiation, with expression persisting throughout osteoblast maturation. Col2.3GFP initiated expression later than Col3.6GFP and identified a more mature osteogenic population. hOC-GFP identified the mature osteoblast. DMP1-GFP identified the terminally differentiated osteoblast, the osteocyte. All reporters created in the laboratory of Dr. David W. Rowe.
the growth plate. Weak expression was detected in periosteal fibroblasts of metaphyseal bone. This Col3.6GFP reporter was not, however specific to the osteogenic lineage as expression was detected in other Type I collagen producing tissues such as skin, aorta, bladder, lung, muscle, and fat (170).

GFP expression via the Col2.3GFP construct was more uniform and appeared to be restricted to more mature osteoblasts. Within calvarial and MSC cultures, GFP expression was concomitant with Bone sialoprotein mRNA expression (day 14 of culture), and was observed within osteocytes, terminally differentiated osteoblasts, within the mineralized portion of the nodule. In vivo it was observed in bone lining cells and osteocytes with the bone matrix (170).

**Osteocalcin**

The Col2.3GFP reporter appeared to identify mature cells of the osteoblast lineage, thus its temporal and spatial expression was compared to an Osteocalcin-GFP reporter (OC-GFP), because Osteocalcin expression has been shown to be largely specific to mature osteoblasts (171). The OC-GFP construct contained a 1.7 kb fragment of the rat Osteocalcin promoter linked to a GFP cDNA. Upon in vivo examination, the OC-GFP reporter and the Col2.3GFP reporter identified different populations. OC-GFP was expressed later in a limited subset of osteoblasts and osteocytes, and was also detected in the central nervous system in vivo. These results suggested that a larger fragment of the Osteocalcin promoter or another integration site would be necessary for more restricted Osteocalcin expression. A second Osteocalcin-GFP reporter (hOC-GFP) was made using a 3.9 kb fragment of the human Osteocalcin promoter (172). This reporter appeared to
mark mature osteoblasts \textit{in vitro} as expression was observed in cells within the mineralizing area of the bone nodule in MSC cultures, also Northern hybridization of FACS-sorted cultured calvarial cells revealed that the fluorescent cells expressed endogenous \textit{Osteocalcin} mRNA. \textit{In situ} hybridization of bone and tooth sections using an \textit{Osteocalcin} mRNA probe compared to adjacent sections viewed by fluorescent microscopy further revealed a colocalization of \textit{Osteocalcin} mRNA and hOC-GFP expression \textit{in vivo}.

\textbf{Dual Type I Collagen and Osteocalcin}

Transgenic hOC-GFP (Osteocalcin) mice were bred to Col3.6GFP (the preosteoblast to mature osteoblast reporter) mice to create transgenic animals which contain two distinguishable GFP-reporter variants, GFPtopaz and GFPcyan, respectively, permitting multiplex analysis of osteoblast differentiation (172). Three different stages of osteogenic differentiation were observed: First, cells with strong Col3.6GFP expression and hOC-GFP expression. These cells are located within the mineralizing region of the bone nodule \textit{in vitro}, or within regions of active bone formation on the bone surface and within the vascular areas of cortical bone \textit{in vivo}. Second, cells with strong Col3.6GFP expression and no hOC-GFP expression. These cells are located in and around the mineralizing bone nodule \textit{in vitro}, and at sites of new osteoblast differentiation in the growth plate \textit{in vivo}. Third, cells with weak or no Col3.6GFP expression cells and hOC-GFP expression. These cells are located within the bone nodule adjacent to the region of mineralization \textit{in vitro}. \textit{In vivo} they may be identified as hOC-GFP-positive-only cells as
the stronger hOC-GFP expression overpowers the weak Col3.6GFP expression. They are located within the endosteal surface of long bone and in the femoral neck region.

Dentin Matrix Protein 1

Kalajzic et al. describe a GFP reporter to mark the latest stage of the osteoblastic lineage, the osteocyte. Dentin matrix protein (DMP1) is restricted to the terminally differentiated osteoblast, the osteocyte. Testing of several fragments of the mouse DMP1 promoter identified a construct encompassing 7.8 kb of the promoter region and 4.4 kb of the first exon/first intron/and part of the second exon which yielded osteocyte specific GFP expression in osseous tissues. However, strong GFP expression was also detected in the brain (173).

While these reporters are able to identify different populations and/or different stages of cells within the osteoblast lineage, they also have limitations. By using discrete proximal fragments, other necessary distal native regulatory sequences are omitted resulting in reporters which may not completely reflect the expression of the endogenous gene. Also, these reporter constructs were all randomly integrated and transgene expression is likely influenced by the integration event.

1.3.4.3 Expression Reporter Systems

Tanaka et al. have recently reported the development of a highly sensitive Gene-Trap vector system which can be used as a reporter of gene expression (174) (For more information on Gene-Trap vectors, see Section 6.4.2, Figure 6.2). This Gene-Trap
system involves random insertional mutagenesis to introduce a neomycin resistant GFP expression cassette within ESCs. The Gene-Trap construct is transfected into a population of ESCs; ESCs that have integrated the construct into a functional gene are selected for in G418 containing medium, and the resultant ‘library’ of ESCs with unique integration events are subcloned and screened by 3’ rapid amplification of cDNA ends (3’ RACE). The 5’ end of the construct contains a splice acceptor sequence which ‘traps’ the upstream exons and allows the reporting of gene expression; the 3’ end of the cassette contains a splice donor sequence causing the reporter to be fused with any downstream exons. Therefore, insertion of the Gene-Trap vector results in a GFP-fusion protein under control of the endogenous promoter. This system was designed to be more sensitive by using a variant of GFP with a higher fluorescence intensity, Venus, as well as enhancing translation of the product by adding tandem repeats of a 9-nt sequence from the 5’UTR of the homeobox gene Gtx. Using this system, a gene-trap insertion within the Twist2 gene was obtained, generating a Twist2-VenusGFP ESC line. To analyze the spatiotemporal expression of the Twist2 reporter, chimeric embryos were generated by tetraploid aggregation and fluorescent expression was compared to whole-mount in situ hybridization with a Twist2 RNA probe. GFP expression patterns within day 9.5 and 10.5 embryos paralleled the in situ mRNA hybridization analysis for Twist2, with signal detected in mesenchymal cells in limb buds, somites, and tongue primordia (174).

While these initial results with the Twist2 Gene-Trap reporter are promising, there are also many caveats to this system. One main issue of the Gene-Trap system is that the loss of function mutation results in hemizygous expression of the endogenous gene. Twist2 is a major negative regulator of osteoblast differentiation and Twist deficiencies
have been shown to result in early expression of osteogenic genes as well as premature osteoblast development (32;175). Also, depending on the area of insertion within the gene, a Gene-Trap reporter can yield a neomorphic or dominant negative mutation, rather than the desired null mutation. Finally, isolation of similar reporters for other osteogenic genes relies solely on the arbitrary insertion of the reporter cassette within a gene of interest.

A gene specific approach has been taken by the Gene Expression Nervous System Atlas (GENSAT) BAC Transgenic Project, which has created GFP-BAC reporters for thousands of genes in order to analyze development of the central nervous system (176). BACs are modified to replace the gene of interest with an EGFP reporter cassette. The modified BAC is introduced by pronuclear injection of fertilized oocytes, and randomly integrated in multiple copies. Multiple transgenic lines are derived from the resulting founder mice, and compared to each other as well as to published data to confirm correct expression of the transgenic reporter. BACs containing Sox9 and Type I Collagen have been used to create transgenic lines, however published data has focused only on expression within the central nervous system. To use these reporters in an in vitro ESC osteogenic culture system, ESC lines would first have to be derived from the ICM of the transgenic animals. Also, the founder line from which the ESCs would be derived should undergo extensive characterization to ensure expression is not altered due to random integration of the BAC. Alternatively, if the modified BAC was available, a reporter-ESC line could be created by transfection of the BAC into ESCs.
1.4 Statement of Thesis

Embryonic stem cells (ESCs) are an ideal candidate for stem cell therapy based on their pluripotentiality and the ability to generate a large number of cells due to their nearly infinite self renewal capacity. ESCs for transplantation can be engineered to be histocompatible with the recipient by somatic cell nuclear transfer, or introduced in utero to a preimmune system and ultimately recognized as ‘self’. Because ESCs have been shown to form teratomas when transplanted in vivo, and because mature osteoblasts lack sufficient proliferative capacity, identification of therapeutic intermediates is a critical step in the translation of this approach to the clinic. ESC derived osteogenic progenitors can be used for the therapy of multiple musculoskeletal afflictions including severe fractures, resection of cancerous bone, and osteogenesis imperfecta.

Due to the limitations and ethical constraints on the use of hESCs, murine ESCs provide an important proof of concept for osteoprogenitor enrichment and subsequent in vitro and in vivo studies. Using the in vitro osteogenic nodule formation assay previously described, this thesis will develop and test a murine ESC-based system for identifying osteoprogenitors using BAC-based gene-specific osteogenic reporters targeted in single copy to the Hprt locus. Based on published in vivo and in vitro data regarding gene expression in osteogenesis, as well as our data on gene expression from ESC-derived osteogenic cultures, we have selected specific genes with expression characteristic of various stages in osteogenic differentiation and we hypothesize: (i) because Brachyury has been shown to be able to enrich for mesodermal populations, and axial and appendicular bone is of mesodermal origin, that a Bry-GFP+ osteoculture-derived
population is enriched for osteogenic progenitors, (ii) because \textit{Sox9} marks a bipotential osteochondroprogenitor and has been shown to be essential for intramembranous bone formation, that a \textit{Sox9-EGFP} \textsuperscript{+} osteoculture-derived population is enriched for osteogenic progenitors, (iii) because \textit{Twist2} has been demonstrated to play a major role in the regulation of osteoblast differentiation, that a \textit{Twist2-EGFP} \textsuperscript{+} osteoculture-derived population is enriched for osteogenic progenitors, (iv) because \textit{Osterix} is specific to the osteogenic lineage, marking a later timepoint in osteoblast differentiation, that an \textit{Osx-EGFP} \textsuperscript{+} osteoculture-derived population is enriched for lineage committed osteogenic progenitors, and (v) because subsets of bone in the viserocranium and neurocranium have been shown to be derived from ectodermal cranial neural crest cells, and as \textit{Nestin} is a known marker of neural crest cells, that a \textit{Nestin-GFP} \textsuperscript{+} osteoculture-derived population is enriched for neural crest derived osteogenic progenitors. (See Figure 1.8)
**Figure 1.8:** Specific transcriptional reporters can be applied to identify intermediate populations within the osteogenic culture.
CHAPTER 2
MATERIALS AND METHODS

2.1 Cell Culture and General Techniques

2.1.1 Cell lines and culture

The BK4 ESC line, deficient in Hypoxanthine guanine phosphoribosyl transferase, is a subclone of E14TG2a (177). ESCs were grown on murine embryonic fibroblasts (feeders) in DMEM-H (Dulbecco’s Modified Eagle Medium) (Life Technologies/Invitrogen, Gemini) supplemented with 15% fetal bovine serum (FBS) (Atlanta Biologicals), 0.1 mM 2-mercaptoethanol (Sigma), 2 mM Glutamax (Invitrogen), and leukemia inhibitor factor (LIF) conditioned supernatants (~ 1000 U/ml). ESCs were grown to sub-confluence, trypsinized to a nearly single cell suspension, and split 1:3 – 1:10 every two to three days. ESC cultures were maintained in humidified incubators at 37°C and 5% CO₂. Confluent plates of ESCs were trypsinized and frozen in a 5:4:1 ratio of FBS, ESC medium, and dimethyl sulfoxide (DMSO) (Sigma), respectively.

2.1.2 Osteogenic differentiation of ESC clones

The differentiation of ESCs into osteoblasts has been described previously (117;120). Briefly, ES cells were trypsinized to single cell and densely replated onto their original 10 cm tissue culture plates in ESC medium. After 24 hrs, the ESC colonies were flushed off the tissue culture plates with ESC medium, replated without LIF onto four 10 cm Petri dishes (bacterial grade plastic), and allowed to form embryoid bodies
(EBs). After two days of incubation at 37°C, the EBs were trypsinized to single cells, filtered through a 70 μm cell strainer (BD Falcon), and plated at 2E³ cells/10 cm dish (25 cells/cm²) – 1E⁵ cells/10 cm dish (1,300 cells/cm²) onto tissue culture plates in osteoprogenitor medium [IMEM (Iscove’s Modified Eagle Medium) (Life Technologies/Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen), 0.1 mM 2-mercaptoethanol, 2 mM Glutamax, and 0.2 mM ascorbic acid (Sigma)]. After seven days, β-glycerophosphate (βGP)(Sigma) was added to the osteoprogenitor medium at a final concentration of 2 mM to aid in mineralization. The cultures were allowed to differentiate for 21 days to form mature bone nodules. For re-plate studies after FACS, the osteogenic differentiation method is the same as listed above.

2.1.3 Fixation and staining of osteogenic cultures

Fixation and staining of osteogenic culture plates has been described previously (117;120). Plates were fixed in 10% neutral buffered formalin (Fisher) at room temperature for two hours. The plates were washed in distilled water (dH₂O) and stained with 2.5% silver nitrate in dH₂O for one hour on a UV light source (von Kossa stain). Plates were washed with dH₂O and counterstained for thirty seconds in methyl green (Vector Labs). The total number of colonies was counted in triplicate and the percentage of colonies that mineralized was determined.

2.1.4 ESC genomic DNA purification

ESCs were plated densely, without embryonic fibroblasts, and allowed to grow to confluency. ESCs were collected, pelleted and frozen at -80°C. After at least 1 hr
incubation at -80 °C, DNA was purified using the Wizard Genomic DNA purification Kit (Promega) following the manufacturer’s protocol for mammalian cells. DNA was resuspended in 100 μl of DNA Rehydration Solution.

2.1.5 PCR analysis of genomic ESC clone DNA for integration of transgene

Approximately 1 μg of ESC DNA was amplified in a 25 μl reaction containing 2.5 mM MgCl2, 0.2 mM dNTPs, and 1.5 U Taq DNA polymerase (Promega). Samples were PCR-amplified for 29 cycles as follows: 30 sec at 95°C, 30 sec at 60°C and 1 min at 72°C, followed by a final extension at 72°C for 5 min. Samples were electrophoresed through 2% agarose and stained with ethidium bromide for visualization. See Table 2.1 for a list of primers used.

2.1.6 RT-PCR analysis

RNA was collected from ESCs or ESC-derived osteogenic cultures with TriReagent (Molecular Research Center, Inc.) according to the manufacturer’s protocol. Two micrograms of ESC RNA was reverse transcribed with the RETROscript Kit (Ambion) using an oligo(dT) primer and following the manufacturer’s protocol. Samples were PCR-amplified for 29 cycles as follows: 30 sec at 95°C, 30 sec at 54°C and 1 min at 72°C, followed by a final extension at 72°C for 5 min. For the CAGGS-EYFP ESC line, primers amplified across chicken beta-actin intron 1 to create a 465 bp product if correct splicing of the RNA had occurred or a 1364 bp product if the intron was still present. The following primers were used: 5’-TCTGACTGACCGTGTTACTC-3’ and 5’-AAGTCGTGCTGCTTCATGTG -3’. For primers used to evaluate the Hprt-Nestin-GFP
ESC clones see Table 2.1. For primers used to evaluate Hprt-Osx-EGFP BAC clones see Table 2.2.

2.1.7 SYBR Green real-time quantitative RT-PCR

QRT-PCR was performed with the ABI PRISM 7900HT and 7300 Sequence Detection Systems and the Quantitect SYBR Green PCR kit (Qiagen) following manufacturer’s protocols. Serial dilutions of ESC cDNA were used to generate standard curves for the primer set. Relative mRNA expression was normalized to ribosomal protein L7 (Rpl7). An assay was designed to detect only RNA such that one primer spanned an exon/exon boundary. Duplicate or triplicate dilutions of 1:8 and 1:32 were analyzed for all cDNA preparations. The amplification program included an initial denaturation step at 95°C for 15 mins, followed by 50 cycles of 94°C for 15 sec, 53°C for 30 sec, and 72°C for 30 sec; fluorescence was measured at the end of each extension step. To verify the specificity of the primers, a melting curve was generated by heating the product to 95°C for 15 sec, cooling it to 53°C, and then slowly heating it at 0.03°C/sec to 95°C; fluorescence was measured during the slow heating phase. For quantification of EYFP in the CAGGS-EYFP ESC clones, the following primers were used: 5’-CGTTACTCCCACAGCTCCTG-3’ and 5’-GTGTACAGGATCGGTCTC-3’. For the primers used in the characterization of the osteogenic FACS sorted populations, see Table 2.2.
2.2 Creation and Evaluation of the EYFP ESC Reporter

2.2.1 Construction of CAGGS-EYFP-PNTlox2 vector

To construct the pCAGGS-EYFP-PNTlox2 constitutive reporter, an initial vector pCAGGS-EYFP, was generated by a blunt end ligation of: 1) an AgeI/MscI fragment from pCAGGS-Cre (a gift from J. Miyazaki M.D., Ph.D., Osaka University School of Medicine) which removed the Cre gene; and 2) a Smal/Xhol fragment from pIRES-EYFP (BD Biosciences/Clontech) containing the EYFP gene and polyadenylation signal sequence. A SalI/StuI fragment from pCAGGS-EYFP (containing the EYFP gene driven by a CMV enhancer and Chicken beta-actin promoter) was then blunt-end ligated into the HpaI site, which is located 5’ to a neomycin resistance gene flanked by two loxP sites in the same orientation, of pPNTlox2. (see Figure 2.1 for a schematic of the final construct)

2.2.2 Electroporation of CAGGS-EYFP-PNTlox2 into ESCs

A 5.7 kb NotI/Smal fragment of CAGGS-EYFP-PNTlox2, containing EYFP driven by the CMV enhancer and chicken beta-actin promoter, and a loxP flanked neomycin resistance gene, was electroporated into the BK4 ESC line at 3 nM/L. Following the electroporation, cells were plated on ten 10 cm plates with G418 resistant feeders. After 24 hours, 0.2 mg/mL G418 was added to the cells. Ten days after selection was added, random integrants were selected by their G418 resistance. Individual colonies were picked for expansion and verification of the integration event.
2.2.3 Metaphase counts of EYFP ESC clones

Cells were incubated for 2 hours in the presence of vinblastine to arrest cells in
the mitotic phase of the cell cycle. Cells were then trypsinized, removed from the plate,
and pelleted. A 0.075 M KCl hypotonic solution was added to the cells and they were
placed at 37°C for 30 minutes. The cells were then fixed in Carnoy’s fixative (consisting
of 1 part glacial acetic acid and 3 parts methanol) and subsequently lysed and spread on
slides by dropping cells from a pasteur pipet held 12 inches above the slide. After drying,
slides were stained in a Giemsa/Wright stain solution for visualization by microscopy.

2.2.4 Teratoma formation from EYFP ESC clone

One million CAGGS-EYFP ESCs or control BK4 ESCs in PBS were injected
bilaterally in the subcutaneous dorsum of 129/P3J mice. Mice were euthanized after 5
weeks, teratomas removed, and either fixed in 4% paraformaldehyde or freshly
embedded in Tissue-Tek O.C.T. Compound embedding medium (Sakura Finetek U.S.A.)
for cryosectioning. Embedded teratomas for cryosectioning were snap frozen and stored
at -80 °C. Upon sectioning, teratomas were cut into 10 micron sections, rehydrated in
PBS, and coverslipped with Aqua Poly/Mount (Polysciences, Inc.) for visualization by
fluorescent microscopy.

2.2.5 Fluorescence microscopy of EYFP ESC clones and teratoma tissue sections

For characterization of the CAGGS-EYFP ESC clones, time points included ESC
cultures, day 2 EBs, day 7 osteoculture (OC), day 14 OC, and day 21 OC. Cells and
teratoma sections were analyzed on an Axiovert s100 inverted Zeiss microscope (Carl Zeiss Incorporated) using a GFP/YFP cube/filter.

2.3 Creation and Evaluation of Gene Specific EGFP Reporters

(For an overall flow chart depicting the construction and characterization of BAC reporters, see Figure 2.1)

2.3.1 BAC clones

BAC clones RP24-291O6 (Twist 2) and RP11-141M17 (Sox9) were obtained from CHORI BAC PAC Resources. The Twist2-BAC is a C57BL/6J mouse BAC clone from the RPCI-24 library and is constructed on the pTARBAC1 vector backbone. The Sox9-BAC is a human BAC clone from the RPCI-11 library and is constructed on the pBAC3.6 vector backbone. BAC clone CTD-3071P8 (Osterix) was obtained from Invitrogen. The Osterix-BAC is a human BAC clone from the CITB D library and is constructed on the pBeloBAC11 vector backbone. BAC clones spanning specific genes were identified using BAC alignments at the UC Santa Cruz Genome Bioinformatics web site. BAC containing bacteria (strain DH10B) were propagated at 32°C in LB medium containing 12.5μg/ml chloramphenicol. The sequence content of newly obtained BAC clones was verified by comparison of restriction endonuclease-generated fragments of the BACs to expected patterns as determined by the genomic sequence available from the UC Santa Cruz Genome Bioinformatics web site. For restriction endonuclease digestion of BAC DNA, approximately 2 μg of BAC DNA was digested with the appropriate restriction endonuclease(s) and electrophoresed through 0.7% agarose in 1X TAE; fragments were visualized by ethidium bromide staining.
2.3.2 Preparation of BAC DNA

BAC DNA was purified from bacteria grown in 5 ml LB medium overnight at 32°C using standard alkaline lysis plasmid mini-preparation. Wild-type BAC DNA in bacteria was grown in LB-chloramphenicol (12.5 μg/ml). *Hprt* modified BAC DNA in bacteria was grown in LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml). *Hprt* modified + EGFP recombineered BAC DNA in bacteria was grown in LB-chloramphenicol (12.5 μg/ml), -kanamycin (12.5 μg/ml) and LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml) after the Cre/loxP excision of kanamycin. Bacteria cultures were resuspended in 50 mM Tris-HCl, pH 7.5/10 mM EDTA, pH 8 and lysed by addition of an equal volume of 0.2 N NaOH/1% SDS. Proteins were precipitated in 3M K/5M OAc, and BAC DNA was precipitated from the supernatant with isopropanol. BAC DNA pellets were washed with 70% ethanol and resuspended in 40 μl 1X TE, pH 8. To avoid shearing of BAC DNA, samples were transferred between tubes by pouring. Larger-scale preps were prepared using the NucleoBond Plasmid Maxi Kit (BD Bioscience) by following the manufacturer’s protocol for low copy number plasmids.

2.3.3 Preparation of BAC clone glycerol stocks

Unmodified, *Hprt* modified, *Hprt* modified + EGFP recombineered, and *Hprt* modified + EGFP recombineered with excised kanamycin BAC clones were grown in LB-chloramphenicol (12.5 μg/ml), or LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml), or LB-chloramphenicol (12.5 μg/ml), -kanamycin (12.5 μg/ml) or LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml) respectively, overnight at 32°C to a density of approximately 1-2x10⁹ cells/ml. Bacterial cultures were stored at -80°C in an
equal volume of glycerol solution [65% (vol/vol) glycerol, 0.1 M MgSO₄ and 25 mM Tris-HCl, pH 8].

2.3.4 In vitro Cre/loxP modification of BAC DNA

The Hprt BAC-modifying fragment, 13.2 kb pJDH8A/246b, was purified by ethanol precipitation following digestion with NotI. For in vitro Cre/loxP-mediated modification of BAC DNA, 200 ng of the modifying fragment and 200 ng of maxi-preparation BAC DNA were combined in a reaction with Cre recombinase (BD Bioscience/Clontech) following the standard protocol for in vitro recombination from the Creator DNA Cloning Kit (BD Bioscience/Clontech). The recombination reaction was incubated at room temperature for 15 min followed by heat inactivation at 70°C for 10 min. Three microliters of the reaction was electroporated into ElectroMAX DH10B cells (Invitrogen) using a chilled 1 mm cuvette in a BTX Electro Cell Manipulator ECM 600 (BTX, San Diego, CA) set at 1.2 KV, 50 μF, and 129 ohms. The electroporated cells were immediately transferred to 1 ml of room temperature Super Optimal broth with Catabolite repression (SOC) media and incubated with shaking, for 1 hr at 32°C. Cells were plated on LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml) and incubated 48 hours at 32°C. Restriction endonuclease digestion patterns of BAC DNA were used to screen dual chloramphenicol-, ampicillin-resistant colonies for properly modified BACs.

2.3.5 Construction of recombineering vectors

Recombineering vectors were generated by ligating 1) a 2.9 kb SalI/EcoRI fragment of pBSII(+)SK (Stratagene); 2) a EcoRV/EcoRI fragment from a PCR
amplified gene-specific 5’ region of homology (Homology Arm 1); 3) a 1.6 kb EcoRI/XhoI fragment from pIRES2-EGFP (BD Biosciences Clontech) which had been modified with an AflII-XhoI-AflII linker inserted at the AflII site (the fragment contained an IRES-EGFP-SV40 polyA tail construct); 4) a 1.4 kb XhoI/PstI fragment of pOGGFP (containing kanamycin flanked by frt sites (119); 5) a PstI/SalI fragment from a PCR amplified gene-specific 3’ region of homology (Homology Arm 2). The gene-specific fragments were generated by PCR amplification from the respective BAC. (See Table 2.3 for the primers used to generate homology arms.) After incubation at 95°C for 5 minutes, samples were amplified for 30 cycles: 30 sec at 95°C, 30 sec at 58°C and 1 min at 72°C. The recombineering construct fragment, (a ~3.5 kb EcoRV/BssIII fragment of pSox9-EGFP and pOsx-EGFP, and a 4.1 kb EcoRV/PvuI fragment of pTwist2-EGFP), containing the gene specific homology arms flaking the IRES-EGFP-SV40 polyA-tail construct and the frt-flanked kanamycin selection gene, was gel purified with the Sephaglas BandPrep Kit (Amersham Pharmacia).

2.3.6 Recombineering of BAC DNA

Modification of BACs with the Red bacteriophage recombination genes has been described previously (178). Briefly, 200 ng of Hprt modified BAC DNA was electroporated (1.2 KV, 50 μF, and 129 ohms in a 1 mm cuvette at 4°C) into the DH10B-derived E. coli strain EL250. This bacterial strain contains a defective lambda prophage that expresses the Red recombination genes under the regulation of the arabinose-inducible pBAD promoter. The integrity of the transferred BAC was verified by restriction endonuclease digest and agarose gel electrophoresis through a 0.7% gel.
EL250 bacteria containing Hprt modified BAC DNA were grown overnight at 32°C in 5 ml of LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml). The overnight culture was diluted 1:50 in 50 ml of LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml) and allowed to grow at 32°C until reaching an OD<sub>600nm</sub> reading of 0.6. Half of the 50 ml LB culture was transferred to a new flask, and the Red recombination genes were induced by placing the culture at 42°C for 15 min in a shaking water bath. To make the cells electrocompetent, the culture was allowed to cool in a ice slurry for 15 min, pelleted, washed, and resuspended in 250 μl of ice-cold 10% glycerol (in sterile water). Red-induced EL250 bacteria containing Hprt modified BAC DNA were electroporated (as above) with 50 to 100 ng of the EGFP recombineering fragment and plated on LB-chloramphenicol (12.5 μg/ml), -kanamycin (12.5 μg/ml) at 32°C for 24 hrs to select for recombination of the EGFP fragment into the BAC. Dual chloramphenicol-, kanamycin-resistant colonies were screened by restriction endonuclease digestion for proper recombination of the EGFP fragment into the BAC.

An EL250 colony containing a properly modified BAC was grown overnight at 32°C in 5 ml of LB-chloramphenicol (12.5 μg/ml), -kanamycin (12.5 μg/ml). The overnight culture was diluted 1:50 in 50 ml LB-chloramphenicol (12.5 μg/ml), -kanamycin (12.5 μg/ml) containing 0.1% L-arabinose to induce Flpe expression, and the culture was allowed to grow at 32°C until reaching an OD<sub>600nm</sub> reading of 0.6. The arabinose-induced culture was diluted 1:10 in 50 ml of LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml) and allowed to grow at 32°C for an additional hour. Cells were plated on LB-chloramphenicol (12.5 μg/ml), -ampicillin (12.5 μg/ml) at 32°C for 24 hrs. Colonies were randomly picked and replated on LB-chloramphenicol (12.5 μg/ml), -
kanamycin (12.5 μg/ml) to identify clones that had excised the kanamycin resistance cassette. Restriction endonuclease digestion and agarose gel electrophoresis through a 0.7% gel were used to verify removal of kanamycin from the BAC.

2.3.7 Construction of Nestin-EGFP plasmids

The Hprt-Nestin-EGFP plasmid was generated by ligating 1) an 8.7 kb SmaI fragment from pNestin-GFP (containing a 5.8 kb fragment of the Rat Nestin promoter, EGFP cDNA, SV40 polyA tail, and 1.8 kb fragment of the Rat Nestin second intron (a gift from Grigori Enikolopov, Ph.D., Cold Spring Harbor Laboratory); and 2) NotI linearized pMP8NEBΔLacZ (containing 5’ and 3’ homology to the Hprt locus, as well as Hprt complementary sequences)(143). A second Nestin-EGFP-Neo construct was made for random integration into ESCs. It was generated by ligating 1) the 8.7 kb SmaI fragment from pNestin-GFP; and 2) HpaI linearized pOSdupdel (containing a loxP-flanked NeoR marker)(a gift from Oliver Smithies, The University of North Carolina at Chapel Hill). A Pmel/NheI digest isolated a 9.8 kb Nestin-EGFP-Neo fragment for electroporation.

2.3.8 Transgene targeting in ESCs

Modified BAC DNA was purified with BD Bioscience’s Nucleobond Plasmid Maxi Kit, resuspended in sterile 1X TE, pH 8 and linearized with I-SceI. Sterile 10X PBS was added to yield a final concentration of 1-3 nM linearized BAC DNA in a total volume of 400 μl. Polyamine buffer (1X) was added to the final solution (0.4 μl of 1000X 30 mM spermine/70 mM spermidine stock). Modified Hprt-Nestin-EGFP
plasmid construct was linearized with *Pmel* and purified by ethanol precipitation. ESC medium was used to bring the final concentration of DNA to 3 nM in a total volume of 400 μl. ESCs (1x10⁷ to 4x10⁷ cells total) were electroporated with the DNA/PBS or DNA/medium solution in a 2 mm cuvette using a BTX Electro Cell Manipulator ECM 600 set at 270 V, 50 μF, and 360 ohms. Immediately following the electroporation, cells were transferred to ESC medium. For electroporation and cell culture conditions for the randomly integrated Nestin-EGFP ESC see section 2.2.2.

2.3.9 HAT selection

ESCs were grown on murine embryonic fibroblasts (feeders), with or without functional *Hprt*, in DMEM-H (Life Technologies/Invitrogen) supplemented with 15% fetal bovine serum (FBS) (Atlanta Biologicals), 0.1 mM 2-mercaptoethanol (Sigma), 2 mM Glutamax (Invitrogen), and LIF conditioned supernatants (~ 1000 U/ml). Following a targeting electroporation, cells were plated on four 10 cm plates with *Hprt*+ feeders for 24-48 hours in normal ESC medium. Homologous recombinants were then selected for in ESC medium supplemented with HAT (0.016 mg/ml hypoxanthine, 0.01 mM aminopterin, and 0.0048 mg/ml of thymidine). After 10 to 14 days in HAT medium, individual colonies were picked for expansion and verification of the desired recombination event.

2.3.10 Pulse Field Gel Electrophoresis of BAC DNA

Approximately 1 μg of I-SceI-linearized BAC DNA was electrophoresed through 1% agarose in 0.5X TBE for 24 hr at 200 volts with an initial switching time of 1 sec, a
final switching time of 25 sec and a switching ratio of 1. Fragments were visualized by ethidium bromide staining.

2.3.11 Flow Cytometry

ESCs, EBs and Osteogenic cultures were dissociated by trypsinization into single cell suspensions, washed, and resuspended in 1X PBS into no less than 0.5 - 1.0 X 10^6 cells per ml. Cells were analyzed on a FACScan flow cytometer (Becton Dickson) with an argon ion laser excitation setting of 488 nm. Forward- and side-scatter light distribution gates were used to exclude debris, clumps, and dead cells. EGFP fluorescence was read through a 530 nm bandpass filter. For the dually expressing Twist2-ESC line, a histogram using forward-scatter light vs. Fl-1 fluorescence of the CAGGS-EYFP ESC culture was set as background and gates were set to include cells with a shift in EGFP fluorescence greater than background, 0.7% compensation was used. A minimum of ten thousand events was taken from each sample. Data was analyzed using FlowJo flow cytometry analysis software (TreeStar Inc.).

2.3.12 Fluorescent Activated Cell Sorting (FACS) of EBs and Osteogenic Cultures

EBs and Osteogenic cultures were dissociated by trypsinization into single cell suspensions, washed, and resuspended in 1X PBS at no less than 5 X 10^6 cells per ml. Cells were analyzed on a MoFlo High Performance Cell Sorter (Cytomation) with an incident beam of 488 nm; laser output powers were set at 20 mW. EGFP fluorescence was read through a 530/40 nm bandpass filter. Forward- and side-scatter light distribution gates were used to exclude debris, clumps, and dead cells. For the dual
EYFP/EGFP Twist2-ESC line, a histogram using forward-scatter light vs. FL-1 fluorescence of the CAGGS-EYFP ESC culture was set as background, and gates were set to include cells with a shift in EGFP fluorescence greater than background. Gates were set such that approximately 1% or fewer of the single labeled EYFP cells fell within the positive collection gate. Sorted cells were collected in osteoprogenitor medium supplemented with antibiotics. Sort data was analyzed using FlowJo flow cytometry analysis software (TreeStar Inc.). Immediately post sorting, an aliquot of the sort population was stained with 0.4% Trypan blue solution (Sigma) to exclude dead cells, and cell count was determined using a hemocytometer. Cells were pelleted by centrifugation and resuspended in osteoprogenitor medium with antibiotics. A typical primary osteoculture will yield approximately 20-40 colonies when seeded with $2 \times 10^3$ cells/10 cm dish. However upon FACS and replating, a replating density of $1 \times 10^4$ cells/10 cm dish was necessary to obtain sufficient colony formation. Thus, cells were replated for the osteogenic culture at $1 \times 10^4$ cells/10 cm dish, except in Twist2-EGFP density studies where cells were replated at higher densities ($1 \times 10^5$ cells/10 cm dish). Osteogenic differentiation of sorted and replated cells is the same as described in section 2.1.2. The control unsorted population in sort experiments did not undergo the sort procedure but is otherwise harvested and manipulated similarly. The parent control osteoculture was the unmanipulated osteoculture from which the sort culture was originally derived.
**Construction of Reporter**

Obtain a BAC clone containing a centrally located gene of interest from a public repository.

- Introduce the BAC *Hprt* modifying fragment (pJDH8A/246b) in an *in vitro* Cre/loxP reaction. This fragment contains regions of homology to the *Hprt* locus as well as sequences to complement the deleted locus.

- Create the EGFP Recombineering Vector containing an IRES-EGFP-pA cassette and a Kanamycin resistance marker, all flanked by two arms of homology to the gene of interest.

- Add the EGFP reporter sequences to the BAC via homologous recombination within *E. coli* (recombineering).

- Excision of the Kanamycin resistance marker from the BAC via L-arabinose induced expression of flpe, for a flpe/flt mediated excision of the kanamycin resistance maker.

- Targeting of the Hprt-EGFP-BAC to ESCs via electroporation of the linearized Hprt-EGFP-BAC. Selection in HAT medium allows for identification of correctly targeted events.

**Characterization of Reporter**

- Confirm integration of BAC by PCR, RT-PCR

- Assess EGFP fluorescence of cells from osteogenic cultures by Flow Cytometry

- FACS to enrich for GFP* population

- Assess gene expression in sorted populations by QRT-PCR

- Assess osteogenic colony formation of sorted populations

- Assess bone formation *in vivo*
  Transplant EGFP* and EGFP* sort populations to histocompatible mice and assess bone formation within osteogenic scaffolds as well as engraftment potential and homing capability of the cells.

**Figure 2.1:** Flow chart of construction and characterization of BAC reporters.
<table>
<thead>
<tr>
<th>Primer Pair</th>
<th>Primer 1 (5' → 3')</th>
<th>Primer 2 (5' → 3')</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAGGS-EYFP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EYFP</td>
<td>ACGTAAACGGCCACAAGTTTC</td>
<td>AAGTCGCTGCTGCTTCATGTG</td>
<td>188</td>
</tr>
<tr>
<td><strong>Twist2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' Hprt</td>
<td>TGGTTTTTCGAATCTTCCTGTAT</td>
<td>ATCCACAAACATTTTGCGCAC</td>
<td>269</td>
</tr>
<tr>
<td>3' Hprt</td>
<td>TTGCAGATTAGCGATGATGAACC</td>
<td>CTTACTGTTTTTGCAAGAGGACC</td>
<td>782</td>
</tr>
<tr>
<td><strong>Sox9</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' Hprt</td>
<td>TGGTTTTTCGAATCTTCCTGTAT</td>
<td>ATCCACAAACATTTTGCGCAC</td>
<td>269</td>
</tr>
<tr>
<td>3' Hprt</td>
<td>TTGCAGATTAGCGATGATGAACC</td>
<td>CTTACTGTTTTTGCAAGAGGACC</td>
<td>782</td>
</tr>
<tr>
<td><strong>Sox9/IRES (2)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFP</td>
<td>ACGTAAACGGCCACAAGTTTC</td>
<td>AAGTCGCTGCTGCTTCATGTG</td>
<td>188</td>
</tr>
<tr>
<td><strong>Osterix</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' Hprt</td>
<td>TGGTTTTTCGAATCTTCCTGTAT</td>
<td>ATCCACAAACATTTTGCGCAC</td>
<td>269</td>
</tr>
<tr>
<td>3' Hprt</td>
<td>TTGCAGATTAGCGATGATGAACC</td>
<td>CTTACTGTTTTTGCAAGAGGACC</td>
<td>782</td>
</tr>
<tr>
<td><strong>Nestin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' Nes promoter</td>
<td>GCACTCAGTTCTCCTGAAAGG</td>
<td>ACGTGCGCCCATGCTTTTAG</td>
<td>238</td>
</tr>
<tr>
<td>Mid Nes promoter</td>
<td>TCATGAGATTTCCACACCAAC</td>
<td>TGAGGCTTATAGGGTTGAA</td>
<td>399</td>
</tr>
<tr>
<td>3' Nes promoter</td>
<td>CTACTTGTCTGAAAGAAGTT</td>
<td>ACGTGCGCCCATGCTTTTAG</td>
<td>430</td>
</tr>
<tr>
<td>Nestin 2nd intron</td>
<td>TGGAAACCGGCGGTGTGAG</td>
<td>CACATGCAAGATTCCACAGT</td>
<td>336</td>
</tr>
<tr>
<td>3' Nestin 2nd intron</td>
<td>TGGAAACCGGCGGTGTGAG</td>
<td>ACGTGCGCCCATGCTTTTAG</td>
<td>630</td>
</tr>
<tr>
<td><strong>EGFP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Alu</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFP</td>
<td>ACGTAAACGGCCACAAGTTTC</td>
<td>AAGTCGCTGCTGCTTCATGTG</td>
<td>188</td>
</tr>
<tr>
<td><strong>Nestin RT-PCR</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hprt</td>
<td>CTCCCTCCTGACGATCAGGC</td>
<td>CATCGCTAATCAGGACGC</td>
<td>71</td>
</tr>
<tr>
<td>EGFP</td>
<td>ACGTAAACGGCCACAAGTTTC</td>
<td>AAGTCGCTGCTGCTTCATGTG</td>
<td>188</td>
</tr>
<tr>
<td>Gene</td>
<td>RT-PCR Primer 1 (5’ → 3’)</td>
<td>RT-PCR Primer 2 (5’ → 3’)</td>
<td>QRT-PCR Primer 1 (5’ → 3’)</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Osx short isoform</td>
<td>CAGGTTCCCCCAGGAGGA*</td>
<td>AAGTCTGCTGCTTCATGTG</td>
<td>Twist2</td>
</tr>
<tr>
<td>Osx long isoform</td>
<td>TCCTCCCTGCTTTGAGGAGGA*</td>
<td>AAGTCTGCTGCTTCATGTG</td>
<td>Twist2</td>
</tr>
<tr>
<td>EGFP</td>
<td>ACGTAAACGGCCACAAGTTTC</td>
<td>AAGTCTGCTGCTTCATGTG</td>
<td>Twist2</td>
</tr>
<tr>
<td>Osx</td>
<td></td>
<td></td>
<td>Twist2</td>
</tr>
<tr>
<td>Runx2</td>
<td>GGAACCAAGAAGGCAAGACA</td>
<td>AATGCACCTAAATCAGTCAG</td>
<td>Twist2</td>
</tr>
<tr>
<td>Sox9</td>
<td></td>
<td></td>
<td>Twist2</td>
</tr>
<tr>
<td>Bry</td>
<td></td>
<td></td>
<td>Twist2</td>
</tr>
<tr>
<td>Nestin</td>
<td></td>
<td></td>
<td>Twist2</td>
</tr>
<tr>
<td>ColIa1</td>
<td>CCGGAAGAAATAGCTATCACCC</td>
<td>ACCAGGAAAGCAGAGAGTC</td>
<td>Myx11</td>
</tr>
<tr>
<td>Ocn</td>
<td></td>
<td></td>
<td>Myx11</td>
</tr>
<tr>
<td>Myx11</td>
<td>AGGTGGAGGACATGGCAAGAG</td>
<td>GAGTAGGTATAGTAGGAGGCC</td>
<td>CD45</td>
</tr>
<tr>
<td>Lpl</td>
<td>AGGGCTCTTGCCCTGAGTTGTA</td>
<td>AATGGGAAACCTTGGTAGGGGC</td>
<td>CD45</td>
</tr>
<tr>
<td>Ocn</td>
<td></td>
<td></td>
<td>CD45</td>
</tr>
<tr>
<td>Myx11</td>
<td>AGGTGGAGGACATGGCAAGAG</td>
<td>GAGTAGGTATAGTAGGAGGCC</td>
<td>CD45</td>
</tr>
<tr>
<td>Col10a1</td>
<td>CATAAAAGGCCACCTTGCTA</td>
<td>CCTCTTACTGGAATCCCTTTACTC</td>
<td>CD45</td>
</tr>
<tr>
<td>Rpl7</td>
<td>CCCTGAAAGACACTTCCGAAAGG</td>
<td>GCCCTTCTGCCTGACTACAGC</td>
<td>CD45</td>
</tr>
</tbody>
</table>

* Primer sequences from Milona et al. (179)
<table>
<thead>
<tr>
<th>Primer Pair</th>
<th>Primer 1 (5' → 3')</th>
<th>Primer 2 (5' → 3')</th>
<th>Size bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twist2 5’ Homology Arm 1</td>
<td>AGTATGATATCGCCTCGAAATCACAGGCTTT</td>
<td>AGTATGATATCCCTTCCACTCCAGCTCTTCTTCTACTG</td>
<td>184</td>
</tr>
<tr>
<td>Twist2 5’ Homology Arm 1 Larger</td>
<td>ATGATATCAGGATCTCCAACCTTCATA</td>
<td>AGTATGATATCCCTTCCACTCCAGCTCTTCTTCTACTG</td>
<td>855</td>
</tr>
<tr>
<td>Twist2 3’ Homology Arm 2</td>
<td>AGTATCTCTGAGGCTCCCTTCCTGACACAGCTC</td>
<td>AGTATCTGTGCGACATTTGACCTCTCGTCGCTCT</td>
<td>100</td>
</tr>
<tr>
<td>Sox9 5’ Homology Arm 1</td>
<td>GTATGATATCCGAGCCTGCTCCGATCT</td>
<td>TATGATATCCGACCTCTGCTCGTCGCTCA</td>
<td>184</td>
</tr>
<tr>
<td>Sox9 3’ Homology Arm 2</td>
<td>AGTATCTCTGAGGCGGCTGAGTACAAGAAGGA</td>
<td>AGTATCTGTGAGGCTGACCTCTGCTCGTGCTCT</td>
<td>174</td>
</tr>
<tr>
<td>Osx 5’ Homology Arm 1</td>
<td>GTATGATATCGAAGTGCTGTGACTGCTCCAG</td>
<td>TATGATATCCGAGGCTGAAAGGTCAC</td>
<td>168</td>
</tr>
<tr>
<td>Osx 3’ Homology Arm 2</td>
<td>AGTATCTCTGAGGCTCTTGGCAGCAGATGCTCTA</td>
<td>AGTATCTGTGACGAAGGCTGCTCTAGCATACACCTT</td>
<td>280</td>
</tr>
</tbody>
</table>
CHAPTER 3
CREATION OF AN ESC LINE CONTAINING A CONSTITUTIVE
EYFP REPORTER

3.1 Introduction

The bioluminescent jellyfish *Aequorea victoria* naturally fluoresces when calcium binds the photoprotein aequorin, subsequently exciting Green fluorescent protein (GFP) to produce a green light (180;181). In the early 1990’s the cloning and expression of GFP in *Escherichia coli* and *Caenorhabditis elegans* opened the door for the use of GFP as a novel transcriptional reporter (182-184). GFP expressed in eukaryotic or prokaryotic cells can be excited by blue or UV light, and requires no additional substrates or gene products for the production of the fluorescence (182). Native GFP yields a maximum emission at 395 nm excitation, however early FACS instruments used a 488 nm argon laser, resulting in sub-optimal excitation of GFP. Therefore the identification of GFP mutants with an optimal excitation at 488 nm was of great value. Cormack et al. introduced random amino acid substitutions flanking the chromophore sequence to produce GFPmut1, also known as EGFP (185). EGFP has an excitation maximum at 488 nm and has a fluorescence intensity 35 fold higher than native GFP. Native GFP has been further optimized to have altered excitation and emission spectral profiles, creating reporters such as Enhanced Yellow Fluorescent Protein (EYFP) and Enhanced Cyan Fluorescent Protein (ECFP) (186;187). Distinct color variants of GFP allow for the expression and tracking of multiple reporters in a single cell using fluorescence resonance
energy transfer (FRET) or multi-colored flow cytometry. Constitutive expression of EYFP has been shown to be non-toxic to ESCs, the cells were used to successfully generate chimeras, and the resultant chimeras allowed germline transmission of the transgenic genome. The EYFP mouse line was further bred to homozygosity without obvious negative impact, suggesting that EYFP expression was developmentally neutral (188).

To create constitutive EYFP expressing ESCs and mouse lines, EYFP cDNA was placed under control of a strong chimeric cytomegalovirus (CMV) enhancer/chicken beta-actin promoter (AG) (188). pCAGGS contains the human CMV enhancer, AG promoter, beta-actin intron, rabbit beta-globin poly-adenylation signal, and an SV40 origin of replication (189). The level of expression of randomly integrated genes depends on factors such as copy number, site of integration, and strength of promoter; therefore when creating expression vectors, strong promoters are often utilized. In a comparison of different promoters, all linked to a gene for β-galactosidase (LacZ) which permitted the incorporation of a quantitative reporter assay, the AG promoter was shown to be more active than the Rous sarcoma virus (RSV) long terminal repeat, and the CMV enhancer had higher activity than both the AG and the RSV constructs. However, the highest level of β-galactosidase activity was achieved by combining the CMV enhancer with the AG promoter, suggesting that the CMV enhancer and the beta-actin enhancer additively stimulate transcription from the AG promoter (189;190).

This strong enhancer/promoter construct allows for ubiquitous and constitutive expression of GFP variants in both ESCs and mouse models. Early reports with pCAGGS-EGFP mouse models resulted in uniform GFP expression with the exception of
hair and red blood cells. GFP was also expressed throughout all life stages, from pre-implantation embryo to adult. All the transgenic mouse lines were normal and healthy suggesting that the high level of EGFP expression was non-toxic (191). Similar results have been reported for pCAGGS-EYFP and pCAGGS-ECFP (188).

Constitutively expressing GFP cells can be useful for transplantation studies. GFP expressing bone marrow cells or ESC-derived cells can be identified at all times, regardless of cell fate during differentiation and integration. Arnhold et al. used this strategy to monitor transplanted ESC-derived neural precursor cells (192). CAGGS-EGFP ESCs were differentiated in a neural culture, staged according to immunocytochemical staining of parallel cultures for Nestin, and implanted into the striatum of Wistar rats. Due to GFP expression, the grafts were clearly visible in brain sections, and spreading of transplanted cells from the center of the graft was observed over time. Arnhold’s methods relied on immunocytochemical staining to identify the neural precursor cells for transplantation. As the neural culture conditions produced a very pure and homogenous population, the entire culture could then be harvested for transplantation. However, this is not feasible in culture systems where a mixed population exists. In these situations, though the cells can be tracked once implanted, the starting population would be heterogeneous, confounding the determination of which cell type arrived and engrafted in a particular locale, and negative outcomes arising from an undesired cell could complicate interpretation. Therefore it is highly valuable to engineer dually expressing stem cells that allow for enrichment of a specific population based on a gene-specific fluorescent reporter, as well as continuous tracking of the transplanted cells based on a constitutively active different fluorescent reporter. This chapter documents
the construction of CAGGS-EYFP HPRT ESC lines which can serve as host to an array of gene-specific osteogenic reporters. This host ESC line contains a deleted \textit{Hprt} locus which allows for single-copy targeting of the gene-specific reporters with direct forward selection. Importantly, the random integration of CAGGS-EYFP in this host line allows for constitutive expression of EYFP and continual tracking upon transplantation.
3.2 Results

3.2.1 Genomic Analysis of the Constitutive EYFP Reporter in ESCs

The CAGGS-EYFP-PNTlox2 construct contains a CMV enhancer and the chicken *beta-actin* promoter upstream of EYFP allowing for strong and constitutive expression of EYFP. The construct also contains the NeoR gene flanked by 2 loxP sites, which confers resistance to the drug G418. Because the NeoR gene is flanked by loxP sites, after selection of integrated clones, the NeoR gene can be removed to allow the cell to contain as little foreign DNA sequence as possible (Figure 3.1A). Figure 3.1B displays an agarose gel of the CAGGS-EYFP-PNTlox2 construct DNA digested with the restriction endonucleases *BamHI* and *SnaBI*, verifying the proper cloning of the construct. A 5.7 kb fragment obtained from a *SmaI/NotI* digestion was then isolated, gel purified by phenol-chloroform extraction, and electroporated into the BK4 ESC line. The electroporation produced an average of 45 colonies per 10 cm plate, using a total of 10 plates, yielding a frequency of one G418 resistant cell per $6.7 \times 10^4$ cells electroporated. From this, a total of 48 colonies were further expanded for analysis.

A subset of the expanded G418-resistant ESC clones (the first 11 mature clones) were initially tested for the presence of EYFP by PCR (Figure 3.2A). All clones contained the EYFP sequences and produced a PCR amplification product of the correct base pair size. To further test the integrity of the construct, RNA was isolated from these clones to assess for correct mRNA splicing of the transcript. Because the CAGGS-EYFP-PNTlox2 construct contains the chicken *beta-actin* intron, primers were designed to amplify over the intronic sequence, producing a smaller product if mRNA splicing
**Figure 3.1**: pCAGGS-EYFP-PNTloxp2 construct. (A.) Diagram of 11 kb vector containing the CMV enhancer, chicken beta-actin promoter, intron, ~100 bp remaining Cre sequence, EYFP cassette and pA sequence, rabbit beta-globin pA signal and the loxP flanked Neo selection cassette. Large black triangles represent loxP sites. NotI/Smal restriction enzyme cut sites shown to isolate the 5.7 kb fragment for electroporation into ESCs. (B.) Restriction enzyme digestion with BamHI and (C.) SnaBI representing the properly cloned vector, visualized on a 1% agarose gel. Expected band sizes are listed.
Figure 3.2: Integration of pCAGGS-EYFP into ESC clones. (A.) PCR analysis of a subset of G418-resistant ESC clones for the presence of EYFP visualized on a 2% agarose gel. The expected product size of 188 bp is marked by the arrow. Negative control (-) is BK4 ESC DNA. Positive control (P) is pCAGGS-EYFP-PNTlox2 plasmid. (B.) A schematic of a portion of the pCAGGS-EYFP construct with arrows representing primers designed to span the chicken beta-actin intron in order to make an RNA specific assay. (C.) RT-PCR analysis of CAGGS-EYFP in a subset of ESC clones visualized on a 2% agarose gel. The expected product size in the presence of mRNA splicing is 465 bp.
occurred and a large product if splicing had not occurred (Figure 3.2B). RT-PCR amplification of mRNA from the 11 clones demonstrated that the CAGGS-EYFP-PNTlox2 construct resulted in correct splicing to remove the chicken beta-actin intron (Figure 3.2C).

EYFP mRNA expression levels were also quantified in the original subset of 11 clones by QRT-PCR. As described earlier, randomly integrated transgenes can have variability in expression based on position and copy number. Figure 3.3 demonstrates the differences in EYFP expression within the clones. Expression varies up to 4 fold in clones where expression is detectable.

In the creation of the CAGGS-EYFP construct, the original CAGGS-Cre plasmid was modified to excise the Cre and replace it with an EYFP cassette. Due to the availability of restriction endonuclease recognition sites, the removal of the Cre sequence left behind approximately 100 bp, including an ATG translational start site. The EYFP cassette was placed downstream of the remaining Cre sequence. Therefore, the final construct was sequenced to verify that if the upstream ATG in the Cre sequence was utilized as a translational start, the ATG of the GFP sequence would be in the correct frame (Figure 3.4).

These data demonstrate the integration of the CAGGS-EYFP construct in ESC clones, correct splicing and expression of an mRNA product, and sequence that will yield a correct reading frame for two different translational start sites. However, a suitable clone to serve as the host for an array of gene-specific reporters must fulfill certain qualitative criteria. To serve as a tracking reporter upon transplantation to a host model, the clone must have ubiquitous EYFP expression within all cells as well as constitutive
Figure 3.3: QRT-PCR analysis of CAGGS-EYFP ESC clones. Each assay was run in duplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7 (Rpl7), displayed relative to clone 18.
Figure 3.4: Sequence of Cre/EYFP junction in pCAGGS-EYFP-PNTlox2. To create pCAGGS-EYFP, pCAGGS-Cre was altered to remove the Cre sequence (remaining Cre sequence shown in gray text) and introduce an EYFP cassette (EYFP sequence shown in black text). However, ~100 bp of the Cre sequence was left behind on the 5’ end. EYFP was placed downstream of this remaining Cre sequence. The EYFP translational start is 22 amino acids downstream of the Cre start site and in frame. Translational starts sites are indicated by underlining.
EYFP expression throughout differentiation. The euploid clone must also demonstrate the ability to undergo osteogenic differentiation.

3.2.2 Functional Analysis of the EYFP Constitutive Reporter in ESCs

ESCs if maintained properly, in contrast to many other cell lines, display remarkable chromosome stability. However it is well appreciated that changes in chromosome number compromise not only chimera formation and germline transmission of the ESC genome but also in vitro differentiation. Thus one of the first qualitative analyses of the EYFP clones determined the number of chromosomes present. Metaphase counts of 32 clones revealed 12 euploid clones (40 chromosomes) (Table 3.1). These 12 clones became the focus of further analysis and aneuploid clones were excluded.

The uniformity of EYFP expression of the 12 euploid clones was evaluated by fluorescent microscopy. Because the EYFP reporter line will be used to track transgenic cells upon transplantation into a host, ubiquitous expression of EYFP is necessary. For the visualization by fluorescent microscopy, ESC cultures of the clones were split onto gelatinized 6-well plates. Some of the mouse fibroblastic feeder layer cells of the original ESC cultures were also transferred along into the 6-well plate and can be seen in the photographs. These feeder cells are negative for EYFP and can be identified by their characteristic fibroblastic morphology. Fluorescent microscopy revealed a large variation in uniformity of expression of the clones, ranging from no expression, partial expression, to full expression (Figure 3.5). Based on fluorescent microscopy, five clones (clones 23, 34, 35, 36, and 40) were determined to have ubiquitous expression of EYFP.
<table>
<thead>
<tr>
<th></th>
<th>Clones:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euploid</td>
<td>4, 10, 15, 20, 21, 23, 30, 34, 35, 36, 40, 43</td>
</tr>
<tr>
<td>Aneuploid</td>
<td>1, 2, 6, 8, 11, 13, 16, 17, 18, 19, 22, 26, 29, 37, 38, 39, 41, 44, 47, 48</td>
</tr>
</tbody>
</table>

**Table 3.1:** Euploid clones as determined by metaphase spread counts. At least 70% of metaphases analyzed per clone must have a chromosome count of 40 to be considered euploid. Slides were stained with a Giemsa/Wright stain and viewed using a 60X bright field objective to photograph and count metaphase spreads. A minimum of 10 spreads was counted per clone. (Clones 3, 7, 9, 14, 27, 31, 42, 45 were undetermined as they did not survive expansion. Clones 5, 12, 24, 25, 28, 32, 33, 46 were excluded as initial fluorescent microscopy demonstrated non-uniform EYFP expression, therefore eliminating these clones from further evaluation.)
Figure 3.5: Fluorescent and Nomarski (differential interference contrast) microscopy images of CAGGS-EYFP ESCs (32X magnification). Note: in clone 35 non-fluorescent fibroblastic feeders are present.
The osteogenic differentiation potential of these five euploid clones with uniform expression was determined by *in vitro* osteogenic culture. ESCs were allowed to aggregate into EBs for two days, after which the EBs were dispersed to single cell and allowed to differentiate in osteogenic medium for 21 days. At day 21, plates were stained with silver nitrate to detect mineralization and with methyl green as a counter stain. Typically, the osteogenic culture will yield 20-40 colony forming units (CFUs) when seeded with $2 \times 10^3$ cells per 10 cm dish; approximately 60%-80% of these CFUs show significant mineralization, yielding a frequency of CFU-osteogenic of approximately 1 in 100 EB-derived cells (117;119). Figure 3.6 reveals the total colony forming units and percentage of colonies with mineralization from osteocultures of the CAGGS-EYFP ESC clones. While all 5 clones were able to form mineralizing colonies, clone 36 showed the highest percent of mineralized colonies (81%).

The clones were also observed by fluorescent microscopy for constitutive and ubiquitous expression of EYFP throughout the osteogenic culture. Figure 3.7 displays clones 34 and 36 ranging in timepoints from ESC to day 21 of the osteoculture. EYFP expression is maintained throughout the duration of the culture. EYFP is also ubiquitously expressed within all the cells of the differentiating cultures. It is important to note that the layering of cells and matrix can create areas of brighter intensity that, depending on the gain, can reduce the apparent fluorescence of individual cells in less layered areas of the colony. However, upon increased magnification and examination of the cultures, EYFP expression was detected in all cells.

The intent of a ubiquitous tracking reporter is that ESC progeny be visible regardless of the differentiation path or engraftment site. One classic property of ESCs is
Figure 3.6: Osteogenic differentiation of CAGGS-EYFP ESC clones. (A.) Osteogenic cultures stained at day 21 with silver nitrate/methyl green. (B.) Total colony forming units (average CFU/3 plates) and percentage of colonies with mineralization after 21 days in culture. Osteocultures plated at 2x10^3 cells/10 cm plate.
Figure 3.7: Analysis of EYFP expression in CAGGS-EYFP clones throughout the osteogenic culture. Fluorescent and Nomarski (differential interference contrast) microscopy images of CAGGS-EYFP clones 34 and 36 ranging from ESC to Day 21 of the osteoculture. Note: in clone 34 ESC photo non-fluorescent fibroblastic feeders are present.
that they are able to form teratomas when placed \textit{in vivo}. Teratomas are benign growths defined by the broad tissue types represented, including essentially all developmental lineages. Thus, teratoma formation can be used as a quick tool to further test the ubiquitous nature of the EYFP reporter. One million CAGGS-EYFP ESCs or non-transgenic BK4 ESCs were injected bilaterally into the subcutaneous dorsum of 129/P3J mice. After 5 weeks \textit{in vivo}, teratomas were removed and snap frozen for cryosectioning. EYFP teratoma sections observed by fluorescent microscopy showed strong and ubiquitous expression within all cells, while the BK4 teratoma sections demonstrated no fluorescence (Figure 3.8).
Figure 3.8: Expression of EYFP in a teratoma derived from CAGGS-EYFP ESCs. (A.) Section of a five week old teratoma derived from control BK4 ESCs as a negative control for autofluorescence (B.) Expression of EYFP within sections of a parallel teratoma derived from CAGGS-EYFP ESC
3.3 Discussion

This chapter describes the creation and characterization of a constitutive and ubiquitous EYFP reporter whose fluorescence can be observed in ESCs, in ESC-derived osteogenic cultures, and in ESC-derived teratomas. While CAGGS-EGFP/EYFP/ECFP ESC lines have previously been created in other laboratories, it was necessary for this project to introduce the CAGGS-EYFP construct into an ESC line able to receive other genetic modifications via homologous recombination into the *Hprt* locus. This newly established CAGGS-EYFP-HPRT ESC line can now serve as a host to *Hprt* targeted transgenes, specifically the osteogenic reporters described in the following chapter. This approach creates a dual reporter ESC line which includes a tissue/stage reporter for enrichment from heterogeneous differentiation cultures and a ubiquitous reporter for *in vivo* tracking.

Electroporation of the CAGGS-EYFP plasmid into ESCs produced clones which showed integration of the EYFP sequences by PCR, and RT-PCR demonstrated correct splicing of the transgene across the *beta-actin* intron (Figure 3.2). Fluorescent microscopy was the ultimate determinant in establishing the function of the CAGGS-EYFP ESC reporter cell lines. The CAGGS-EYFP reporter construct proved to be functional through fluorescent microscopy by visualization of EYFP, however, while all tested clones demonstrated integration of the EYFP transgene, fluorescent microscopy revealed significant variation in fluorescence of different clones (Figure 3.5). This variability was expected as the CAGGS-EYFP construct is randomly integrated into the
ESC genome where variation in copy number and integration site typically yield significant differences in expression.

The quality of the CAGGS-EYFP ESC clones was also an important issue as at least one of these clones would be chosen to serve as a host to a second fluorescent reporter, and ultimately participate in murine in vivo stem cell therapy models. Of the 32 EYFP ESC clones analyzed, 12 were euploid (Table 3.1), and of these 12 clones, 5 clones were demonstrated to have uniform EYFP expression in ESC cultures (Figure 3.5). These clones maintained the capacity to undergo osteogenic differentiation (Figure 3.6) and EYFP fluorescence was observed throughout the osteogenic culture (Figure 3.7). As a final test of the ubiquitous nature of the CAGGS-EYFP reporter, CAGGS-EYFP ESC derived teratomas were generated and EYFP expression was observed in all cells (Figure 3.8). This rigorous validation process yielded 5 clones, one of which (CAGGS-EYFP ESC clone 36) was chosen to be the host for gene-specific BAC-based reporter constructs, described in the following chapter.
CHAPTER 4
CREATION OF GENE SPECIFIC EGFP ESC REPORTERS

4.1 Introduction

Recombinant DNA methodology facilitated by amplification in *Escherichia coli* is the cornerstone of molecular biology. Plasmid constructs can be created to facilitate genetic manipulations by homologous recombination or random integration of transgenic sequences in a variety of model systems including yeast, worm, fly, mouse, rat, and human cell lines. However, there are limitations of standard cloning technology. The methods used to create these sometimes complicated constructs depend on the availability of unique restriction endonuclease cleavage sites, and plasmid based vectors have a limited size capacity (~20 kb). Recently, the DNA double-strand break and repair recombination machinery in *E. coli* has been implemented in the manipulation of larger segments of DNA within bacterial artificial chromosomes (BACs), a process termed ‘Recombineering.’

The earliest form of recombineering relied on *E. coli* strains which contained a mutant linear-double stranded (ds) DNA exonuclease. RecBCD is a multisubunit enzyme containing helicase and strong linear exonuclease activity (193-195). A *recBC* mutation however, allows for the introduction of linear DNA to the cell without the otherwise rapid degradation which would normally occur (194;196). This system allowed for introduction of linear dsDNA or PCR products into plasmids as well as chromosomal DNA (196-199). However, a disadvantage of these RecBCD mutant strains is that the recombination machinery is constitutively active, which can cause rearrangements and
recombination events resulting in mutation. Also, RecBCD deficient strains grow very poorly.

The addition of Chi sites to linear dsDNA facilitates recombineering in wild-type *E. coli* by protecting the DNA from degradation by RecBCD. Chi sites (5’-GCTGGTGG-3’) normally mark the end of exonuclease activity and the loading of a synapsis protein subsequent to strand exchange with an intact homolog and repair (200;201). The helicase activity of RecBCD is still active, creating a single stranded DNA tail which can then participate in strand exchange (201). This technique however requires large areas of homology and is very inefficient.

Circular targeting constructs can be used in wild-type *E. coli* as RecBCD will not degrade circular DNA, but the integration of circular DNA requires RecA as well as relatively large regions of homology (~500 bp). Many commonly used strains of *E. coli* are recA’, necessitating the introduction of the recA gene within the targeting vector (196;202). The need for more homology and the increased risk of other unwanted recombination events are the disadvantages of this system.

Advances in recombineering came with application of the RecET homologous recombination system, encoded within the RAC prophage of *E. coli* K12 strains, permitting recombination of linear dsDNA with significantly shorter areas of homology (≥ 42 bp). This system was further honed by the creation of the pBAD-ETγ plasmid, which placed the exonuclease (*recE*) under an inducible promoter (pBAD), and the loading protein (*recT*) and the RecBCD exonuclease inhibitor (*gam*) under constitutive promoters (EM7, Tn5), allowing three times as many recombinants as compared to native RecET machinery (203). The bacteriophage λ-encoded Red recombination system was
also introduced which is analogous to the RecET system; under this system the homologous recombination of linear DNA was approximately 50-70 times more efficient than previous systems, however, homologies had to be more than 1,000 bp in size (196;204;205). A plasmid analogous to pBAD-ETγ was also created for this system called pBAD-αβγ (206). Advantages of both of these plasmid systems are that they can be used in any *E. coli* strain, the exonuclease activity of RecBCD is controlled, and recombination occurs only upon the induction of the introduced exonuclease, therefore decreasing the probability of unwanted events due to prolonged exposure to active recombination machinery. The downside however, is that these plasmid based systems are poorly maintained and over-expression of the RecBCD exonuclease inhibitor can be toxic (205).

Finally, one of the most efficient recombineering systems involves using an *E. coli* strain containing a λ prophage with its Red recombination genes under control of a temperature-sensitive λ cl-repressor (194). The represor is active at 32°C, but when cells are shifted to 42°C for 5-15 minutes, the represor is inactivated allowing recombination activity as well as the inhibition of RecBCD exonuclease activity. The genes are under control of the λ-pL promoter, allowing the brief expression to be at a very high level. The short period of recombination activity limits unwanted recombination events. The stably integrated λ-prophage system, versus previous RecET and λ recombination plasmids, yields 50-100 fold more efficient recombination, and homologous regions can be as short as 30-50 bp.

Copeland et al. applied this prophage system to recombine BACs (178). To create a BAC transgenic mouse line with neural specific *Cre* expression, an IRES-
eGFPcre-FRT-kan-FRT cassette was targeted to the *Eno2* locus of a 250 kb BAC. Primers including 42 nt of homology to the *Eno2* locus were created to amplify the targeting cassette, placing the cassette downstream of the *Eno2* stop codon and upstream of its polyA site. Of $10^8$ electroporated cells, approximately $5 \times 10^3$ kanamycin resistant colonies were isolated. Because the recombineering occurred in the EL250 strain of *E. coli* which contains an arabinose-inducible *flpe* gene, the kanamycin resistance marker was then removed by inducing Flpe expression. A transgenic mouse line was created by pronuclear injection of this recombineered Eno2 BAC, demonstrating neural specific Cre expression.

Our group has used this same prophage based recombineering system to create an Osteocalcin-GFP reporter BAC (119). The recombineered BAC was modified by introducing an *Hprt* targeting cassette within the BAC vector backbone by Cre/loxP recombination (Figure 4.1) (118;119). This recombineered BAC-based GFP reporter was targeted in single copy to the *Hprt* locus and demonstrated appropriate Osteocalcin-GFP reporter expression in murine ESC-derived osteogenic nodules. As with smaller recombinant constructs, this targeting method eliminates the complications of copy number, position effects, and insertional mutagenesis which can be seen with random integration and also allows a means of direct forward selection in HAT medium of even a rare recombination event. This chapter describes the construction of a series of targeted BAC-based osteogenic reporters to identify and enrich for therapeutic osteogenic progenitors.
Figure 4.1: Hprt modification of BAC clones. BAC modifying fragment pJDH8A/24b (containing human Hprt promoter and exon 1, and mouse exons 2 and 3) is introduced to the BAC vector backbone by Cre/loxP recombination. Subsequent linearization with I-SceI allows the BAC genomic insert to be flanked by 5' Hprt homology as well as 3' Hprt homology/Hprt rescuing sequences. (Figure from (207)).
4.2 Results

4.2.1 BAC Clones

The Twist2-BAC, RP24-291O6, is a mouse BAC from the RPCI-24 library. This BAC is constructed on the pTARBAC1 vector backbone which is 13 kb in size. The genomic DNA sequence within the BAC is 158 kb in size with the Twist2 gene located approximately in the middle (Figure 4.2). The Twist2 gene is flanked by 71 kb of upstream sequence and 41 kb of downstream sequence (Table 4.1A). BACs were chosen with the gene of interest centrally located and flanked by generous up- and downstream sequences in attempt to mimic endogenous expression by maintaining surrounding enhancer and promoter elements. Upon Hprt modification and EGFP recombineering of the BAC, the total size of the Hprt-Twist2-EGFP-BAC was 186 kb (Table 4.1B).

The Sox9-BAC, RP11-141M17, is a human BAC from the RPCI-11 library and is constructed on the pBAC3.6 vector backbone which is 12 kb in size. The genomic DNA sequence is 174 kb and the Sox9 gene is positioned to have maximum upstream flanking sequence. Sox9 is flanked by 130 kb upstream sequence and 38 kb downstream sequence (Table 4.1A). A Sox9-BAC was chosen with a large upstream sequence because Sox9 has been shown to be regulated by upstream distal cis-acting regulatory elements (208-210). Upon Hprt modification and EGFP recombineering, the Hprt-Sox9-EGFP-BAC was a total size of 197 kb (Table 4.1B).

The Osterix-BAC, CTD-3071P8, is a human BAC from the CITB-D library and is constructed on the pBeloBAC11 vector backbone which is 7.5 kb in size. This BAC contains a genomic sequence 157 kb in size with the Osx (Sp7) gene located generally
Figure 4.2: Physical maps of the genomic inserts of BACs to be recombineered to create EGFP reporters. (A.) Mouse BAC clone RP24-291O6 containing the \textit{Twist2} gene. (B.) Human BAC clone RP11-141M17 containing the \textit{Sox9} gene. (C.) Human BAC clone CTD-3071P8 containing the centrally located \textit{SP7/Osterix} gene.
### Table 4.1 A

**BAC Clones**

<table>
<thead>
<tr>
<th>Gene of Interest</th>
<th>BAC clone</th>
<th>Species</th>
<th>Size (kb)</th>
<th>Upstream Sequence (kb)</th>
<th>Downstream Sequence (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twist2</td>
<td>RP24-291O6</td>
<td>Mouse</td>
<td>158</td>
<td>71</td>
<td>41</td>
</tr>
<tr>
<td>Sox9</td>
<td>RP11-141M17</td>
<td>Human</td>
<td>174</td>
<td>130</td>
<td>38</td>
</tr>
<tr>
<td>Osterix</td>
<td>CTD-3071P8</td>
<td>Human</td>
<td>157</td>
<td>57</td>
<td>90</td>
</tr>
</tbody>
</table>

### Table 4.1 B

**BAC Clone Sizes (kb)**

<table>
<thead>
<tr>
<th>BAC</th>
<th>Genomic DNA</th>
<th>BAC Vector</th>
<th>Recombineering Fragment</th>
<th>Hprt Homology</th>
<th>Total Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twist2</td>
<td>158</td>
<td>13</td>
<td>1.5</td>
<td>12</td>
<td>186</td>
</tr>
<tr>
<td>Sox9</td>
<td>174</td>
<td>12</td>
<td>1.5</td>
<td>12</td>
<td>197</td>
</tr>
<tr>
<td>Osterix</td>
<td>157</td>
<td>7.5</td>
<td>1.5</td>
<td>12</td>
<td>178</td>
</tr>
</tbody>
</table>
in the middle. \textit{Osx} is flanked by 57 kb upstream and 90 kb downstream (Table 4.1A). One other BAC was available with \textit{Osx} more centrally located, but the very large 230 kb size of this BAC made it undesirable for our use. After \textit{Hprt} modification and EGFP recombineering the Hprt-Osx-EGFP-BAC was 178 kb in size (Table 4.1B).

4.2.2 Modification of BACs to contain \textit{Hprt} homology

The Twist2-BAC, Sox9-BAC, and Osx-BAC were modified to contain \textit{Hprt} homology with the modifying vector pJDH8A/246B. pJDH8A/246B contains 3.8 kb of 5’ homology to the \textit{Hprt} locus, 4.5 kb of \textit{Hprt} complementary sequences including human \textit{Hprt} exon 1, and 3 kb of 3’ homology to the \textit{Hprt} locus (118). The linearized modifying vector was introduced by an \textit{in vitro} Cre-mediated recombination event at the \textit{loxP} site of the BAC vector backbone (Figure 4.1). After the \textit{in vitro} modification of the BACs, the recombination reaction was electroporated into DH10B cells and recombinants were plated on agar with selection for chloramphenicol-, ampicillin-resistance (48 hrs). The integration of the \textit{Hprt} modifying fragment into the BAC clones alters the restriction digest pattern allowing for the verification of proper modification by restriction endonuclease-generated BAC fragment patterns (Figure 4.3). In the original screening of the Hprt-Twist2-BAC by restriction endonuclease digestion, out of 5 tested colonies, 4 were correctly modified (80%). Of the 16 tested \textit{Hprt} modified Sox9-BAC colonies, 11 were correctly modified (68%). Of the 12 tested \textit{Hprt} modified Osx-BAC colonies, 3 were correctly modified (25%). Previous modifications in the lab had an average of 56% correct integrants. Inappropriate recombination events included colonies which showed an incorrect BAC fragment pattern, or intensely bright bands which
Figure 4.3: Identification of properly modified *Hprt* BACs. (A.) Restriction endonuclease analysis of the unmodified Sox9-BAC (un) as compared to modified Hprt-Sox9 BAC (mod). (B). Restriction endonuclease analysis of the unmodified Osx-BAC (un) as compared to modified Hprt-Osx BAC (mod). Appearance of new bands resulting from *Hprt* modification are identified by arrowheads.
suggested that an unexplained high copy-number plasmid formed during the recombination reaction.

4.2.3 EGFP Recombineering Vectors

All recombineering vectors were designed to have an IRES-EGFP-polyA cassette with a frt-flanked kanamycin selectable marker, flanked by two arms of homology to a specific gene, allowing for homologous recombination of the fragment into the specific locus on the BAC (Figure 4.4). All homology arm amplification primers were designed to contain restriction endonuclease sites to allow for cloning of the PCR fragments into the recombineering constructs. This modification of the BAC by a homologous recombination integration event occurs within a strain of Red recombination competent EL250 E. coli cells. The Red recombineering system allows for the modification of the BACs without the necessity of identifying unique restriction endonuclease sites for cloning of the large construct. The frt-flanked kanamycin marker allows for the selection of modified BACs after the homologous recombination event within the E. coli. Because kanamycin is flanked by frt sites, the kanamycin marker can then be removed through Flpe-mediated recombination before the BAC is introduced to ESCs. The IRES-EGFP cassette is placed on the BAC downstream of the translational start, producing a bicistronic transcript with two translational products: a cap-dependent truncated peptide from the disrupted gene and cap-independent GFP.

The pTwist2-EGFP recombineering fragment was engineered to insert into exon 1 on the Twist2-BAC. This replacement event created a 23 amino acid truncated Twist2 peptide followed by a stop codon, with the IRES-EGFP cassette immediately
Figure 4.4: Modification of BACs to express the EGFP recombineering fragment. A general gene of interest from the BAC is shown with its exons. The EGFP recombineering vector contains short arms of homology to the locus, an IRES-EGFP-pA cassette, and a *frt* flanked kanamycin selectable marker. During recombineering, a replacement homologous recombination event inserts the recombineering fragment into the selected locus. Following identification of properly modified BACs, induction of Flpe-recombinase in bacteria with arabinose promotes the excision of the kanamycin resistance marker. Transcription produces a bicistronic mRNA, from which a truncated peptide is produced by cap-dependent translation and GFP is produced by cap-independent translation initiated at the IRES. (Figure and legend modified from (207))
downstream. Upon original construction of the recombineering fragment, homology arms were created by PCR amplification of the Twist2 locus on the BAC to be approximately 100-200 bp in size (5’ Homology Arm 1 - 184 bp and 3’ Homology Arm 2 - 100 bp). However, initial attempts at recombineering the fragment into the Twist2 locus proved to be unsuccessful. The recombineering fragment was rebuilt with the 5’ Homology Arm 1 increased in size to 855 bp to encourage the homologous recombination event (see Table 2.3 for primers used). The reverse primer used to amplify 5’ Homology Arm 1 was engineered to create a stop codon (UAG) downstream of the Twist2 exon 1 translational start site. The 5-way ligation to create pTwist2-EGFP (including Twist2 5’Homology Arm 1, IRES-EGFP-pA cassette, frt flanked kanamycin, 3’Homology Arm 2, and pBSIISK vector backbone) was verified by restriction endonuclease digestion (Figure 4.5).

The pSox9-EGFP and pOsx-EGFP constructs were made in the same manner as pTwist2-EGFP (Figure 4.4). pSox9-EGFP was designed to recombine into the Sox9-BAC in Sox9 exon 1. The integration of the recombineering fragment downstream of the translational start resulted in a 16 amino acid truncated Sox9 peptide with an engineered stop codon (UGA). The 5’Homology Arm 1 was designed to be 184 bp and the 3’Homology Arm 2 was 174 bp in length.

The homology arms for pOsx-EGFP (5’Homology Arm 1 - 168 bp and 3’Homology Arm 2 - 280 bp) were engineered to place the pOsx-EGFP recombineering fragment in Osterix exon 3 of the Osx-BAC. Osterix has three alternatively spliced isoforms yet all three isoforms contain exon 3; therefore by placing EGFP in exon 3, all isoforms will express the reporter. The insertion of the recombineering fragment, with its
Figure 4.5: Restriction enzyme digests of properly cloned recombineering vectors. (A.) NotI restriction digest of pTwist2-EGFP. (B.) NotI/SalI restriction digest of pSox9-EGFP. (C.) XbaI restriction digest of pOx-EGFP. Restriction enzyme digests visualized on a 1% agarose gel. Expected fragment sizes of each enzyme digest are listed.
engineered stop codon downstream of the translational start site of exon 3, resulted in a 39 amino acid truncated peptide for the short Osterix isoform and a 57 amino acid truncated peptide for the long Osterix isoform.

4.2.4 Recombineering of BACs to contain EGFP

The correctly ligated EGFP-recombineering vectors were then used to modify their respective BAC by homologous recombination within the EL250 bacterial strain. BAC containing bacteria were induced for recombination by incubation at 42°C for 15 minutes, subsequently 50 to 100 ng of the linearized EGFP-recombineering fragment was electroporated into the cells. Dual chloramphenicol-, kanamycin-resistant colonies were screened for by restriction endonuclease digestion for proper recombination of the EGFP fragment into the BAC.

A pTwist2-EGFP recombineering fragment, obtained by an EcoRV/PvuI digest, was used in recombineering the Hprt-Twist2-BAC to contain the EGFP sequence (Figure 4.6). Out of 6 tested colonies, preliminary restriction endonuclease digestion screening demonstrated all with a correct BAC banding pattern. Two correct integrants were chosen for large scale DNA amplification and further verification by endonuclease restriction digestion.

A pSox9-EGFP recombineering fragment was obtained by an EcoRV/BssHII digest. Of the 8 colonies tested from this recombineering attempt, all 8 colonies contained Hprt-Sox9-BAC DNA with correct EGFP- recombineering fragment integration. Four colonies were chosen for further verification by endonuclease restriction digestion (Figure 4.7 A,B).
Figure 4.6: Proper EGFP recombineering of Twist2 BAC. (A.) An EcoRV/PvuI digest of pTwist2-EGFP results in a 4.16 kb recombineering fragment. (B.) Hprt-Twist2-EGFP BAC DNA from kanamycin resistant recombineered colonies was digested with Hpal and visualized on a 0.7% agarose gel. Arrows mark appearance of new bands resulting from integration of EGFP. un= unrecombineered Hprt-Twist2 BAC DNA.
Figure 4.7: Proper EGFP recombineering of Sox9 and Osterix BACs. (A.) Diagram of pSox9-EGFP recombineering vector with marked restriction enzyme sites. (B). Hprt-EGFP-Sox9 BAC DNA digested with BamHI restriction enzyme and visualized on a 0.7% agarose gel. (C.) Diagram of pOsx-EGFP recombineering vector with marked restriction enzyme sites. (B). Hprt-EGFP-Osx BAC DNA digested with XhoI restriction enzyme and visualized on a 0.7% agarose gel. un= unrecombineered Hprt modified BAC DNA.
Lastly, a pOsx-EGFP recombineering fragment, from an EcoRV/BssHII digest, was able to correctly recombineer the Hprt-Osx-BAC to contain the EGFP-recombineering fragment in 7 out of 8 colonies (Figure 4.7 C,D).

Two colonies containing correctly recombineered integrants each for Hprt-Twist2-EGFP-BAC, Hprt-Sox9-EGFP-BAC, and Hprt-Osx-EGFP-BAC were chosen for kanamycin removal. Upon the addition of 0.1% L-arabinose to the bacterial cultures, Flpe induced expression allowed for a Flpe/frt mediated excision of the kanamycin resistance marker. Colonies which excised kanamycin were selected for on LB-chloramphenicol, -ampicillin plates and replicated plated on LB-chloramphenicol, -kanamycin plates to verify lack of kanamycin resistance. At least 8 kanamycin excised colonies were tested for each Hprt-EGFP-BAC construct by restriction endonuclease digestion and nearly 100% of the colonies displayed a correctly excised kanamycin banding pattern. One colony was chosen for each Hprt-EGFP-BAC construct for large scale DNA preparation (Figure 4.8).

4.2.5 Targeting of Hprt-EGFP-BAC constructs in ESCs

Modified Hprt-EGFP-BAC DNA was linearized with I-SceI. Linearization at the I-SceI site placed the EGFP recombineered genomic sequence at the center of the construct, flanked by regions of homology to the Hprt locus (Figure 4.1). Concentrations ranging between 1.3 – 3.0 nM of linearized Hprt-EGFP-BAC DNA were electroporated into ESCs for targeting to the Hprt locus. Electroporated cells were allowed to recover for two days in ESC medium and were switched on the second day for forward selection
Figure 4.8: Proper modification and kanamycin removal of BACs.  (A.) HpaI restriction enzyme digestion of various stages of Twist2-BAC modification.  (B.) HpaI restriction enzyme digestion of various stages of Sox9-BAC modification.  (C.) EcoRI restriction enzyme digestion of various stages of Osx-BAC modification.  Visualization on 0.7% agarose gel.  Unmodified = original BAC, Hprt Mod = Hprt modified BAC, Hprt + GFP= Hprt modified and EGFP recombineered BAC, Hprt + GFP – Kan= Hprt modified and EGFP recombineered BAC with kanamycin removal.  Arrow heads represent new bands resulting from integration of modifying constructs. Shifting arrow representative of kanamycin removal.
in HAT medium. Only cells which have a complemented Hprt locus are able to survive HAT selection.

Hprt-Twist2-EGFP-BAC DNA was electroporated into CAGGS-EYFP-ESC clone 36 in four separate attempts (Table 4.2). Concentrations of electroporated DNA ranged from 1.44 nM to 1.57 nM. From the four electroporations, 7 ESC colonies survived HAT selection and expansion of the clone. One clone, T1, had normal ESC colony morphology of circular shape, sharp edges, appropriate size, and contained a correctly targeted Twist2 BAC as determined by PCR amplification analysis (Figure 4.9A). Primers were designed to test for the ends of the integrated BAC construct at the 5’Hprt region and in the 3’Hprt rescue sequence, as well as transgenic Twist2-EGFP sequences in the middle of the construct (Primer Table 2.1). All of the other clones contained only the 3’Hprt rescuing sequence, allowing them to survive selection, while the Twist2-EGFP and 5’BAC end were absent. Electroporations were also performed using the BK4 ESC line (a non-EYFP containing ESC line). Hprt-Twist2-EGFP-BAC DNA was introduced into BK4 ESCs in 14 electroporations with the concentration of electroporated DNA ranging from 1.44 nM to 3.05 nM. A total of 4 colonies survived selection and expansion for analysis, T17-T19,T22. None of the clones derived from the BK4 electroporations contained a correctly integrated Twist2 BAC.

Hprt-Sox9-EGFP-BAC DNA was also introduced into CAGGS-EYFP-ESC clone 36 (4 electroporations), as well as into BK4 ESCs (14 electroporations), to target the BAC to the Hprt locus (Table 4.2). Electroporation concentrations ranged from 1.34 nM to 3.03 nM. In total, 18 clones survived HAT selection and expansion, however none contained a correctly integrated BAC. Again, integration of the BAC was tested by
Table 4.2  Summary of Electroporations of BAC Reporters into ESCs

<table>
<thead>
<tr>
<th>BAC Reporter</th>
<th>Cell Line</th>
<th># of Electroporations</th>
<th># of HAT Resistant Colonies</th>
<th># of Correctly Targeted Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twist2</td>
<td>CAGGS-EYFP 36</td>
<td>4</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Twist2</td>
<td>BK4</td>
<td>14</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Sox9</td>
<td>CAGGS-EYFP 36</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Sox9</td>
<td>BK4</td>
<td>14</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Osterix</td>
<td>BK4</td>
<td>3</td>
<td>17</td>
<td>5</td>
</tr>
</tbody>
</table>
Figure 4.9: Verification of recombination of BACs into the Hprt locus of HAT-resistant ESC clones. (A.) PCR analysis of clone T1 testing for 5’ and 3’ Hprt sequences as well as central Twist2-EGFP sequences. T1= Twist2 clone T1, - = 2.36 CAGGS EYFP ESC. (B.) PCR analysis of clone S18 testing for 5’ and 3’ Hprt sequences as well as central Sox9 recombineering fragment sequences. S18 = Sox9 clone S18, - = BK4 ESC, + = Hprt-Sox9-EGFP-BAC DNA. (C.) PCR analysis of Osterix clones for 5’ and 3’ Hprt sequences as well as central Osterix recombineering fragment sequences. S18 = Sox9 clone S18, - = BK4 ESC, + = Hprt-Osx-EGFP-BAC DNA. (D.) Alu fingerprinting PCR of Osterix clones. + = Hprt-Osx-EGFP-BAC DNA, - = BK4 ESC, 8 = known incorrect integrant included for comparison. Visualization on a 1.5% agarose gel.
primers designed to amplify the 5’ and 3’ ends of the BAC as well as middle Sox9-EGFP sequences within the BAC (see Primer Table 2.1). All clones contained the 3’Hprt rescuing sequences, yet no clones contained transgenic Sox9-EGFP sequences. One clone, S18, did amplify both the 5’ and 3’ ends of the BAC, on detailed PCR analysis however, no Sox9-EGFP sequences were found to be contained within the clone (Figure 4.9B).

The Hprt-Osx-EGFP-BAC was targeted to the Hprt locus of BK4 ESCs in 3 electroporations (Table 4.2). The three electroporations produced 17 viable clones, 5 of which integrated a correctly targeted BAC (Osx2,5,7,10,12). Integration of the BAC was first assessed by PCR amplification of the 5’ and 3’ BAC ends, transgenic Osx sequences, and EGFP sequences (Figure 4.9C). As the ESCs are murine and the Osx-BACs is human, Alu DNA fingerprinting was used to look at the integrity of the whole BAC. Alu sequences are a series of short, interspersed, repeated DNA elements distributed throughout primate genomes. PCR of oppositely oriented Alu sequences can be used as a method of obtaining a fingerprint, or a unique pattern of bands, for a large genomic region. Alu PCR of the 5 correct Hprt-Osx-EGFP-BAC ESC clones revealed a banding pattern identical to that of the Hprt-Osx-EGFP-BAC positive control (Figure 4.9D). These results suggest the integration of the entire BAC within these clones.

4.2.6 Creation of Nestin-EGFP ESC

To create the Hprt-Nestin-EGFP targeting fragment, pNestin-GFP (211) was modified to have Hprt regions of homology. pNestin-GFP contains a 5.8 kb fragment of the rat Nestin promoter, an EGFP cDNA, an SV40 polyA tail, and a 1.8 kb fragment of
the rat *Nestin* second intron. This Nestin construct was ligated into pMP8NEBΔLacZ (143) which flanked the construct with regions of 5' *Hprt* homology and 3' *Hprt* homology as well as *Hprt* sequences capable of complementing the deletion (Figure 4.10A). Restriction endonuclease digestion was used to verify the correct ligation of the pNestin-GFP fragment into pMP8NEBΔLacZ (Figure 4.10B). Upon verification of the correct construct, Hprt-Nestin-EGFP DNA was targeted to the *Hprt* locus of BK4 ESCs. One 3 nM electroporation of linearized Hprt-Nestin-EGFP DNA resulted in the formation of 17 clones which survived selection and expansion. An initial PCR amplification screen was performed on the Hprt-Nestin-EGFP ESC clones to see if they contained the newly integrated EGFP sequence; EGFP sequence was amplified from all the clones (Figure 4.11A). To further investigate the integration event, a subset of clones was tested by PCR with primers to amplify *Hprt* regions flanking the construct, sequences within the rat *Nestin* promoter, and within the rat *Nestin* 2nd intron (Figure 4.11B,C). Again, all analyzed clones contained the queried sequences. Finally, clones were tested by RT-PCR for the presence of a complemented and functional *Hprt* locus, demonstrating correct integration of the targeted construct (Figure 4.11D).

A second Nestin-EGFP ESC was created by the random integration of Nestin-EGFP into the ESC. The pNestin-GFP fragment was this time ligated into pOSdupdel to include a *loxP*-flanked NeoR selection marker. Restriction endonuclease digestion identified two correct constructs, with forward and backward orientation relative to pOSdupdel (Figure 4.12A,B). A *Pmel*/*Nhel* digestion linearized the constructs as well as removed the vector backbone, yielding a 9.8 kb Nestin-EGFP-Neo fragment for electroporation. A 3 nM electroporation into BK4 ESC was performed with both the
Figure 4.10: Hprt-Nestin-EGFP Targeting Fragment. (A.) Diagram of pHprt-Nestin-EGFP construct containing 5.8 kb fragment of the rat Nestin Promoter, EGFP cDNA-SV40 pA tail, 1.8 kb fragment of the rat Nestin 2nd Intron, all flanked by 5′ Hprt homology as well as 3′Hprt homology and rescuing sequences. (B.) EcoRV restriction endonuclease digestion to verify correct cloning of pHprt-Nestin-EGFP. Visualized on a 1% agarose gel. Vector = pMP8NEBΔLacZ
Figure 4.11: Verification of targeted Hprt-Nestin-EGFP in ESC clones. (A.) PCR screen of a subset of Hprt-Nestin-EGFP ESC clones to verify the integration of EGFP. (B.) PCR verification of a subset of Hprt-Nestin-EGFP ESC clones for the integration of 5’, internal, and 3’ sequences of the rat Nestin Promoter. (C.) PCR verification of a subset of Hprt-Nestin-EGFP ESC clones for the integration of internal and 3’ rat Nestin 2nd Intron sequences. (D.) RT-PCR analysis of Hprt-Nestin-EGFP ESC clones for the presence of a rescued and functional Hprt locus. Visualization on 1% agarose gel. Expected product sizes are listed.
**Figure 4.12:** pNestin-EGFP-Neo random integration construct. (A.) Diagram of the Nestin-EGFP random integration construct in both forward and backward orientation relative to pOSdupdel vector. (B.) Restriction enzyme digestion revealing the correct cloning of pNestin-EGFP-Neo. fwd = forward orientation of Nestin-EGFP in pOSdupdel vector, bkwd = backward orientation of Nestin-EGFP in pOSdupdel vector. (C.) PCR analysis of Nestin-EGFP-Neo ESC clones for the integration of the rat Nestin promoter, rat Nestin $2^{nd}$ intron, and EGFP sequences. F = forward orientation clones, B = backward orientation clones, - = BK4 negative control DNA, + = pNestin-EGFP-Neo plasmid DNA. Expected PCR product sizes are listed. Visualization on a 1% agarose gel.
forward-orientation and backward-orientation constructs. Both electroporations resulted in over 100 Nestin-EGFP ESC colonies surviving G418 selection; 24 colonies were picked to analyze of the forward orientation, and 36 colonies were picked of the backward orientation. A subset of the forward (F) and backward (B) orientation clones were tested by PCR amplification for the presence of the rat *Nestin* promoter, 2nd intron, and EGFP sequences (Figure 4.12C). Of the 10 tested clones, only one, 28B, did not contain sequences for the complete Nestin-EGFP construct.
4.3 Discussion

In this chapter we have modified gene-specific BACs to contain GFP, as well as *Hprt* homology, and subsequently targeted these constructs to the *Hprt* locus within ESCs to serve as potential reporters for osteogenic progenitors. The Hprt-Twist2-EGFP-BAC was electroporated into ESCs in 18 attempts, yielding 1 correctly integrated clone. The Hprt-Sox9-EGFP-BAC was also electroporated into ESCs in 18 attempts, yielded no correctly integrated clones. The Hprt-Osx-EGFP-BAC DNA was electroporated into ESCs in 3 attempts, yielding 5 correctly integrated clones. The large size of the Sox9 and Twist2 BAC constructs is likely responsible for their low targeting efficiency (Twist2-BAC is 186 kb, Sox9-BAC is 197 kb, while Osx-BAC is 178 kb), because BACs have been shown to be highly susceptible to breakage by hydrodynamic shearing, resulting in terminal BAC deletions. The BAC DNA, prior to electroporation, was assessed by pulse field gel electrophoresis (PFGE) and showed no significant degradation; however within the electroporation protocol a BAC DNA and cell resuspension step takes place which could alter DNA integrity. Wide bore pipet tips were always used when handling BAC DNA and caution was taken to handle the DNA gently. BAC DNA is commonly stabilized for microinjection by the addition of polyamines to the microinjection buffer. While the polyamines are not necessary for microinjection, we employed this concept to potentially stabilize our BAC electroporations.

Larger BACs have also been shown to be more susceptible to rearrangements (212-214). While BACs are relatively stable, internal deletions and rearrangements have been observed to occur within the ESC. It is thought that repeat sequences within the
BAC mediate intrachromosomal recombination. This type of event may have impacted the identification of targeted Hprt-Sox9-EGFP-BACs, as one potential clone contained both 5’ and 3’ sequence integration (clone S18, Figure 4.9B), yet no Sox9 or EGFP sequences could be detected, suggesting a large internal deletion. If this was a high frequency deletion, in conjunction with the larger size of the BAC and higher probability for terminal end shearing due to manipulation, it is not surprising that no correct Sox9 clones were obtained. It is also not surprising that the HAT resistant clones successfully recombined at the 3’ BAC end, yet contained large deletions/breaks. Recombination at the 3’ end is essential for Hprt complementation and survival in HAT media. While other clones may have recombined at the 5’ end only, forward selection in HAT media would not have identified these clones.

Although the integration of BACs as large as 200 kb have been reported in transgenic mouse models, these approaches relied on random integration and the integrity of the BAC was not verified (176;215;216). The Gene Expression Nervous System Atlas (GENSAT) BAC Transgenic Project reports limited success in working with genomic fragments 200 kb or greater in size (176). While there have been reports of gene targeting where the construct is a recombineered BAC, these reports are not directly comparable to our methods based on the larger amount of homologous sequences (up to 100s of kb) and the smaller amount of newly integrated sequence (typically less than 2kb) (217-219).

This issue of homology arm size, and more importantly the ratio of homology arm size-to-length of non-homologous intervening sequence, could be the determining factor for the lack of recombination efficiency in our targeting attempts. While it has been
previously reported that the length of sequence between homology arms plays a minor role in homologous recombination efficiency, these reports have not been based on large inserts of DNA, such as BACs (129;220-222). Other reports demonstrate the ability to delete large spans of DNA with a smaller transgenic cassette during knock-out studies, these however focus on the large sequence of DNA at the endogenous locus and not carried within the targeting construct itself (223). Heaney et al. targeted BACs (114 kb and 146 kb in size) to the Hprt locus, yet witnessed a targeting frequency 6 to 7 times less than when using a linearized 15 kb control plasmid (207), confirming that the size of the intervening sequence does play a role in targeting efficiency. The upper size limit of the length of intervening sequence as well as the role of specific intervening sequences on the efficiency of the targeting event are unknown.
CHAPTER 5
EGFP REPORTERS AND ENRICHMENT POTENTIAL

5.1 Introduction

Transplantation of stem cell derived progenitors, tissues, or organs has the potential to repair or restore function to the host at a level unachievable by any mechanical device or drug treatment (224). Ideally, the proliferative capacity of a transplanted stem or progenitor cell would allow for a single cell-based treatment, versus the chronic, and sometimes toxic, nature of many drug based therapies. One of the best characterized examples of stem cell therapy is the introduction of hematopoietic stem cells (HSCs) via bone marrow transplant. For disorders involving the hematopoietic system (leukemia, lymphoma, myeloma, anemia) as well as for other disorders of non-hematopoietic origin where either the pathology or the treatment alter hematopoietic development or function, the transplantation of HSCs has been used to reconstitute the hematopoietic system of the recipient. HSCs, found in bone marrow, spleen, blood, and fetal liver, are self renewing stem cells capable of forming both the myeloid (megakaryocyte, erythrocyte, granulocyte, monocyte) and lymphoid lineages (T cells, B cells, natural killer cells) (225-228). The number of transplanted HSCs correlates positively with the time necessary to reach a critical number of mature blood cells within the recipient (229;230), thus cell surface markers have been identified to facilitate the enrichment and purification of HSCs for transplant (231-234).
Similar attempts at transplantation and repopulation have been made with another bone marrow derived stem cell, the MSC. Because the MSC can differentiate into the osteogenic lineage, donor MSCs were infused into recipients with the genetic disorder Osteogenesis Imperfecta (OI) in an attempt to provide a new healthy population of collagen producing osteoblasts. However, as discussed in Chapter 1, engraftment of donor cells was relatively low (83-85). This outcome could be due to limiting numbers of true MSCs, poor homing, poor engraftment, or even for an inherent inability of the cells to undergo necessary proliferation and differentiation \textit{in vivo}. In any case, the lack of standardized cell surface markers for MSCs hinders definitive experimentation and therapy development.

An alternate approach is to use transcriptional reporters. Lineage and differentiation stage-specific reporters can allow for the enrichment of a range of differentiated osteogenic progenitors, permitting experimental determination of the most therapeutically effective cell (Figure 5.1).
Figure 5.1: Osteoblast differentiation from ESCs. Expected cell types are in black font and relevant gene expression markers are in red font. Lineages in gray have either not been tested or are suggested to be absent due to an absence of mature markers. Gene expression markers such as *Brachyury*, *Twist2*, *Runx2*, *Sox9*, *Osterix*, *Osteocalcin*, and *Nestin* can be used to create genetic reporters in an attempt to identify osteogenic populations. (source Bronson Laboratory)
5.2 Results

5.2.1 Bry-EGFP ESC Reporter

Brachyury (Bry) is an early mesodermal marker, a T-box transcription factor, and has been shown to be required for mesoderm formation and axial development. It is expressed transiently in the nascent and migrating mesoderm of the primitive streak, and persists through notochord development. As mesodermal cells migrate away from the primitive streak, Bry expression is down regulated (235-237). Fehling et al. created a Bry-GFP knock-in ESC reporter line to identify hematopoietic progenitors arising from mesodermal differentiation within embryoid bodies (EBs) (162). The knock-in cassette contains arms of homology to exon 1 of the Bry locus, resulting in the replacement of the first two-thirds of the exon by the GFP-cassette as well as disrupting the Bry gene (Figure 5.2). This Bry-GFP reporter allowed EB cells from various days of culture to be separated into mesodermal populations (Bry+) and neuroectodermal populations (Bry-).

Two color FACS for GFP and a fluorescent tagged antibody to receptor tyrosine kinase Flk1 allowed the Bry+ population to be separated into hemangioblast (Flk1+) and pre-hemangioblast mesoderm (Flk1-) populations. Because axial and appendicular bone is also of mesodermal origin, we hypothesized that a Bry-GFP+ osteogenic culture-derived population would be enriched for osteogenic progenitors.

Flow cytometry was used to assess GFP expression of the Bry-GFP ESC line within the osteogenic culture. Previous data in our laboratory has demonstrated a peak in Bry mRNA expression at Day 4 of the osteogenic culture (D4 OC) with almost complete
Figure 5.2: Brachyury GFP knock-in reporter. The knock-in EGFP targeting vector results in the replacement of the first two-thirds of Brachyury exon 1. LoxP sites are indicated by black triangles flanking the Neo gene. The hatched box indicates the SV40 polyadenylation signal. TK: Herpes simplex thymidine kinase gene. H: HincII (figure and legend from (162)).
down-regulation by D7(116;119). Therefore, flow cytometry of Bry-GFP osteocultures was performed at D2 OC, D4 OC, and D6 OC (Figure 5.3). At D2 OC, Bry-GFP had a mean fluorescence of 9, with the BK4 no-GFP control culture at a mean fluorescence of almost 8, and only about 4% of the Bry-GFP cells had a fluorescent intensity higher than background. D4 OC and D6 OC however, showed a large boost in GFP fluorescence with greater than 50% of Bry-GFP cells higher than background, with a mean fluorescence of 52.8 and 64.8 respectively.

Cells from D4 and D6 OC plated at 1E^5 cells/10 cm dish were harvested, FACS sorted, and the resultant Bry-GFP positive and Bry-GFP negative populations were re-plated in the osteogenic culture. Undifferentiated Bry-GFP ESCs were set as background and cells were analyzed with an incident beam of 488 nm and a 530/40 nm bandpass filter. Forward- and side-scatter light distribution gates were used to exclude debris, clumps, and dead cells. Using undifferentiated Bry-GFP ESCs as the background, D4 OC contained 35.6% positive cells and D6 contained 31.6% positive cells separated into two populations, a high expressing population consisting of 12.8% positive cells and an intermediate population consisting of 18.8% positive cells. Replating the sorted cells at 2E^3 cells/10 cm dish (typical density for osteocultures) yielded no colony formation for either timepoint (data not shown).

The experiment was repeated, replating at a higher density (1E^4 cells/10 cm dish). For this experiment, parallel cultures of non-GFP expressing BK4 ESC derived osteogenic cultures were used as the control giving a more accurate representation of background auto-fluorescence due to the size of the osteogenic cells. Osteocultures were originally plated at 1E^5 cells/10 cm dish, harvested with trypsin prior to sorting, and
Figure 5.3: Flow cytometric analysis of Bry-GFP fluorescence throughout the osteogenic culture. Fluorescent expression of Bry-GFP at days 2, 4, and 6 of the osteoculture, cultures of non-fluorescent BK4 ESCs were set as background. Percent positive cells are indicated, mean fluorescence is indicated.
assessed by FACS at D4 and D6 OC. D4 OC contained 13% EGFP positive cells and D6 contained 14% positive cells (Figure 5.4). Cultures were replated at 1E^4 cells/10 cm dish. Again D6 sorted cultures displayed no colony formation (data not shown). At this higher density however, the D4 positive sort cultures had an average of 6 colonies per plate, while the negative sort population had no colony formation (Figure 5.5). Of the colonies formed on the D4 positive sort plates, about 50% mineralized.

Despite the better colony forming efficiency from the Bry-GFP positive sort population in comparison to the Bry-GFP negative sort population, it was concluded that the Bry-GFP reporter would not yield efficient enrichment of a population of osteogenic progenitors with this low colony forming efficiency. A follow up to these experiments might include looking at gene expression in the sorted populations, sorting cells from EBs rather than osteocultures, or trouble-shooting the poor performance of this cell line post-harvest from the osteoculture and FACS (see section 6.3), however these experiments are beyond the scope of this thesis.
Figure 5.4: FACS of Bry-GFP cells at D4 and D6 OC. Bry-GFP cells were sorted at D4 and D6 OC, parallel cultures of control BK4 osteocultures were set as background. D4 culture sorted for 13% EGFP positive cells and D6 sorted for 14% EGFP positive cells. Representative histograms are included.
Figure 5.5: Colony formation of Bry-GFP sort populations. Bry-GFP osteogenic cultures were FACS sorted at days 4, and 6 OC and replated for 21 days in the osteogenic culture. Cultures were stained at day 21 replate-OC with silver nitrate/methyl green. No colony formation was observed for D6 replate cultures (data not shown). Representative D4 Bry-GFP positive and negative sort plates are shown.
5.2.2 Nestin-EGFP ESC Reporters

A subset of bone is ectodermally derived. Fate mapping of cranial neural crest migration has demonstrated that this population is responsible for bone formation of the viserocranium (jaws and branchial arch derivatives) as well as sections of the neurocranium (including frontal bones and the sagittal suture) (238;239). In vitro experiments have also demonstrated the ability to induce osteogenic differentiation of both cranial as well as trunk neural crest cells, and upon transplantation, the cells are capable of forming bone in vivo (240;241). Nestin is a marker of neural crest cells (242) and we hypothesized that a Nestin-GFP reporter can be used to enrich for neural crest derived osteogenic progenitors from ESC-derived EB or osteogenic cultures.

To create an Hprt-targeted Nestin-EGFP ESC reporter, pNestin-GFP (211) was ligated into pMP8NEBΔLacZ (143) to contain Hprt regions of homology (see Section 4.2.6). This Hprt-Nestin-EGFP construct was electroporated into BK4 ESCs to create 17 Hprt-Nestin-EGFP ESC reporter clones. A subset of the clones was assessed for fluorescent expression by flow cytometry. Analysis by Woll et al. has demonstrated that Nestin mRNA in the osteogenic culture is up-regulated over 200-fold by D2 EB relative to D0 EB expression (119). Therefore Hprt-Nestin-EGFP ESC reporter clones were analyzed by flow cytometry at D0, D2, and D4 of EB differentiation for Nestin EGFP fluorescence. EBs at each respective day of culture were trypsinized to a single cell suspension and assessed on a Becton Dickinson FACScan with a wavelength excitation of 488 nm and emission detection of 530/40 nm bandpass. Hprt-Nestin-EGFP ESC clones however, showed no difference in fluorescence as compared to control EB cultures of BK4 cells (Figure 5.6).
**Figure 5.6**: Analysis of Hprt-Nestin-EGFP ESC reporter clones for EGFP expression. Hprt-Nestin-EGFP clones 1,2,4, and 5 were analyzed by flow cytometry at Day 0, Day 2, and Day 4 of EB differentiation. BK4 D0 EB cultures served as the negative control.
To test the construct, a second Nestin-EGFP ESC reporter line was created by random integration of Nestin-EGFP into BK4 ESCs (see Section 4.2.6) and was also assessed by flow cytometry. ESC clones established from both the forward- and backward-orientation of the Nestin-EGFP construct were assessed for fluorescence at D2 EB differentiation, and the random Nestin-EGFP ESC lines showed no difference in fluorescence as compared to control BK4 D2 EB cultures (Figure 5.7).

The transgenic ESC lines were tested for the presence of the EGFP mRNA transcript within the cells. RT-PCR of cDNA prepared from RNA isolated from the clones detected EGFP mRNA in both lines (Figure 5.8A). There were no exon/exon boundaries within the Nestin construct, thus it was not possible to design primers spanning an intron to rule out amplification of the DNA sequence. However, upon DNase treatment, the EGFP mRNA product was detectable within the cells suggesting the presence of Nestin-EGFP mRNA (Figure 5.8B).

Review of the literature revealed that the 2\textsuperscript{nd} intron fragment of the rat Nestin promoter used in this construct contains two highly conserved elements, which are midbrain and central nervous system enhancer elements, driving Nestin-EGFP reporter expression in these tissues, and perhaps to the exclusion of other tissues (further discussed in section 5.3) (243-245). While follow-ups to these experiments might include looking for fluorescence in ESC-derived neural lineages, these experiments are beyond the scope of this thesis.
**Figure 5.7:** Analysis of Randomly Integrated Nestin-EGFP ESC reporter clones for EGFP expression. A subset of Nestin-EGFP-Neo clones were analyzed by flow cytometry at D2 of EB differentiation. BK4 D2 EBs served as the negative control. F= forward orientation clone, B= backward orientation clone.
Figure 5.8: RNA analysis of Nestin-EGFP clones. (A.) RT-PCR of cDNA isolated from D3 EB Hprt-Nestin-EGFP clones and Nestin-EGFP-Neo randomly integrated clones to detect EGFP mRNA. Expected product size is shown. H= Hprt targeted clone, F= forward orientation random integrant, B= backward orientation random integrant. (B.) DNase treatment of Hprt-Nestin-EGFP clone 5 and subsequent PCR analysis.
5.2.3 Hprt-Osterix-EGFP ESC Reporter

Osterix (Osx), a zinc finger-containing transcription factor, has been shown to be specific to the osteogenic lineage. Osx, detected as early as 13 dpc and robustly by 15.5 dpc and on, is essential for osteogenesis as Osx null embryos have arrested osteoblast differentiation and completely lack bone formation, although cartilage development is normal (33). Osx has been shown to be downstream of Runx2 and dependent on Runx2 for expression (33;34). It has been suggested that Osx expression is necessary for the differentiation of Runx2-positive preosteoblasts into mature osteoblasts (33). As Osx is specific to the osteogenic lineage, marking a later timepoint in osteoblast differentiation, we hypothesized that an Osx-EGFP reporter will allow for the enrichment of a committed osteogenic cell population from ESC-derived osteogenic cultures.

As previously described in section 4.2.5, three electroporations of Hprt-Osx-EGFP BAC into BK4 ESCs resulted in 5 clones with a correctly targeted and intact BAC. Five clones, Osx-2,5,7,10,12, were subsequently assessed for their level of Osx-EGFP fluorescent expression by flow cytometry. Previous mRNA analysis in the laboratory has shown Osx expression in the osteogenic culture to be up-regulated after D7 of the culture, therefore Osx-EGFP reporter clones were tested at D6 – D19 of culture (Figure 5.9). Clones at each specific day of the osteogenic culture were harvested by trypsinization and assessed by flow cytometry on a Becton Dickinson FACScan, wavelength excitation 488nm/emission detection with a 530 nm bandpass filter, collecting no less than 10,000 events per sample. The Osx-EGFP reporter clones showed no difference in fluorescence as compared to control cultures of BK4 cells at the same timepoint of differentiation.
Figure 5.9: Analysis of Osx-EGFP expression throughout the osteoculture. Flow cytometry analysis of Osterix clones 2 and 4, collecting no less than 10,000 events per sample.
Despite the apparent lack of fluorescence in Osx-EGFP transgenic cells from the osteoculture, RT-PCR of cDNA isolated from the clones, with primers designed to span an intron to exclude amplification from DNA, revealed the presence of both the long and short isoforms of transgenic Osx-EGFP mRNA (Figure 5.10) (See Table 2.2 for primer sequences). Woll et al. have demonstrated that Osx mRNA expression levels increase by 10-fold from the ESC to D14 osteoculture (119); this 10-fold increase in mRNA from the whole plate may be due to all cells increasing expression by 10-fold, or by a small subset of cells increasing expression by a 1000-fold or anything in between. If a small subset of the total population of cells was expressing Osx at a high level, we would expect to detect Osx-EGFP fluorescence more readily than at the other extreme. We hypothesized that by increasing the number of cellular events collected by flow cytometry we might see any rare high Osx-EGFP expressors. Clones were again assessed by flow cytometry, this time collecting 10 X the amount of cells as previously. The flow cytometry analysis, even with the collection of >100,000 events, detected little to no population of cells with a significantly increased fluorescence in comparison to the background BK4 control cells (Figure 5.11).

While further experiments could be employed to determine the level of Osx-EGFP RNA and protein expression (QRT-PCR, Western), or to increase expression (treatment with BMP2, IGF-I, MEK1 inhibitors), these experiments are beyond the scope of this thesis.
**Figure 5.10:** Detection of Osx-EGFP mRNA in Hprt-Osx-EGFP ESC derived osteocultures. RT-PCR of cDNA isolated from Osterix clone 10 at D15 and D17 of the osteoculture. Visualization on a 1% agarose gel. Both the short and long *Osterix* isoforms were assayed for using primers designed to span an intron to ensure RNA specificity. For the short isoform, primer 1 spanned exon1/exon3, for the long isoform primer 1 spanned exon2/exon3; the same primer 2 was used for both assays, routed in the EGFP cassette, creating a product of 1,051 bp for both assays. BK4 osteoculture RNA was used as a negative control.
Figure 5.11: Analysis of the Hprt-Osx-EGFP ESC derived osteoculture for a small population of EGFP expressing cells. Flow cytometry analysis of Osterix clone 10, collecting no less than 100,000 events per sample. Solid line represents Hprt-Osx-EGFP osteoculture sample, Dotted line represents control BK4 osteoculture. Mean fluorescence is indicated.
5.2.4 Hprt-Twist2-EGFP ESC Reporter

5.2.4.1 Initial Characterization of the Twist2-EGFP Reporter:

Twist2, a basic helix-loop-helix transcription factor, is first detected in mice at 10.5 dpc in early mesoderm condensations of the limb bud as well as the body wall. During embryogenesis, Twist2 transcripts accumulate in chondrogenic cells which will develop into bone such as the vertebrae, limbs, and digits; expression is also detected in the dermal layer of skin (175;246). Early experiments where Twist2 expression in osteogenic cell lines rapidly decreased upon osteogenic differentiation of the cultures, suggested a role for Twist2 in the negative regulation of osteogenesis (175;247). Further evidence included a loss of osteogenic morphology of the cells and decreased alkaline phosphatase and Type I collagen levels when Twist2 expression was enforced by an exogenous construct (175). Ultimately Karsenty et al. demonstrated the interaction of Twist2 with the DNA binding domain of Runx2, and correlated this with inhibition of Runx2 function and delayed initiation of osteoblast differentiation (32). As Twist2 has been demonstrated to play a major role in the regulation of osteoblast differentiation, we hypothesize that a Twist2-EGFP reporter can be applied to enrich for osteogenic progenitor populations from ESC-derived osteogenic cultures.

Electroporation of Hprt-Twist2-EGFP-BAC DNA into CAGGS-EYFP ESCs resulted in the creation of one correctly targeted clone, T1. Because the Twist2-EGFP reporter was integrated into the constitutively expressing CAGGS-EYFP cell line, T1 was first assessed to determine if EGFP fluorescence was distinguishable from background EYFP fluorescence by flow cytometry. T1 cells were taken at early timepoints of the
osteoculture (ESC, D0 OC, D4 OC) and fluorescent expression levels were compared to CAGGS-EYFP ESCs (Figure 5.12). The dual reporter T1 ESCs and T1 D0 OC showed a shift in fluorescence (mean fluorescence 8.46 and 8.31 respectively) as compared to CAGGS-EYFP ESCs (mean fluorescence 3.56), determining the background level of EGFP expression within the dual expressing line. By D4 of the osteoculture, a greater shift in fluorescence was visible (mean fluorescence of 20.9) with 16.6% of the cells higher than background, demonstrating the expected function of the Twist2-EGFP reporter.

After the verification of the EGFP signal, the fluorescent expression pattern of Twist2-EGFP was assessed throughout the osteogenic culture by flow cytometry. Cells at specific days of the osteoculture were trypsinized and collected from the plate to be assessed by flow cytometry on a Becton Dickinson FACScan, wavelength excitation 488 nm/emission detection with a 530 nm bandpass filter, collecting no less than 10,000 events per sample. Previous Twist2 mRNA analysis in the lab demonstrated a peak in expression at D7 of the osteoculture with decreasing levels upon continued culture and differentiation, therefore T1 EGFP fluorescent expression was assessed at days 2, 4, 6, and 10 of the osteoculture (Figure 5.13). Paralleling the earlier mRNA analysis, EGFP expression was detectable at D2 of the osteoculture, with increased levels observed at D6, and decreased expression by D10. By comparing the fluorescent expression of Twist2-EGFP/EYFP osteogenic cultures to parallel CAGGS-EYFP culture timepoints, a separate EGFP peak was identifiable by flow cytometry in cells from D2-D6 of the osteoculture. This shift in EGFP expression potentially would allow for the sorting and enrichment of osteoprogenitors by FACS.
Figure 5.12: Initial verification of Twist2 EGFP fluorescence. Twist2 EGFP/EYFP cells were taken at different timepoints of the osteoculture (ESC, D0 OC, D4 OC) and fluorescent expression levels were compared to CAGGS-EYFP ESCs using flow cytometry. CAGGS-EYFP ESC had a mean fluorescence of 3.56, Twist2-EGFP ESC of 8.46, Twist2-EGFP D0 OC of 8.31, Twist2-EGFP D4 OC of 20.9.
Figure 5.13: Flow cytometric analysis of Twist2-EGFP fluorescence throughout the osteogenic culture. Fluorescent expression of Twist2-EGFP at days 2, 4, 6, and 10 of the osteoculture compared to parallel osteocultures of control CAGGS-EYFP cells. Mean fluorescence is indicated.
Using a MoFlo High Performance Cell Sorter, it was possible to sort two populations of EGFP expressing cells, those with high EGFP fluorescence (Twist2-EGFP positive cells) and those with lower or no EGFP fluorescence (Twist-EGFP negative cells) at different timepoints throughout the osteoculture (D2-D8 OC). Parallel cultures of CAGGS-EYFP cells were set as background and gates were set to include cells with a shift in EGFP fluorescence greater than the EYFP background. Cells were analyzed with an incident beam of 488 nm and a 530/40 nm bandpass filter. Forward- and side-scatter light distribution gates were used to exclude debris, clumps, and dead cells. Cells from D2 osteoculture yielded an average of 2% positive cells, cells from D4 osteoculture yielded an average of ~14% positive cells, cells from D6 osteoculture yielded an average of ~9% positive cells, and cells from D8 osteoculture yielded ~ 1% positive cells (Figure 5.14). Osteocultures for the D2 sort were plated at a higher density (1E⁶/10 cm plate as compared to D4-D8 OC at 1E⁵/10 cm plate) to obtain enough cells for the sort. Trials were also run with D4 and D6 OC cultures plated at the higher 1E⁶/10 cm plate density, however the increased density caused the osteogenic differentiation culture to mature at a faster rate resulting in an average of ~9.9% positive cells for D4, and ~1.2% positive cells for D6.

Immediately after sorting, RNA was isolated from the positive and negative sort populations. QRT-PCR assays were designed to detect endogenous Twist2 expression as well as Twist2-EGFP specific expression. QRT-PCR analysis demonstrated the fidelity of the gene-specific reporter as Twist2-EGFP expression recapitulated endogenous Twist2 expression in the positive sort samples (Figure 5.15). The RNA expression pattern also paralleled the fluorescent expression pattern of the osteoculture observed by flow
Figure 5.14: FACS of Twist2-EGFP cells throughout the osteoculture. Twist2-EGFP cells were sorted at days 2, 4, 6, and 8 of the osteoculture. Parallel cultures of CAGGS-EYFP cells were set as background and gates were set to only choose cells with a shift in fluorescence greater than EYFP background. D2 cultures sorted for an average of 2% positive cells, D4 for 14% positive cells, D6 for 9% positive cells, and D8 for ~1% positive cells.
Figure 5.15: QRT-PCR analysis of Twist2 endogenous and transgenic expression in FACS positive sort populations throughout osteogenic culture. Immediately after sorting on days 4, 6, and 8 of the osteogenic culture, positive sort cells were processed for mRNA analysis. Assays were designed to detect endogenous Twist2 expression as well as Twist2-EGFP specific expression (see Table 2.2 for primers used). Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7 and displayed relative to the unmanipulated parent control culture at D0 OC. RNA was harvested from D2 positive sort population also, however not enough RNA was isolated for analysis.
cytometry with an observed peak at D6 of the culture. QRT-PCR analysis also demonstrates the strong boost of Twist2 mRNA expression within the osteoculture; with Twist2 mRNA from the D6 OC ~50 fold higher than at D0 OC.

Comparison of endogenous and transgenic Twist2 mRNA expression in both the positive and negative sort samples shows that the positive and negative sort populations had similar levels of expression (Figure 5.16). One might expect a higher expression of Twist2 in the positive sort population as compared to the negative sort population, if the Twist2-EGFP reporter was enriching for a Twist2 expressing population by FACS. A couple of factors may explain this result. First, because of the background levels of EYFP fluorescence, not all Twist2 fluorescing cells were able to be selected into the positive sort population. Thus some Twist2 expression in the negative sort population is expected. Second, because of the lag between mRNA transcription and protein translation, a substantial amount of Twist2 message was likely present in the Twist2-EGFP negative population – with these cells representing a stage earlier in the differentiation process (further discussed in section 5.3, Figure 5.27).

To demonstrate the presence of pre- or early Twist2 expressing cells in the negative sort population, the Twist2-EGFP negative sort population was assessed by flow cytometry immediately and after additional culture. We hypothesized that some of these cells were in the early stages of Twist2 expression, and given time to further differentiate, would progress into fully expressing Twist2 cells. To test this hypothesis, cells were sorted at D6 OC and the resultant negative sort population was subjected to flow cytometry analysis (Figure 5.17). On the day of the sort, as expected, the negative sort population was determined to have 0% EGFP positive cells. However, if this negative
Figure 5.16: QRT-PCR analysis of Twist2 expression in FACS sort populations. Analysis of endogenous and transgenic Twist2 mRNA expression in both the positive and negative sort populations immediately after the sort of D4 and D6 osteocultures. Each bar represents the average of two sort experiments (except Twist2-EGFP D4- sort in which data was obtained for only one sort population), error bars denote standard deviation. Assays were designed to detect endogenous Twist2 expression as well as Twist2-EGFP specific expression (see Table 2.2 for primers used). Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7 and displayed relative to Control D0 OC.
Figure 5.17: Flow cytometry of Twist2-EGFP negative sort population. Fluorescent expression of Twist2-EGFP negative sort cells immediately after a D6 sort (D0), and after replating for two days in the osteogenic culture (D2). Cultures of CAGGS-EYFP ESCs were set as background. Percent positive cells are indicated.
sort population was immediately replated in the osteogenic culture and subsequently harvested and reassessed by flow cytometry two days later, these cells now contained 2.72% EGFP positive cells, lending support to the idea that a premature Twist2 population exists within the negative sort population. This early Twist2 population likely contains significant levels of Twist2 mRNA, contributing to the nearly equivalent Twist2 and Twist2-EGFP mRNA levels in the negative and positive sort populations.

5.2.4.2 Analysis of Twist2-EGFP D2 and D8 Sort Cultures:

Experiments were next performed to test the hypothesis that the Twist2-EGFP reporter could identify and enrich for osteogenic progenitors from the ESC derived heterogeneous osteogenic culture at different days of differentiation (D2 – D8 OC). After sorting for Twist2-EGFP, cells were replated and assessed for colony formation and lineage specificity within the osteogenic culture. The D2 OC Twist2-EGFP positive population formed an average of 7 colonies per plate while the D2 OC Twist2-EGFP negative population formed an average of 22.5 colonies per plate. The control unsorted population, which does not undergo the sort procedure but is otherwise harvested and manipulated similarly, also had higher colony formation than the positive sort population with an average of 15.5 colonies per plate (Figure 5.18). Of the 7 colonies on the positive sort plate, only 3 of the colonies showed any mineralization (42.8%), as compared to 36.3% mineralization for the negative sort and 46.9% mineralization for the unsort population. The replating of the D8 OC Twist2-EGFP positive population was even less efficient with an average of 1-2 colonies per plate and no mineralization
Figure 5.18: Colony formation of Twist2-EGFP sort populations. Twist2-EGFP osteogenic cultures were FACS sorted at days 2, 4, and 6 OC and replated at $1E^4$ cells/10 cm plate for 21 days in the osteogenic culture. Cultures stained at day 21 replate-OC with silver nitrate/methyl green. Average number of colonies formed per plate and percent mineralization are displayed. Cultures were also sorted at D8 OC and replated for the osteogenic culture, however all plates were harvested for mRNA analysis. D8 OC positive sort had an average of 1.5 colonies per plate (0% mineralization). D8 OC negative and unsort populations yielded no colony formation. The parallel unmanipulated control culture for this experiment had an average of 27 colonies (42% mineralization). GFP + sort = positive sort population, GFP – sort = negative sort population, un sort = control replated population.
observed (not shown). No colonies formed after plating the D8 OC Twist2-EGFP negative and unsorted populations. This low colony forming efficiency coupled with low or no mineralization demonstrated the absence of a proliferative osteogenic progenitor within the D2 and D8 positive sort populations. Alternatively, harvesting and sorting at D2 OC may damage a fragile cell population that is unable to proliferate upon replating in the osteogenic culture.

To further analyze the colonies derived from the D2 OC and D8 OC sorted populations, cells were harvested from replated osteogenic cultures 21 days post-sort (D21 replate-OC), and mRNA was prepared for assessment by QRT-PCR (Figure 5.19). Markers of osteogenic differentiation, Type I collagen (Col1a1) (248) and Osteocalcin (Ocn) (249;250), were dramatically lower in the replated cultures from Twist2-EGFP positive and negative populations as compared to a parallel unmanipulated osteoculture from which the sort culture was originally derived. The lack of osteogenic markers in the sorted populations suggests that the sort process is detrimental to osteogenic progenitors. Furthermore, osteogenic gene expression of the D2 OC Twist2-EGFP positive sort cells did not out-perform the negative sort cells, showing no enrichment for osteogenic progenitors.

Markers of competing mesodermal lineages, Myh11 (Myosin heavy peptide 11), an expression marker for smooth muscle (251), as well as Lpl (Lipoprotein lipase), a marker for adipose cells (252), were comparable in the sorted populations and the parallel unmanipulated control cultures, suggesting a lack of osteogenic lineage specificity in the D2 and D8 positive sort populations. CD45, an expression marker for hematopoietic
Figure 5.19: QRT-PCR analysis of (A.) Day 2 and (B.) Day 8 FACS sorted and replated Twist-EGFP populations at Day 21 replate OC. Positive and negative sort populations were analyzed for Col1a1, Ocn, Myh11, Lpl, CD45, and Col10a1 gene expression. D2 data represents the average of two sort experiments, error bars denote standard deviation. Negative sort values from the D8 sort are not available as no colony formation was observed. Insufficient sample was collected for D8 +sort Col10a1 analysis. Control = D21 OC of unmanipulated parallel control osteoculture from which the sort culture was derived. Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7 and displayed relative to Control D0 OC.
cells (253), and Col10a1 (Type X collagen) a marker for hypertrophic chondrocytes (254), also showed no significant changes between the unmanipulated control culture and the positive and negative sort cultures.

### 5.2.4.3 Analysis of Twist2-EGFP D4 Sort Cultures:

The D4 OC Twist2-EGFP positive sort population yielded an average of 25.3 colonies per plate with a percent mineralization of 58.3% (Figure 5.18). The unsorted control osteoculture formed an average of 12.9 colonies per plate with a percent mineralization of 47.5%. The negative sort population performed similarly to the positive sort population with an average of 23.5 colonies per plate and a percent mineralization of 57.2%. The D4 sorted populations demonstrated the best colony formation and percent of colonies mineralized out of all the tested timepoints.

Gene expression analysis was performed on the D4 sort cultures to assess the similarities and differences between the positive and negative sort cultures, as well as to determine the lineage specificity of the populations (Figure 5.20). Directly after the sort (D0 replate-OC), D4 OC Twist2-EGFP RNA analysis reveals an enrichment for osteogenic progenitors within the Twist2-EGFP positive sort cells based on robust Osterix and Runx2 expression compared to the Twist2-EGFP negative sort population. Runx2 and Osterix are transcription factors which are specific to and essential for osteogenesis (17;33). Sox9, a marker of bipotential osteo-chondroprogenitor cells (35), Brachyury (Bry), an early mesodermal marker (255), and Nestin, a marker of the neural crest lineage (242), are all at levels comparable to the unmanipulated control culture for both the positive and negative sort populations. While we might expect Bry or Nestin or
Figure 5.20: QRT-PCR analysis of Day 4 FACS sorted and replated Twist-EGFP populations at D0, D14, and D21 replate-OC. Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7, displayed relative to Control D0 OC. D0 Osterix and Runx2 data represent the average of two sort experiments, error bars denote standard deviation. Control = unmanipulated parallel control osteoculture from which the sort culture was derived.
Sox9 to also be increased in the Twist2-EGFP positive sort relative to the negative sort, these results are in agreement with the Twist2-EGFP positive sort population representing an intermediate osteogenic progenitor, rather than the earliest progenitor for the lineage which might be more likely to express Bry, Nestin, or Sox9.

By D14 replate-OC however, the robust osteogenic expression of the Twist2-EGFP positive sort culture was diminished as both Osterix and Runx2 expression were dramatically decreased, with Osterix expression dropping to almost undetectable levels, and Runx2 expression now at only half of the negative sort population. The mRNA levels for Bry and Nestin are largely unchanging, Sox9 levels however are substantially increased, perhaps due to Sox9 expression in a non-osteogenic lineage.

At D21 replate-OC, cells from the D4 Twist2-EGFP positive sort have higher mRNA levels for the osteogenic marker Col1a1 than the negative sort population. Type 1 collagen is a major structural component of bone, however it is not specific to bone and is also expressed in other tissues such as cartilage, tendon, and fibroblasts. Interestingly, Osteocalcin, a late osteogenic marker which is specific to osteoblasts, is similarly expressed in the control unsorted, positive, and negative sort populations. One explanation for this is that Twist2-EGFP-low, more primitive progenitors are present in all three populations to a similar extent and are responsible for the mineralizing colonies and Osteocalcin expression at D21. It seems likely that the cells that have the high Runx2 and Osx expression directly post-sort are not the cells forming mineralized colonies that express Osteocalcin. Gene expression analysis suggests that while Twist2-EGFP sorting at D4 of the osteoculture may choose for a progenitor with the proliferative capacity to form colonies, this progenitor has limited osteogenic potential and lacks
lineage specificity as both *Myh11* and *Lpl* expression levels are approximately 2-fold higher in plates derived from the positive sort population than plates derived from the negative sort population.

As the D4 Twist2-EGFP negative sort population contained cells with equal colony forming efficiency as the Twist2-EGFP positive sort population, it was next determined if a larger, more encompassing, population of the Twist2-EGFP positive cells could be separated from the Twist2-EGFP negative cell population. This task was complicated by the background levels of EYFP expression obscuring the identification of low EGFP fluorescing cells. However, by comparing the shift in fluorescence of the EGFP expressing cells to the background EYFP levels, it was possible to identify and set gates to enrich for a larger population of the D4 Twist2-EGFP positive cells (Figure 5.21). Gates were set obtaining 53% positive Twist2-EGFP cells, as compared to earlier sorts containing 14% positive cells which had consisted only of the highest EGFP expressors. By increasing the positive sort gate to include the moderate Twist2-EGFP fluorescing cells, the positive sort gates also enriched for a larger population of endogenous and transgenic *Twist2* mRNA expressing cells. QRT-PCR analysis of the sorted populations using the broader gate resulted in higher endogenous and transgenic *Twist2* mRNA expression in the positive sort population as compared to the negative sort population (Figure 5.22A). We also observed that the Twist2-EGFP positive sort cells at D0-replate had even higher *Osterix* and *Runx2* expression compared to the negative sort population, further supporting our earlier interpretation that the Twist2-EGFP positive population is enriched for osteogenic progenitors (Figure 5.22B).
Figure 5.21: FACS of D4 Twist2-EGFP OC to include a larger population of Twist2-EGFP fluorescing cells. (A.) Representation of earlier sort data with the positive sort gate set to include only cells with a shift in fluorescence greater than EYFP background. (B.) Enriching for a larger Twist2-EGFP fluorescing population by increasing the positive gate sort to include moderate Twist2-EGFP fluorescing cells.
Figure 5.22: QRT-PCR analysis of Day 4 larger gate sort at D0 replate-OC. Positive sort gate was enlarged to obtain cells with moderate to high Twist2-EGFP fluorescence. (A.) Analysis of endogenous Twist2 and Twist2-EGFP expression. (B.) Analysis of Osterix and Runx2 expression. Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7, displayed relative to Control D0 OC.
5.2.4.4 Analysis of Twist2-EGFP D6 Sort Cultures:

D6 Twist2-EGFP sort populations were assessed by QRT-PCR immediately post sort (D0 replate-OC), demonstrating enrichment for osteogenic progenitors in the Twist2-EGFP positive population, as shown by the expression of both Osterix and Runx2, which was decreased in the negative sort cells. Of interest, Nestin expression is significantly higher in the Twist2-EGFP positive sort population as compared to both the negative sort cells and the parallel control osteoculture, suggesting the Twist2-EGFP reporter may enrich for osteogenic progenitors derived from neuroectoderm at this stage of the osteoculture (D6), or perhaps it is an indication of Twist2 expression in a competing Nestin-expressing lineage at this stage (Figure 5.23).

A significant portion of colonies in the D6 positive sort osteoculture exhibited characteristic osteogenic morphology at 7 days post sort (D7 replate-OC). The colonies were symmetrical with outer edges consisting of a proliferating zone of cuboidal cells. The middle of the colony was opaque, suggesting enhanced extra cellular matrix deposition, and mineral deposits were visible within the center of the colony (Figure 5.24). Colonies formed from the negative and unsort populations did not display the osteogenic qualities observed in the positive sort population.

By D14 replate-OC however, the characteristic osteogenic morphology of the D6 Twist2-EGFP positive sort population is less discernable with less matrix and mineralization than anticipated based on the morphology of the day 7 colonies (Figure 5.24). Gene expression analysis at this timepoint corresponds with a non-osteogenic phenotype, with very low Osterix and Runx2 expression. Sox9 and Nestin expression
Figure 5.23: QRT-PCR analysis of Day 6 FACS sorted and replated Twist-EGFP populations at D0, D14, and D21 replate-OC. Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7, displayed relative to Control D0 OC. D0 Osterix and Runx2 data represent the average of two sort experiments. Insufficient sample was collected for D14 – sort analysis. Control = unmanipulated parallel control osteoculture from which the sort culture was derived.
Figure 5.24: D6 Twist2-EGFP positive sort colony morphology. Phase contrast microscopy following 3 specific colonies at D7, D14, and D21 replate-OC at 4X.
remain high, again perhaps due to a competing lineage (Figure 5.23).

By D21 replate-OC, the morphology of the D6 positive sort population suggests a non-osteogenic colony with little mineral deposition (Figure 5.24). Von Kossa staining of D6 positive sort culture plates also demonstrates the relative lack of mineralization within a majority of the colonies with an overall percent mineralization of the culture at 28% in comparison to the unmanipulated culture at 42%. The D6 positive sort plates had an average of 13.6 colonies per plate while the negative and unsort populations were only able to form an average of 1-2 colonies per plate (Figure 5.18).

Gene expression analysis corroborates the apparent lack of osteogenesis within the D6 sorted populations at D21 replate-OC, with Osteocalcin and Type I collagen expression levels that are relatively low. Expression levels of Myh11, Lpl, and CD45 were all similar or lower in the positive sort population as compared to the negative sort cells, and both sorted populations yielded lower levels of gene expression characteristic of the competing lineages than the parallel control culture (Figure 5.23).

We hypothesized that the failure of the D6 Twist2-EGFP positive sort colonies to mature into mineralized nodules was due to lower growth factor concentrations resulting from the low colony forming efficiency. Thus, D6 Twist2-EGFP positive sort cells were replated at an increased density, at $1 \times 10^5$ cells/10 cm dish compared to $1 \times 10^4$ cells/10 cm dish, to concentrate trophic factors. Increased density caused the cultures to grow at a faster rate, with D14 cultures plated at $1 \times 10^5$ cells/10 cm dish resembling D21 cultures plated at $1 \times 10^4$ cells/10 cm dish, however Osteocalcin mRNA expression suggested no change in the ability of these colonies to form true osteogenic nodules (Figure 5.25). Lpl mRNA expression suggested few adipose cells within both the positive and negative sort
**Figure 5.25**: QRT-PCR analysis of Day 6 FACS sorted and densely replated Twist-EGFP populations. D6 cultures were replated at $1 \times 10^5$ cells/10 cm dish, cultures were harvested at D14 replate-OC due to dense colony formation. Control cultures were plated at $1 \times 10^4$ cells/10 cm dish and harvested at D21 OC. Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to *ribosomal protein L7*, displayed relative to Control D0 OC. Control = unmanipulated parallel control osteoculture from which the sort culture was derived.
populations, and *Myh11* mRNA expression was also relatively low in the positive sort population. This same density paradigm was tested on the D4 sort populations and while at D7 post replating there is low expression of the osteogenic, myogenic, and adipogenic lineages, by D14 it is apparent that the culture is dominated by the myogenic lineage (Figure 5.26).
Figure 5.26: QRT-PCR analysis of Day 4 FACS sorted and densely replated Twist-EGFP populations. D6 cultures were replated at $1 \times 10^5$ cells/10 cm dish, cultures were harvested at D14 replate-OC due to dense colony formation, an intermediate timepoint of D7 replate-OC was also analyzed. Control cultures were plated at $1 \times 10^4$ cells/10 cm dish and harvested at D21 OC. Each assay was run in triplicate at two different template concentrations. Relative mRNA expression was normalized to ribosomal protein L7, displayed relative to Control D0 OC. Control = unmanipulated parallel control osteoculture from which the sort culture was derived.
Table 5.1 Summary of QRT-PCR timepoints performed on D2-D8 Sort Populations

<table>
<thead>
<tr>
<th></th>
<th>D0 Replate-OC</th>
<th>D14 replate-OC</th>
<th>D21 replate-OC</th>
<th>D7 replate-OC (dense)</th>
<th>D14 replate-OC (dense)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2</td>
<td>+</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- cont</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>+</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>- cont</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>D6</td>
<td>+</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>- cont</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>D8</td>
<td>+</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>- cont</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n/a = insufficient sample population was obtained
5.3 Discussion

In this chapter we have tested the function of a series of gene-specific EGFP reporters and their ability to enrich for osteoprogenitors as determined by gene expression and plating in osteogenic culture conditions. In the creation of these reporters, there were multiple factors influencing prioritization including: the stage at which the gene is first expressed, the fold-increase in expression seen in the osteogenic cultures, the specificity of the marker for the osteogenic lineage, as well as the availability of a BAC clone carrying the locus. As might be expected, these factors impacted the outcome of the experiments: the large boost in Twist2 expression allowed relatively straightforward separation of GFP positive and negative populations; Osterix, with relatively low expression and a more modest boost made GFP detection difficult; and Sox9 and Runx2 evaluation were limited by the availability of BAC clones – with a Sox9 clone too large in size to allow for targeting in the ESC, and a Runx2 clone unavailable that contained the entire coding sequence. Thus two other earlier, and perhaps riskier, reporters were obtained and generated, Bry-GFP and Nestin GFP respectively. The risky aspect of Bry-GFP being the liability of the hemizygous knock-in and of Nestin-GFP being regulatory sequences that had not been previously used for this lineage. Both of these genes have documented relevance to the osteogenic lineage, albeit somewhat earlier than what might be useful.

The Brachyury-GFP knock-in ESC line was previously created by another laboratory. This reporter was originally used to identify hematopoietic precursors arising from mesodermal differentiation within embryoid bodies (162). When applied to the
osteogenic culture, this reporter cell line showed poor colony forming efficiency with sorts at D4 and D6 OC. Previous experiments in our laboratory have attempted to enrich for an osteogenic population by sorting for Bry-GFP positive cells from different day EBs, however sorts at D2 EB resulted in an almost complete lack of colony formation from the positive sort population, with numerous colonies forming from the negative sort population. These results demonstrate that sorting for Brachyury in the D2 EB is either too early or too late of a timepoint to capture the mesodermal osteogenic population, or perhaps that a more moderate level of Brachyury expression is most indicative of the mesodermal lineages. In the osteogenic culture, Brachyury has a boost in expression at D4 and D6 OC, however sorting at these timepoints resulted in poor colony forming efficiency. Minimal colony formation was observed though after sorting at the intermediate timepoint of D4 OC. Sorting for Bry-GFP at D2 OC was not attempted because Brachyury expression is relatively low at D2 OC and because experiments with other reporter lines have revealed challenges with attempting to replate from a D2 OC. Alternatively, as Brachyury expression increases with increasing days of EB differentiation, later day EB populations could be sorted for Bry-GFP fluorescence. One caveat to this approach is that later day EBs have been shown to have decreased colony formation in the osteogenic culture (116;117). Finally it is important to note that the Brachyury-GFP ESC reporter line is a knock-in construct, therefore resulting in hemizygous expression of Brachyury, and a Brachyury haploinsufficient phenotype of tail shortening and mild skeletal defects has been observed in mouse models (256;257), suggesting the possibility for a negative impact on in vitro osteogenesis experiments such as these.
While both the targeted Nestin-EGFP reporter and the randomly integrated Nestin-EGFP reporter lines demonstrated the integration of the construct and RT-PCR analysis revealed the presence of Nestin-EGFP mRNA within the culture, Nestin-EGFP fluorescence was not detected in either line. These reporter lines were derived from a small Nestin promoter construct designed by Mignone et al. originally used to identify neural stem cell populations (211). The original Nestin construct contains a 5.8 kb fragment of the rat Nestin promoter, an EGFP cDNA cassette, and a 1.8 kb fragment of the rat Nestin second intron. While the Nestin promoter is not specific for neural stem cells, it has been demonstrated that the rat Nestin second intron directs reporter activity exclusively to the central nervous system (CNS) (211;245). Deletion studies of the rat Nestin second intron by Yaworsky et al. revealed the presence of a CNS enhancer element and a midbrain-specific enhancer in the 3’ region of the intron. Using a minimal promoter with the 1.8 kb rat Nestin second intron resulted in β-galactosidase activity restricted to cells of the developing nervous system; the midbrain enhancer allowed for expression in the ventral midbrain progenitor cells, and the CNS enhancer drove expression in specific regions of the hindbrain and midbrain (244). This Nestin construct was employed to potentially identify neural crest derived osteogenic progenitors, however, this CNS specificity appears to produce a reporter which is not appropriate for use in the osteogenic culture system. Our experience with this reporter construct underscores the advantages of BAC-based constructs with the potential for gene expression under near endogenous context versus promoter fragments which risk partial or restricted expression of the gene reporter.
The Hprt-Osx-EGFP reporter was tested in five BK4 ESC clones which did not contain constitutive CAGGS-EYFP expression. While DNA analysis demonstrated correct integration of the Osx-EGFP reporter construct, and RT-PCR analysis revealed the presence of transgenic Osx-EGFP mRNA within the osteoculture, no Osx-EGFP fluorescence was detected by flow cytometry. As Osterix expression within the osteogenic culture has only a 10-fold up-regulation relative to D0 OC, it could be that the level of Osx-EGFP expression is too low for efficient detection. It is also possible that while there is a 10-fold boost in expression relative to D0 OC, the base-level at D0 OC itself may be relatively low, making the 10-fold increase still beyond detection. Methods to increase Osterix expression within the osteogenic culture might enable detection of Osx-EGFP. It has been shown that TNF-α regulates Osterix expression by inhibiting the Osx promoter; treatment of cells with a MEK1 inhibitor can diminish TNF-α inhibition of the Osx promoter, allowing for a boost in expression (258). Alternatively, treatment with BMP2 or BMP2 + IGF1 has been shown to significantly increase Osterix expression in hMSCs (259). However, these experiments are beyond the scope of this thesis.

The most promising results were achieved using the dual reporter Twist2-EGFP/EYFP ESC line. The Twist2-EGFP reporter fluorescent expression profile paralleled Twist2 mRNA expression (Figure 5.13 and 5.15), and in sorted populations from the osteogenic culture the Twist2-EGFP mRNA and the endogenous Twist2 mRNA displayed the same patterns (Figure 5.15). One unexpected result was the observation that both the positive and negative sort populations had similar levels of Twist2/Twist2-EGFP mRNA expression. One might expect a higher expression of Twist2 mRNA in the positive sort population as compared to the negative sort population if the Twist2-EGFP
reporter was enriching for a Twist2 fluorescing population by FACS. However, because of the background levels of EYFP fluorescence, not all Twist2-EGFP fluorescing cells were selected into the positive sort population, resulting in cells expressing Twist2 mRNA within the negative sort population. An experiment was designed to test the fate of low Twist2-EGFP fluorescing cells in the negative sort population. These cells in the early stages of Twist2-EGFP fluorescence were given time to further differentiate, and progressed to higher expression of Twist2-EGFP (Figure 5.17). It is hypothesized that because of a lag between mRNA transcription and protein translation, a substantial amount of Twist2 message is present in the Twist2-EGFP low population, (see Figure 5.27). In this paradigm, the negative sort population contains cells with little to low EGFP fluorescence, yet a population of low and a population of high Twist2 mRNA expressors. Conversely, the positive sort population contains cells with high EGFP fluorescence, that contains both high and low Twist2 mRNA expressors. Therefore, sorting based on Twist2-EGFP fluorescence can yield positive and negative sort populations with similar Twist2 mRNA levels. This was demonstrated by broadening the sort gate for Twist2-EGFP fluorescence. By including high and low Twist2-EGFP fluorescent cells in the positive sort gate, most cells with high Twist2 mRNA expression are now included in the positive sort population (Figure 5.22).

When assessing the osteogenic enrichment potential of the Twist2-EGFP reporter, D2 OC Twist2-EGFP positive populations showed poor osteogenic enrichment and D8 OC showed poor colony formation efficiency in all three populations (Twist2-EGFP positive, negative, and unsorted control). D4 positive sort cultures displayed the best
Figure 5.27: Schematic to explain similar Twist2 mRNA levels in the positive and negative FACS sort populations. In this paradigm, Twist2-EGFP protein translation lags behind Twist2 mRNA transcription producing distinct populations of cells. The solid blue bars represent low mRNA expression/low EGFP fluorescence. The gray striped bars represent high mRNA expression and are separated into low EGFP fluorescence (blue striped) and high EGFP fluorescence (yellow striped) populations. The solid yellow bars represent high EGFP fluorescence/low mRNA expression. Sorting for the Twist2-EGFP cells with high EGFP fluorescence subsequently results in both the positive and negative populations containing cells expressing high levels of Twist2 mRNA. By removing the constitutive CAGGS-EYFP background fluorescence, Twist2-EGFP fluorescing cells with low/moderate fluorescence would be able to be identified, therefore allowing a majority of the Twist2 mRNA expressing cells to fall within the positive sort gate.
colony formation efficiency and also revealed enrichment for osteogenic progenitors as Runx2 and Osterix mRNA expression levels were higher in the positive sort population as compared to the negative sort population. However, at D21-replate OC mineralization in the osteocultures from the positive and negative sort populations were not different, and the Twist2-EGFP positive population contained high mRNA expression of myogenic and adipogenic markers. A broader gate for Twist2-EGFP expression may allow for a greater distinction between the colony formation of the positive and negative sort populations. However, lack of lineage specificity is still an issue with this timepoint.

Though D6 sort cultures had a lower colony forming efficiency than the D4 sort cultures, the early osteogenic morphology of the colonies and early gene expression characteristic of osteogenic progenitors in the Twist2-EGFP positive population makes this timepoint more promising for osteogenic progenitor enrichment. High expression of both Osterix and Runx2 were observed on the day of the sort. By D7 replate-OC, the positive sort population yielded colonies of characteristic osteogenic morphology. However, by D14 replate-OC, the osteogenic morphology was lost and colonies failed to focus matrix deposition into the classical nodule structure. Gene expression analysis at D14 replate-OC also revealed a loss of Osterix and Runx2 expression. By D21 replate-OC, colonies completely lacked mineralization and mRNA analysis revealed low levels of Osteocalcin expression in comparison to parallel unmanipulated cultures.

It was apparent that the D6 Twist2-EGFP positive sort population contained an osteogenic progenitor that was able to initiate osteogenesis, yet was lacking an integral signal or interaction for the progression into a mature osteogenic colony. To concentrate secreted trophic factors, the density of the replate culture was increased. The increased
density caused the cultures to grow at a faster rate, an outcome also seen in primary osteocultures plated at higher densities (section 5.2.4.1). However, osteogenesis in the replate culture was not enhanced. Competing smooth muscle and adipose lineages also remained low in the culture suggesting the dominance of a fibroblastic population. Overall, the results suggest that sorting at D6 is enriching for a relatively pure osteogenic population (based on gene expression in the positive sort population and early morphology), yet osteogenesis does not progress. Plating sorted cells at higher densities would increase the concentration of autologous growth factors but not those normally provided by competing lineages. Thus, signals from cells of competing lineages may be necessary for the maturation of these osteoprogenitors. If so, treating sorted cultures with conditioned medium from parallel ESC-derived heterogeneous osteogenic cultures could provide the necessary requirements, and subsequent treatment with exogenous trophic factors could define the critical components (see section 6.3.2). Alternatively, plating sorted cell populations on pre-formed extracellular matrix or an osteoconductive scaffold could permit the desired osteogenic maturation (see section 6.3.1).

*Sox9* has been shown to mark a bipotential progenitor of osteogenesis and chondrogenesis. While a Sox9 BAC was modified to contain *Hprt* homology and EGFP recombineered, the Hprt-Sox9-EGFP BAC reporter construct failed to integrate in its entirety into the ESC. The Sox9 BAC was chosen with a large upstream sequence because *Sox9* has been shown to be regulated by upstream distal cis-acting regulatory elements (208-210). Thus the Hprt-Sox9-EGFP BAC was large in size (197 kb), and this large size could be responsible for the low targeting efficiency. Runx2 has been shown to be both necessary and sufficient for osteoblast differentiation, however a Runx2-EGFP
reporter was not tested because a BAC which contained the entire coding sequence was not available. *Runx2* remains a very viable candidate for a transcriptional reporter that could achieve enrichment for a therapeutically useful osteoprogenitor.
CHAPTER 6
GENERAL DISCUSSION

6.1 Overview

The differentiation potential of ESCs and their ability to generate a large population of cells suggests the usefulness of embryonic stem cell based therapies. However, the heterogeneity of differentiation cultures as well as the existence of teratogenic progenitors within early stem cell populations demonstrates the need to isolate committed tissue-specific stem cells from ESC populations. Also, for a complete understanding of ESC derived tissue upon transfer to a host, it is imperative to monitor homing and engraftment of the transplanted cells for reasons of efficacy and safety.

This report describes the generation of gene-specific EGFP reporters within a constitutively expressing EYFP ESC line. The BAC-based reporter approach is advantageous as it can allow accurate representation of endogenous expression. Small promoter fragments taken out of context can result in incomplete, restrictive reporter expression patterns, for example see refs (169-172). The large size of BACs, ~200 kb, provides ample surrounding genomic sequence, increasing the likelihood that the reporter will faithfully represent the expression of the endogenous gene. Targeting a single copy of the BAC-based reporter to a chosen site of integration removes the variables of position effect and copy number, yet leaves both endogenous copies of the gene unaltered.
To identify an osteogenic progenitor with therapeutic potential that contains the desired characteristics of lineage specificity as well as proliferative capacity, we focused on early markers of osteogenesis. Twist2 has been shown to be a negative regulator of Runx2 function by interacting with the DNA binding domain of Runx2 (32), and controlling the developmental timing of osteogenesis. Twist2 expression at this critical time in osteogenesis, and the significant boost in Twist2 mRNA level that occurs in the ESC-derived osteogenic cultures, make it an important candidate for a transcriptional reporter of the osteoprogenitor. QRT-PCR analysis of mRNA from FACS sorted Twist2-EGFP osteogenic cultures showed higher expression of osteogenic markers within the positive sort cells as compared to the negative sort, demonstrating the enrichment for an osteogenic population.

Because the Twist2 reporter was integrated into a CAGGS-EYFP ESC line, the Twist2-EGFP positive sorted population can also be followed in vivo. Future studies can not only assess transplanted cell contribution to osteogenesis within three-dimensional scaffolds, but also the engraftment of cells transplanted intravenously or intrafemorally. Importantly, engraftment at other sites throughout the organism can also be monitored. One caveat of the dual reporter system is that background YFP fluorescence complicates the separation of EGFP positive and EGFP negative populations.
6.2 Methodological considerations

6.2.1 Targeting efficiency

With these experiments, one major difficulty was targeting large BAC constructs to the \textit{Hprt} locus. Ideally, BACs were chosen with the gene of interest centered in the BAC genomic sequence and surrounded by large areas of flanking sequence. These flanking sequences allow for the expression of the gene in a near endogenous context. However these results, as well as others (176), report limited success in the integration of large BACs. The low efficiency of integration of large BACs can be attributed to many factors such as the increased susceptibility of large BACs to rearrangements and internal deletions, high frequency terminal end breakage, and the ratio of homology arm size-to-length of non-homologous intervening sequence (207;212-214;217). Many of these issues persist in the context of random integration of large genomic fragments.

The ratio of homology arm size-to-length of intervening sequence may be the determining factor for the low recombination efficiency of large BACs. One way to address this issue is to limit the size of heterologous intervening sequence – to only choose BACs of limited size. Targeting has been successful with BACs ranging in size from 114 – 178 kb, however BACs ~190 kb and greater have yielded mixed results. Obtaining a BAC with a centered gene of interest, ample flanking sequence, and limited size, restricts the availability of BACs within the existing public BAC libraries. For example, in the UCSC Mouse Genome Browser, the complete Twist2 gene is available on 8 different BACs. However, of these 8 BACs, 4 BACs have severely limited 5’ or 3’ flanking sequences (as low as ~10 kb). Of the 4 remaining available BACs, 3 would
result in a total BAC size greater than 200 kb after modification. Subsequently, only one Twist2 BAC was appropriate for our purpose, resulting in the modified 186 kb BAC used in these experiments.

One method to obtain manageable BACs is through the process of BAC trimming. BAC trimming is a method of recombineering to remove unwanted genes from a BAC, or to decrease the overall size of a BAC. In this system, PCR is used to amplify an antibiotic resistance cassette with primers containing homology arms to the BAC. The homology arms on the selection cassette can be homologous to regions of the BAC with up to 100 kb of intervening sequence. Using the homologous recombination machinery in E. coli, a replacement event occurs exchanging the intervening sequence with the selection cassette, and thus a deletion in the resultant BAC (217;260;261) (Figure 6.1). BAC trimming could be applied as an initial BAC modification step to create appropriately sized BACs for subsequent use in the creation of gene-specific GFP reporters. The BAC trimming step need not interfere with the loxP site located on the BAC vector backbone, therefore still allowing for the Cre-mediated Hprt modification of the BAC. A second step of BAC recombineering would then integrate the IRES-EGFP cassette into the gene of interest. The feasibility of this approach is supported by the fact that many protocols employ two rounds or more of BAC modification by recombineering (217;261).

Another approach to making the homology arm-to-intervening sequence ratio more favorable is obviously to increase the size of the homology arms. Frequency of homologous recombination in ESCs has been shown to be dependent on length of homology between the targeting vector and the target locus (127;220;222;262).
Figure 6.1: BAC trimming. Amp and Gent represent the two PCR products used in the two trimming rounds, conferring resistance to ampicillin and gentamicin. The oligonucleotide homology arms A, B, C, and D were taken from positions as indicated on the wt BAC. A and D flank the BAC vector. The regions replaced by the Amp and Gent PCR products are depicted on the wt BAC in light gray; the remaining BAC region after trimming is in black. The oligonucleotides for the trimming steps each had 63 nucleotides of homology to the BAC. Figure and legend adapted from (217).
Experiments by Deng and Capecchi show an exponential relationship between homology length and targeting frequency. Total homology (both arms) of 10 kb vs 2 kb yielded a 100-fold increase in targeting frequency (222). Total homology of 14 kb was shown to saturate the recombination system, while the lower size limit for one arm was reported at 1 kb. Again it is important to note that these results were based on smaller plasmid DNA constructs and not BAC based reporters, however the message of the importance of homology arm size should still be recognized.

While our lab has been able to target 114 kb – 186 kb BACs to the Hprt locus using homology arms which were 3 and 3.8 kb in size (for a total homology of 6.8 kb), perhaps these homologous sequences are too small to efficiently allow the recombination of the larger sized BACs. Making changes to the homology arm size in the Hprt targeting construct could potentially increase the efficiency of the Twist2- and Sox9-BAC electroporations. Anecdotally, the on-line guidelines for design and assembly of BAC targeting vectors by the Vanderbilt Center for Stem Cell Biology now highly recommend, when designing homology arms, that the long arm be at least 5 kb in length and the short arm at least 2 kb, with an absolute minimum size of 4 kb and 1 kb for the long and short arms (263).

6.2.2 Dual reporter system

The dual EYFP/EGFP reporter system is dependent on the ability to distinguish individual emission maxima that are separated by less than 20 nm (EGFP at 510 nm and EYFP at 527 nm). The simultaneous detection of these two emission peaks can be separated when excited with a standard argon laser tuned to 488 nm, however a complex
filter configuration must be employed. In this system, a 510/20 nm bandpass filter allows for EGFP detection, a 550/30 nm bandpass filter is used for EYFP detection, and a 525 nm short pass dichroic mirror is necessary to separate the two signals. Applying this filter configuration allows for the simultaneous detection of EGFP and EYFP, identifying populations that express EGFP only, EYFP only, and cells that co-express both proteins (264;265).

For these experiments, a 530/40 nm bandpass filter was used to detect the fluorescent protein variants. For the dually expressing Twist2-ESC line, fluorescence of the single labeled CAGGS-EYFP ESC culture was set as background, and gates were set to choose cells with a shift in EGFP fluorescence greater than background. This approach allowed only for the enrichment of cells expressing the most robust levels of EGFP, however it did not fully identify the complete EGFP/EYFP co-expressing population. Therefore, a significant population of Twist2-EGFP positive cells was left behind in the negative population.

Ideally, the more complex filter configuration would allow for a more distinct separation of the single EYFP expressing (negative sort) population from the dual expressing EYFP/EGFP (positive sort) population within the Twist2-EGFP reporter line. Alternatively, initial in-vitro characterization of the Twist2-EGFP reporter, or other osteogenic reporters, could take place in a non-EYFP fluorescent host ESC line. Upon determination of a transcriptional reporter and timepoint that identified a potentially therapeutically useful progenitor, the reporter could then be moved into the CAGGS-EYFP background ESC line for further in-vivo studies.
6.3 Enhancement of the osteoculture

6.3.1 Bio-adhesive surfaces

Within the current methods of this research, ESC-derived osteogenic cultures at varying days of differentiation are trypsinized to single cell suspensions, sorted by FACS, and replated in osteogenic medium. The replated cells are then assessed for their potential as osteogenic progenitors. However, these osteoprogenitor cells are disrupted and removed from the original osteoculture consisting of a rich extracellular matrix and subsequently transferred in relatively low density to tissue culture polystyrene. As it has been shown that interactions with the extracellular matrix can affect osteoblast survival, proliferation, differentiation, and matrix mineralization (266), it is easily hypothesized that the addition of bio-adhesive surfaces to the tissue culture plates could foster the adhesion and differentiation of the replated cells. Cell culture on substrates coated with the extracellular matrix components Fibronectin or Type I collagen, have resulted in increased expression of osteogenic genes, increased Alkaline phosphatase activity, and increased commitment to the osteogenic lineage (266-269). Therefore, future experiments employing the use of matrix coated tissue culture plastic could more clearly elucidate the osteogenic potential of the EGFP sorted progenitor cells.

Osteoconductive scaffolds have also been developed to encourage osteogenic differentiation. Synthetic scaffolds can promote bone formation by providing cell anchorage sites, mechanical stability, and structural guidance (270). Osteogenic differentiation of ESCs has been demonstrated on 3D scaffolds consisting of D,DL,L
polylactic acid, poly(D,L-lactic-co-glycolic acid)/hydroxyapatite, alginate hydrogels, synthetic nanofibers, and ESC derived cartilage matrix (271-275).

6.3.2 Trophic factors

In an effort to enhance osteogenic commitment and differentiation upon replating, sorted cultures could potentially be treated with osteoinductive trophic factors. Bone morphogenic proteins (BMPs), growth factors belonging to the transforming growth factor β (TGF-β) superfamily, have been demonstrated to promote osteogenic differentiation. Treatment with BMP-2 of MSCs, C2C12 myogenic cells, and C3H10T1/2 non-osteogenic cells, have all demonstrated an enhanced commitment to the osteogenic lineage (276-279). ESCs cultured with BMP-2 have shown an increase in matrix deposition and upregulation of Osteocalcin and Alkaline phosphatase expression (108). Similarly, composite scaffolds have been treated with BMP-2 for the in-vivo culture of ESC derived osteogenic progenitors resulting in enhanced formation of new bone as compared to the untreated scaffolds (274). BMP-2’s osteoinductive effects are mediated through Smad signaling, most notably Smad1 and Smad5, and the upregulation of Runx2 (276;280-282). Other factors such as Insulin-like growth factor-1, Melatonin, and Interleukin-6-type cytokines have all also been shown to promote osteogenic differentiation (108;283-285).
6.4 Alternative avenues

6.4.1 Lineage selection

As lineage directed differentiation of ESC cultures is not 100% efficient, cultures result in heterogeneous populations containing progenitors of multiple cell types. Therefore, this report has developed an approach to create gene-specific reporters to identify and enrich for lineage specific progenitors to use for therapeutic purposes. The methods described in this report are based on the paradigm of creating a ‘pure’ population by isolating the cells of interest from a mixed population, however an alternate approach can be applied which only allows for the survival of the lineage of interest.

Genetic selection can be applied to produce a relatively pure population of lineage restricted precursors from a heterogeneous culture. Constructs have been created which contain the promoter of a lineage specific gene (α-cardiac myosin heavy chain (MHC) promoter for the cardiomyocyte lineage (286), Sox2 for the neural lineage (287)) and cDNA encoding Aminoglycoside phosphotransferase conferring resistance to G418 selection. The transgene also contained hygromycin resistance for original selection upon electroporation of the construct into ESCs. Upon identification of hygromycin-resistant ESC lines, differentiation of the transgenic-ESCs was induced by the removal of LIF. After 8 days of differentiation, G418 selection was added to the cultures therefore eliminating α-cardiac MHC-negative non-cardiac cells. While original differentiation conditions produced cultures which contained less than 1% of cardiomyocytes, G418-selected cultures contained greater than 99% cardiomyocytes. Further, cardiomyocytes
from the G418-treated cultures were transplanted into the left ventricle of mice and were able to successfully engraft (286).

This approach introduces a simple method to generate relatively pure populations by removing competing lineages from the culture, and given the observed sensitivity of the osteoprogenitors to disruption and replating, may be a more viable alternative. It is even hypothesized that combinatorial approaches using multiple promoters could be applied to isolate a specific lineage (286). Applying this approach using genes such as Twist2, Runx2, and Osterix, could potentially allow for the production of a pure in vitro osteogenic population.

6.4.2 Existing constructs

An alternative approach in attempting to identify osteogenic progenitors from an ESC derived culture would be to use osteogenic GFP reporters which have already been created. The International Gene-Trap Consortium and the Gene Expression Nervous System Atlas (GENSAT) BAC Transgenic Project have GFP reporters that could be of potential use.

The Conventional Gene-Trap vector system is generally applied as a mutagenesis/gene knock-out tool. The system generates random loss-of-function mutations while reporting gene expression. This is accomplished by the introduction of a gene trap vector which contains a splice acceptor, a GFP reporter, and a polyadenylation signal. The splice acceptor ‘traps’ upstream exons allowing GFP to report endogenous expression of the trapped gene. The polyA signal placed immediately downstream of the GFP sequence results in a truncation of the fusion protein (Figure 6.2A). This system
**Figure 6.2:** Gene trapping. (A.) Conventional gene trap vectors use a splice acceptor to take advantage of endogenous transcription and truncate the mRNA, leaving the gene 5' of the insertion site intact, followed by the vector sequence containing the selection/reporter construct. A polyA signal is placed at the 3' end of standard vectors, causing translation to end and producing a truncated fusion protein. (B.) PolyA gene trap vectors contain a promoter signal and a transcriptional start site, allowing genes to be trapped that are not normally expressed, or are expressed at very low levels under experimental conditions. A splice donor sequence is present at the end of the gene trap cassette, causing the mRNA product of the vector construct to be fused with any downstream exons. Since these vectors do not have their own polyA sequences to signal the end of translation, only cell lines in which the vector inserts upstream of a terminal exon will produce the selectable reporter tag. Figure and legend taken from (288). Note: The Twist2-VenusGFP Gene-Trap vector by Tanaka et al. (section 1.3.4.3) is a hybrid of the Conventional and PolyA systems, containing the splice acceptor of the Conventional vector along with the splice donor of the PolyA vector, therefore resulting in a GFP fusion protein under control of the endogenous promoter.
reports endogenous gene expression, but also results in hemizygous expression of the gene (288).

A second category of Gene-Trap vectors have also been developed which are termed PolyA gene trap vectors. These vectors randomly integrate into the intron of a gene and contain an exogenous constitutive promoter to drive transcription of the downstream exons (Figure 6.2B). Therefore, these gene trap vectors allow for the reporting of genes which are not normally expressed or expressed at very low levels. These vectors also result in hemizygous expression of the gene (288).

The International Gene Trap Consortium has available both a Twist2 and a Runx2 reporter ESC line which have been created with the Conventional Gene-Trap vector system (for more information on the Twist2 Gene-Trap reporter see section 1.3.4.3). Ideally, the Conventional vector system can allow for correct spatiotemporal expression of the gene, however there are also many limitations to this system. By trapping the endogenous gene to create a reporter fusion protein, the Gene-Trap system results in hemizygous expression of the endogenous gene. Both Twist and Runx2 have been shown to display defects in osteoblast development upon deficiencies in gene expression (27;31;32;175). Also, a potential problem with the random integration of the Gene-Trap vector is that it does not always result in a knock-out of the endogenous gene, but can result in a dominant negative or neomorphic mutation. Therefore, Gene-Trap reporters must be fully characterized. The international Gene Trap Consortium also has reporter ESC lines available for Nestin and Coll1a1, however these reporters are documented as being created with the PolyA constitutive vectors, or no vector information is provided.
The Gene Expression Nervous System Atlas BAC Transgenic Project has created GFP-BAC reporters in an attempt to analyze development of the central nervous system (176). Two genes which have been used to create transgenic lines are Sox9 and Col1a1, however published data has only focused on expression within the central nervous system. To use these reporters in an in vitro ESC osteogenic culture system, ESC lines would first have to be derived from the ICM of transgenic blastocyst stage embryos. GENSAT BACs are modified to replace the endogenous gene with an EGFP reporter cassette, therefore reporter expression remains under control of the endogenous promoter and surrounding enhancer sequences. However, the modified BAC is randomly integrated by pronuclear injection of fertilized oocytes. Random integration of the BAC introduces variability due to copy number and position effects. Therefore, transgenic lines from which the ESCs would be derived should undergo extensive characterization to ensure expression is not altered due to random integration of the BAC.
6.5 Importance of lineage specificity and stem cell therapy

Because ESCs have the ability to differentiate into the osteogenic lineage, the possibility of using ESC derived tissue for transplantation and therapy has become a major focus of study. Multiple musculoskeletal afflictions have the potential to be treated with ESC derived osteogenic progenitors including severe fractures, resection of cancerous bone, and the brittle bone disease osteogenesis imperfecta (OI). The concept of using ESCs for regenerative medicine is particularly appealing as ESCs have the ability to self renew for prolonged periods without undergoing senescence, allowing for the generation of an almost infinite source of therapeutic cells. Differentiated ESCs for transplantation can also be engineered to be histocompatible with the recipient by somatic cell nuclear transfer. However, the heterogeneity of ESC cultures and the formation of teratomas in vivo present significant hurdles. To utilize ESC derived progenitors for transplantation, a highly efficient method allowing for the selection of pure, lineage-specific cells must be employed. Also, while it is necessary to have lineage specificity of transplanted cells, it is also essential to have the ability to track the fate, homing specificity and engraftment efficiency of the ESC-derived cells upon transplantation.

The ability to design lineage specific differentiation protocols and effectively test the ESC-derived cells will allow stem cell therapy to become a large scale reality. In the context of osteogenic differentiation, this thesis has begun to characterize a Twist2-EGFP reporter, as well as other potential osteogenic markers including Brachyury, Sox9, Osterix, and Nestin; and an excellent candidate, Runx2, awaits a suitable genomic fragment to harness for a transcriptional reporter. The gene-specific reporter/tracking
reporter paradigm, developed and tested in this thesis, can be applied to any lineage-specific stem cell from differentiating ESC cultures.
Reference List


71. Haynesworth SE, Baber MA, Caplan AI. Cytokine expression by human marrow-
derived mesenchymal progenitor cells in vitro: effects of dexamethasone and IL-1 

72. Long MW, Robinson JA, Ashcraft EA, Mann KG. Regulation of human bone 
marrow-derived osteoprogenitor cells by osteogenic growth factors. J.Clin.Invest 

73. Rickard DJ, Kassem M, Hefferan TE, Sarkar G, Spelsberg TC, Riggs BL. 
Isolation and characterization of osteoblast precursor cells from human bone 

74. Rider DA, Nalathamby T, Nurcombe V, Cool SM. Selection using the alpha-1 
integrin (CD49a) enhances the multipotentiality of the mesenchymal stem cell 
population from heterogeneous bone marrow stromal cells. J.Mol.Histol. 

75. Kassem M, Abdallah BM, Saeed H. Osteoblastic cells: differentiation and trans-

76. Pelagiadis I, Dimitriou H, Kalmanti M. Biologic characteristics of mesenchymal 
stromal cells and their clinical applications in pediatric patients. 

77. Chapel A, Bertho JM, Bensidhoum M, Fouillard L, Young RG, Frick J et al. 
Mesenchymal stem cells home to injured tissues when co-infused with 
hematopoietic cells to treat a radiation-induced multi-organ failure syndrome. 

Intervertebral disc cell therapy for regeneration: mesenchymal stem cell 

Xenoreactivity and engraftment of human mesenchymal stem cells transplanted 

grade mesenchymal stem cells transdifferentiated under xenofree conditions 
2009.

81. Ohgushi H, Goldberg VM, Caplan AI. Repair of bone defects with marrow cells 


187. Living Colors User Manual. BD Biosciences Clontech . 2001. Ref Type: Electronic Citation


263. Design and Assembly of Targeting Vectors. Vandebilt Center for Stem Cell Biology. 6-15-2008. Ref Type: Electronic Citation


269. Stephansson SN, Byers BA, Garcia AJ. Enhanced expression of the osteoblastic phenotype on substrates that modulate fibronectin conformation and integrin receptor binding. Biomaterials 2002;23(12):2527-34.


Jean E. DeMarco
220 Short St. Apt. 1
Hummelstown, PA 17036 USA
Phone (717) 566-5582, Fax (717) 531-7667, email: jed24@psu.edu

EDUCATION:
2003-2009  The Pennsylvania State University College of Medicine, Hershey PA
            Ph.D. Graduate Degree Program in Cellular and Molecular Biology
1997-2001  St. Lawrence University, Canton NY
            B.S., Biology
            Minor in Literature

PROFESSIONAL EXPERIENCE:
2003-2009  Graduate Assistant
            Penn State College of Medicine, Hershey PA
            Sarah K. Bronson, Ph.D.
2001-2003  Research Technician
            Penn State College of Medicine, Hershey PA
            Teresa L. Wood, Ph.D.

AWARDS & SCHOLARSHIPS:
NIH Training Grant T32-GM08619 (2005-2006)
Beta Beta Beta Biological Honor Society (2001)
Irving Bachelor Literature Honor Society (2001)
Dean’s List

PUBLICATIONS & SELECTED ABSTRACTS:
DeMarco JE, Bronson SK. Dual Reporter ESCs for Gene-Specific Enrichment as well as Cell
Tracking. (in preparation).

Organogenesis. (in preparation).

Therapeutic Osteogenic Populations. (Oral Presentation) 20th Annual Graduate Student Research
Forum, Penn State College of Medicine.

DeMarco JE, Bronson SK (2007). Defining Embryonic Stem Cell Derived Therapeutic
Osteogenic Populations. (Poster) American Society of Cell Biology, Washington D.C.

Loladze AV, Stull MA, Rowzee AM, DeMarco JE, Lantry JH, Rosen CJ, Leroith D, Wagner KU,
Hennighausen L, and Wood TL (2006). Epithelial-Specific and Stage-Specific Functions of
Insulin-Like Growth Factor-I During Postnatal Mammary Development. Endocrinology 147(11):
5412-23.