IDENTIFICATION AND VALIDATION OF \textit{DLG2} AS A NOVEL TUMOR SUPPRESSOR IN OSTEOSARCOMA

by

Yang Washington Shao

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Medical Biophysics
University of Toronto

© Copyright by Yang W. Shao 2016
ABSTRACT

Biological events including mechanisms underlying tumorigenesis are frequently well conserved within mammals. We here report a cross-species comparative genomics study across human, dog, and mouse in osteosarcoma (OS). Genomic profiling in spontaneous human and dog OS elucidated a promising tumor suppressor gene candidate, *Disks Large Homolog 2 (DLG2)*. *DLG2* copy number losses are observed in 29% of spontaneous human OS and 45% of canine OS. The tumor suppressive ability of *DLG2* is then examined *in vitro* using human and dog cell lines, and *in vivo* using murine xenograft models. Restoration of *DLG2* expression selectively reduces cell motility in *DLG2*-deleted cell lines. *DLG2* expression also reduces tumorigenic capacity in multiple *DLG2*-deleted cell lines, as determined by soft agar assays and xenografts in NOD/SCID mice. Finally, osteoblast-specific knockout (KO) of *Dlg2* gene accelerates tumor progression in a murine model of OS. RNA-seq profiling shows that *DLG2* overexpression in *DLG2* KO cells affect a range of cellular activities with RhoA pathway and actin-organization at the top of the list. Therefore, by leveraging biological conversation across species, we identify *DLG2* as a novel tumor suppressor in OS. The principles of this project can be widely applied to
other human cancers, offering an approach to shortlist potential candidate genes within large chromosomal altered regions, and prioritize their functional validation.
ACKNOWLEDGMENTS

Foremost, I would like to thank my supervisor Dr. Rama Khokha for all your guidance, continuous support and encouragement during my Ph.D study. You have been a tremendous mentor for me with your immense knowledge, inspiration, enthusiasm, and criticism. I also thank you for the critical reading of this thesis.

I thank the Department of Medical Biophysics at University of Toronto for giving me the opportunity to study here. I greatly appreciate the insightful discussions and comments from all my supervisory committee members, Dr. Vuk Stambolic, Dr. Lakshmi Muthuswamy and Dr. Jason Moffatt. Your advice on both research, as well as on my career, has been invaluable.

This research project would not have been possible without the help of many people including the Wood lab from University of Guelph, and the Malkin Lab from University of Toronto.

I am thankful to all of the past and present members in the Khokha lab, Sam Molyneux, Marco DiGrappa, Yan Chen, Aditya Murthy, Jackson Hartland, Alison Aiken, and many others. The group has been a source of friendships, as well as good advice and collaboration. All of you made my study in this lab a cherished memory. I thank the PMH Flow Cytometry and Animal facility for their technical support, and our lab manager Iris Fang, for making the lab organized and functional. I am honored to have had the opportunity to supervise a talented and hard-working summer student, Chris Watling. I am grateful for all help in many aspects of my project.

Finally, I dedicate this thesis to my parents, Huawu Shao and Xiaolan Li, my sister, Lucy Shao, and my wife, Cherry Wang, for their constant support and unconditional love.
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS**  
iv

**TABLE OF CONTENTS**  
v

**LIST OF FIGURES**  
vii

**LIST OF TABLES**  
ix

**LIST OF ABBREVIATIONS**  
x

## Chapter 1  Introduction

1.1 Osteosarcoma  
1.1.1 Overview of Osteosarcoma  
1.1.2 Etiology of Osteosarcoma  
1.1.3 Metastasis of Osteosarcoma  
1.1.4 Models of Osteosarcoma  
1.2 Identifying cancer ‘driver’ genes using genomic tools  
1.2.1 Cancer Genome  
1.2.2 Array-based Comparative Genomic Hybridization  
1.2.3 Comparative Oncogenomics  
1.3 Rationale and hypothesis  

## Chapter 2  Materials and Methods

2.1 aCGH  
2.2 Cell culture  
2.2.1 Culture medium  
2.2.2 Lentivirus production and transduction  
2.2.3 Cell proliferation assay  
2.2.4 3D Matrigel colony formation assay  
2.2.5 Scratch-wound assay  
2.2.6 Transwell migration assay  
2.3 Mice  
2.3.1 Establishment of tumor cell line xenografts in immunodeficient mice  
2.3.2 Establish of Osteoblast-specific Dlg2 null mice  
2.4 Cloning
2.5 RNA-seq 43
2.6 Biochemical analysis 43
2.7 X-ray and MicroCT scan 45
2.8 Histology and immunohistochemistry 45
2.9 Cell cycle analysis by flow cytometry 45
2.10 Quantitative real time PCR 46
2.11 Gene expression microarray analysis 48
2.12 Statistics 48

Chapter 3 Results 50
3.1 Cross-species Oncogenomics Identifies DLG2 as a Tumor Suppressor Candidate in Osteosarcoma 51
3.1.1 Optimization of aCGH analysis 51
3.1.2 Cross-species oncogenomics analysis identify DLG2 gene 56
3.2 Functional Validation of DLG2 as a Tumor Suppressors in vitro and in vivo 64
3.2.1 Functional validation of DLG2 in OS cell lines 64
3.2.2 Conditional Dlg2-null mouse for studying of OS 82

Chapter 4 Discussion 90
4.1 Oncogenomics 91
4.2 Cross-species comparative oncogenomics 93
4.3 Tumor suppressor function of DLG2 in human cancer 94
4.4 Conditional Dlg2 KO mouse model 96

Chapter 5 Summary and Future Directions 97
5.1 Summary and Key Findings 98
5.2 Future Directions 99
5.2.1 Role of DLG2 as a scaffold protein in bone cells 99
5.2.2 The role of small GTPs in DLG2-deficiency derived osteosarcoma 100
5.2.3 Other candidate genes in significant changed CNVs identified by this study 100
5.3 Concluding Remarks 101

References 102
LIST OF FIGURES

| Figure 1-1. Human and canine ortholog genes .......................................................... 21 |
| Figure 1-2. Timeline for somatic mutations acquired by the cancer cells ....................... 23 |
| Figure 1-3 Comparative oncogenomics approach to identify candidate cancer driver genes ..... 31 |
| Figure 2-1. Targeting strategy for conditional Dlg2 null mouse model ............................. 39 |
| Figure 2-2. Construction of DLG2 expression vector ..................................................... 42 |
| Figure 2-3. Schematics of RNA-seq work flow on tumor samples .................................... 44 |
| Figure 2-4. Gating pathway used for cell cycle analysis by flow cytometry ....................... 47 |
| Figure 3-1. aCGH analysis pipeline .................................................................................. 54 |
| Figure 3-2. Test of different segmentation methods using a training dataset ....................... 55 |
| Figure 3-3. Whole genome landscape of CNAs in human and canine OS ............................ 57 |
| Figure 3-4. Identification of a putative tumor suppressor gene locus .................................. 58 |
| Figure 3-5. Top five mutated oncogenes and tumor suppressors in OS ............................... 60 |
| Figure 3-6. Identification of DLG2 as a tumor suppressor in OS ....................................... 61 |
| Figure 3-7. DLG2 in OS and other human cancers ............................................................ 62 |
| Figure 3-8. Selection of OS cell lines with DLG2 deletion .................................................. 68 |
| Figure 3-9. Lentivirus-based delivery of DLG2 re-expression in various cell lines ............... 69 |
| Figure 3-10. Effect of DLG2 re-expression on proliferation in normal and low FBS conditions 71 |
| Figure 3-11. DLG2 re-expression reduces colony formation under 3D-matrigel conditions ...... 72 |
| Figure 3-12. DLG2 re-expression reduces colony number in soft agar assay ....................... 73 |
| Figure 3-13. DLG2 re-expression reduces the tumor xenograft growth in vivo ..................... 74 |
| Figure 3-14. Intergrative analysis of aCGH and gene expression from canine tumors .......... 75 |
Figure 3-15. List of top differentially expressed genes between $DLG2$-null and $DLG2$-wildtype group .......................................................................................................................................................................................... 77

Figure 3-16. $DLG2$ re-expression cells show reduced motility in scratch-wound assay .......... 78

Figure 3-17. $DLG2$ re-expression cells show reduced motility in transwell migration assay ..... 79

Figure 3-18. Effect of $DLG2$ re-expression on cell cycle progression ............................................. 81

Figure 3-19. Osteoblast-specific $Dlg2$ KO accelerates OS tumor growth in mice ....................... 83

Figure 3-20. Whole body LacZ profiling of $Dlg2$ deleted, LacZ-expressing mice under Col1a-Cre promoter ........................................................................................................................................................................... 84

Figure 3-21. Ki-67 staining of mouse OS .............................................................................................. 87
LIST OF TABLES

Table 1-1. Clinical features of osteosarcoma................................................................. 3
Table 1-2. Familial syndromes predisposed to OS ......................................................... 6
Table 1-3. Mouse models of OS ..................................................................................... 18
Table 2-1. PCR primers for genotyping............................................................................. 40
Table 3-1. DLG2 mutations in human cancer................................................................. 63
Table 3-2. Properties of each cell lines in various in vitro and in vivo assays................. 67
Table 3-3. Top differentiated expressed genes by RNA-seq. .................................. 88
Table 3-4. Top altered gene ontology pathways identified by RNA-seq in Dlg2-/- tumor. 89
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td>anchorage-independent growth</td>
</tr>
<tr>
<td>AP-1</td>
<td>activator protein 1</td>
</tr>
<tr>
<td>AWS</td>
<td>adaptive weights smoothing</td>
</tr>
<tr>
<td>BS</td>
<td>Bloom syndrome</td>
</tr>
<tr>
<td>CaMP</td>
<td>Cancer Mutation Prevalence</td>
</tr>
<tr>
<td>Cat K</td>
<td>Cathepsin K</td>
</tr>
<tr>
<td>CBS</td>
<td>circular binary segmentation</td>
</tr>
<tr>
<td>CCRT</td>
<td>Col1a-Cre, Rb1fl/fl, Tp53 fl/fl</td>
</tr>
<tr>
<td>CDK</td>
<td>cyclin-dependent kinase</td>
</tr>
<tr>
<td>aCGH</td>
<td>array comparative genomic hybridization</td>
</tr>
<tr>
<td>CGH</td>
<td>comparative genomic hybridization</td>
</tr>
<tr>
<td>chapsyn-110</td>
<td>channel-associated protein of synapse-110</td>
</tr>
<tr>
<td>CIN</td>
<td>chromosomal instability</td>
</tr>
<tr>
<td>CNA</td>
<td>copy-number alterations</td>
</tr>
<tr>
<td>CNV</td>
<td>copy number variation</td>
</tr>
<tr>
<td>Col-1a</td>
<td>collagen type 1a</td>
</tr>
<tr>
<td>COSMIC</td>
<td>Catalogue of Somatic Mutations in Cancer</td>
</tr>
<tr>
<td>GARNL3</td>
<td>GTPase activating Rap/RanGAP domain-like 3</td>
</tr>
<tr>
<td>DLG2</td>
<td>Disks Large Homolog 2</td>
</tr>
<tr>
<td>DWD</td>
<td>distance weighted discrimination</td>
</tr>
<tr>
<td>DX</td>
<td>doxorubicin</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein Barr virus</td>
</tr>
</tbody>
</table>
ECM  extracellular matrix
EMT  Epithelial-Mesenchymal-Transition
ERM  Ezrin, Radixin and Moesin
ET-1  endothelin-1
FAK  focal adhesion kinase
FBS  fetal bovine serum
FISH  fluorescence in situ hybridization
GEMM  Genetically Engineering Mouse Models
GEO  Gene Expression Omnibus
GISTIC  Genomic Identification of Significant Targets in Cancer
GLAD  gain and loss of DNA
GK  guanylate kinase
HBV  hepatitis B virus
H&E  hematoxylin & eosin
HMM  hidden Markov model
HPV  human papilloma virus
IGV  Integrative Genomics Viewer
IMPC  International Mouse Phenotyping Consortium
LFS  Li-Fraumeni syndrome
IL  interleukin
LOH  loss of heterozygosity
IP  immunoprecipitation
MAGUK  membrane-associated guanylate kinases
MDM2  mouse double minute 2 homolog
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP</td>
<td>matrix metalloproteinase</td>
</tr>
<tr>
<td>MOTO</td>
<td>Murine Osteocalcin driven T-antigen Osteosarcoma</td>
</tr>
<tr>
<td>MTX</td>
<td>methotrexate</td>
</tr>
<tr>
<td>NFATc1</td>
<td>nuclear factor of activated T-cells</td>
</tr>
<tr>
<td>NOD-SCID</td>
<td>nonobese diabetic/severe combined immunodeficiency</td>
</tr>
<tr>
<td>ODN</td>
<td>oligodeoxynucleotide</td>
</tr>
<tr>
<td>OPG</td>
<td>osteoprotegerin</td>
</tr>
<tr>
<td>OS</td>
<td>osteosarcoma</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
</tr>
<tr>
<td>PCD</td>
<td>programmed cell death</td>
</tr>
<tr>
<td>PEDF</td>
<td>pigment epithelial-derived factor</td>
</tr>
<tr>
<td>PI3K</td>
<td>PI3 kinase</td>
</tr>
<tr>
<td>PKC</td>
<td>protein kinase C</td>
</tr>
<tr>
<td>PSD-93</td>
<td>postsynaptic density protein 93</td>
</tr>
<tr>
<td>PDZ</td>
<td>PSD95/DLG/ZO-1</td>
</tr>
<tr>
<td>PFA</td>
<td>paraformaldehyde</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative real time PCR</td>
</tr>
<tr>
<td>RANKL</td>
<td>receptor activator of nuclear factor κB ligand</td>
</tr>
<tr>
<td>RB</td>
<td>retinoblastoma</td>
</tr>
<tr>
<td>RECK</td>
<td>reversion-inducing cysteine rich protein with Kazal motifs</td>
</tr>
<tr>
<td>RECQL</td>
<td>RECQ helicase-like</td>
</tr>
<tr>
<td>RMA</td>
<td>Multi-array average</td>
</tr>
<tr>
<td>RPKM</td>
<td>Reads Per Kilobase of transcript per Million mapped reads</td>
</tr>
<tr>
<td>RTS</td>
<td>Rothmund-Thomson syndrome</td>
</tr>
</tbody>
</table>
SD standard deviation
SDS Sequence Detection System
SGEF Src homology 3 domain-containing guanine nucleotide exchange factor
SH3 Src homology 3
SV structural variation
TCP Toronto Centre of Phenogenomics
TGF-β transforming growth factor beta
TRC The RNAi Consortium
uPA urokinase plasminogen activator
VEGF vascular endothelial growth factor
WS Werner syndrome
WT wild type
Chapter 1

Introduction
1.1 Osteosarcoma

1.1.1 Overview of Osteosarcoma

Osteosarcoma (OS), although a relatively uncommon cancer with an overall incidence of 5 cases per million people per year, is the most common primary bone malignancy and the second highest cause of cancer-related death in the pediatric age group (Mirabello et al., 2009a, 2009b). OS shows a bimodal age distribution. Children and adolescent are most affected with the peak OS incidence corresponding to the rapid skeletal growth period, followed by a lower incidence in age between 25-60 years. The incidence of OS peaks again in individuals over 60 years of age.

The clinical features of OS are summarized in Table 1-1. OS occurs in virtually all bone types (Skubitz and D’Adamo, 2007). It is commonly found in the metaphyseal region of long bones within the medullary cavity, and penetrates the cortex of the bone to extend into surrounding soft tissues. The two most common sites of OS are the distal femurs and the proximal tibia, which contain large growth plates with high proliferative activity and undergo extensive bone turnover (Clark et al., 2008). OS, a malignant mesenchymal neoplasm, is characterized by a highly cellular tumor composed of pleomorphic spindle-shape cells capable of producing a defective osteoid matrix (immature bone) (Dahlin, 1975). Histologically, OS is categorized as different subtypes based on the predominant pattern of differentiation, including osteoblastic, fibroblastic, chondroblastic, giant cell-rich, telangiectatic and others (Klein and Siegal, 2006). Studies have shown that the histologic subtypes do not influence the clinical outcome of OS (Gorlick R, 2003). However, loss of differentiation in OS results in a significant decrease in patient’s survival.
### Table 1-1. Clinical features of osteosarcoma

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>Bimodal age distribution, with first and larger peak of incidence at age 10-25, and second peak beyond 60 years of age</td>
</tr>
<tr>
<td><strong>Clinical presentation</strong></td>
<td>Pain as the most common symptom (lasting a range of 1-8 months); pathologic fractures are uncommon occurring in less than 5% of patients</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td>Any bone in the body but most commonly metaphyseal portion of appendicular bones; the distal femur accounts for one third of all cases, followed by proximal tibia, and proximal humerus</td>
</tr>
<tr>
<td><strong>Histologic appearance</strong></td>
<td>Malignant spindle cell tumor that produces osteoid; various histological subtypes named based on predominant cell types within tumor and the extent of matrix</td>
</tr>
<tr>
<td><strong>Radiographic appearance</strong></td>
<td>Lytic and blastic bone lesion shows as a “sunburst” appearance with Sharpey’s fiber stretched out perpendicular to the bone; periosteal elevation related to a soft tissue mass produces a “Codman's triangle”</td>
</tr>
<tr>
<td><strong>Dissemination</strong></td>
<td>Approximately 20% of OS metastasize to lungs as the major site as well as to other bones</td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td>Surgery to resect all sites of bulk disease for cure; all high-grade OS, regardless of stage, are treated with chemotherapy, which most typically includes cisplatin, doxorubicin, and high-dose methotrexate</td>
</tr>
<tr>
<td><strong>Prognosis</strong></td>
<td>Age, tumor localization, tumor mass, metastases and response to chemotherapy are all prognostic factors; molecular events associated with disease aggressiveness and chemotherapy response may also serve as prognostic factors. Human with TP53 or RB1 mutations are prone to develop OS.</td>
</tr>
</tbody>
</table>

Table adapted from (Gorlick and Khanna, 2010).
Pretreatment factors that influence prognosis include patient age, primary tumor site and tumor mass, as well as the presence of clinical detectable metastatic disease (Pakos et al., 2009). It is well established that OS in distal sites has a relatively favorable prognosis, while OS on axial location has a considerably worse outcome. Large primary tumor volume also correlates with an adverse prognosis. About 20% of patients have radiographically detectable metastases at diagnosis, and the lung is the most common metastatic site (Meyers et al., 2005). Patients with localized disease normally have a better prognosis (5-year survival rate: 60-80%) than patients with metastasis (5-year survival rate: 15-30%) (Bielack et al., 2002). The prognosis for patients with metastatic disease is determined mainly by the metastatic site, the number of metastases, and the surgical resectability of these metastases (Harris et al., 1998) (Bacci et al., 2008). The survival rate increases to 40% if the cancer has spread only to the lungs or if all the tumors including metastases can be surgically removed (Statistics from American Cancer Society, http://www.cancer.org). Other factors associated with a better prognosis include diagnosis at a younger age, the female gender and a good response to chemotherapy. The prognosis of OS patients has been improved greatly in the past four decades, mainly benefiting from the improvement in surgery and multimodal chemotherapy (Ferguson and Goorin, 2001). However, long-term survival is achieved by radical surgical procedure such as amputation, which greatly influences the patient’s quality of life.

1.1.2 Etiology of Osteosarcoma

It has been a challenge to identify the molecular mechanisms critical to the development of OS owing to the genomic complexity at the time of disease presentation. However, recent advances in molecular biology, cancer genomics and the development of new laboratory OS models are providing insights into the molecular etiology of OS, as summarized below.
1.1.2.1 Bone Growth and Tumorigenesis

Several studies have demonstrated that rapid bone growth rate in puberty is correlated with OS development (Cotterill et al., 2004; Gelberg et al., 1997; Withrow et al., 1991). Since epiphyseal growth plates of the distal femur and proximal tibia are responsible for the increase in height during puberty, approximately 60% of all OS present tumors around the knee. Patients with long-standing Paget’s disease, a bone remodeling disorder characterized by excessive bone breakdown followed by an compensatory increase in bone formation, also have a higher incidence of developing OS (McNairn et al., 2001).

1.1.2.2 Environmental Factors

Several physical, chemical and biological agents have been reported as carcinogens for OS, among which ultraviolet and ionizing radiation are the best examples. Radiation exposure is implicated in about 2% of OS cases (Savage et al., 2013), although an interval of 10–20 years between exposure and OS formation has been noted (Longhi et al., 2003). As radiotherapy is commonly used in children as a treatment for solid tumor, 5.4% of them will develop a secondary cancer among which 25% are OS (Paulino and Fowler, 2005). It has been reported that chemical carcinogens, such as bulky-adduct-forming agents, can induce OS in rodent models by inducing chromosomal instability (Jones, 2011).

1.1.2.3 Inherited Human Diseases Predispose to OS

Several familial syndromes predisposed to OS are summarized in Table 1-2 (Calvert et al., 2012). The genes (to be discussed in detail below) associated with these syndromes typically
<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Description</th>
<th>Gene</th>
<th>Location</th>
<th>Gene function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retinoblastoma (RB)</td>
<td>Malignant tumor of the retina. Most common malignant cancer of the eye in young children. Survivors with germline mutations are at greater risk for subsequent secondary OS.</td>
<td>RB1</td>
<td>13q14.2</td>
<td>Tumor suppressor Cell cycle regulator</td>
<td>(Hansen et al., 1985)</td>
</tr>
<tr>
<td>Li-Fraumeni syndrome (LFS)</td>
<td>Associated with soft tissue sarcomas, breast cancer, and other cancer types. OS is the most prevalent bone sarcoma and presents at a younger age.</td>
<td>TP53</td>
<td>17p13.1</td>
<td>Tumor suppressor DNA damage response Cell cycle regulator</td>
<td>(Birch et al., 1994; Malkin, 2011)</td>
</tr>
<tr>
<td>Rothmund-Thomson syndrome (RTS)</td>
<td>Associated with poikiloderma, craniofacial, short stature and skeletal abnormalities. Increased incidence of OS at a younger age.</td>
<td>RecQL4</td>
<td>18q24.3</td>
<td>RecQ family member with single-stranded DNA annealing activity Lacks DNA helicase activity</td>
<td>(Kitao et al., 1999; Macris et al., 2006; Quan et al., 2007)</td>
</tr>
<tr>
<td>RAPADILINO syndrome</td>
<td>Extremely rare. Associated with short stature, craniofacial abnormalities and diarrhea, but not poikiloderma. 13% develop OS.</td>
<td></td>
<td></td>
<td></td>
<td>(Siitonen et al., 2003, 2009)</td>
</tr>
<tr>
<td>Werner syndrome (WS)</td>
<td>Associated with short stature, bilateral cataracts, cutaneous abnormalities, and premature aging. 7% develop late-onset OS in the fourth decade of life.</td>
<td>WRN</td>
<td>8p12-p11.2</td>
<td>RecQ family member with 3'-5' exonuclease activity</td>
<td>(Goto et al., 1996; Yu et al., 1996)</td>
</tr>
<tr>
<td>Bloom syndrome (BS)</td>
<td>Associated with sun-sensitive skin, short stature, facial dysmorphism, lung disease and mental retardation. Predisposed to common carcinomas as well as OS.</td>
<td>BLM</td>
<td>15q26.1</td>
<td>RecQ family member DNA helicase</td>
<td>(Ellis et al., 1995; German, 1997)</td>
</tr>
<tr>
<td>Paget’s disease</td>
<td>Excessive breakdown of bone with abnormal bone formation and remodeling, resulting in bone pain with weak and malformed bone. &lt;1% develops OS.</td>
<td>TNFRSF11A</td>
<td>18q21-q22</td>
<td>RANK signaling</td>
<td>(Hansen et al., 2006; White and Rushbrook, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SQSTM1</td>
<td>5q31</td>
<td>IL-1/TNF signaling</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAPK8</td>
<td>5q35qter</td>
<td>RANK signaling</td>
<td></td>
</tr>
</tbody>
</table>
encode proteins necessary to stabilize the genome, and their malfunction can lead to defective maintenance of DNA stability.

1.1.2.4 Chromosomal Abnormalities

Genomic studies have demonstrated that OS has a relatively high rate of chromosomal abnormalities, known as chromosomal instability (CIN) in human cancer, which include copy number variations (CNVs) and structural variations (SVs) (Bridge et al., 1997; Chen et al., 2014a; Sadikovic et al., 2009a; Selvarajah et al., 2008; Squire et al., 2003b). CIN plays a significant role in the development of OS, resulting in complicated structural and numerical alterations as well as extensive variability between cells (Bayani et al., 2007). OS genomes exhibit complex karyotypes, characterized by recurrent amplifications and deletions of chromosomal segments, and gross chromosomal abnormalities (Lau et al., 2004). Studies of pre-therapeutic biopsy specimens have identified several common genomic alterations in OS, including amplifications of chromosomes 6p21, 8q24, and 12q14, as well as loss of heterozygosity (LOH) of 10q21.1, which also correlate with poor prognosis (Smida et al., 2010). Examples for numerical chromosomal abnormalities associated with OS include loss of chromosomes 9, 10, 13, and 17, and gain of chromosome 1 (Ta et al., 2009). However, the high rates of CNVs and SVs in OS have made it difficult to distinguish driver mutations from passenger mutations, which has slowed the development of molecular-targeted therapies for this disease compared to other cancer types.

1.1.2.5 Tumor Suppressor Gene Dysfunction

The TP53 (encodes p53 protein) and RB1 (encodes Retinoblastoma protein, pRb) are two well-known tumor-suppressor genes. Mutations in these two genes have proven to be involved in OS
pathogenesis (Marina et al., 2004). p53 is activated by different stresses, including DNA damage, hyper-proliferative signals, nutrient deprivation, hypoxia, oxidative stress and ribonucleotide depletion, which consequently promote diverse responses such as cell cycle arrest, DNA repair, apoptosis and senescence, and lead to tumor suppression (Bieging et al., 2014; Muller and Vousden, 2013). p53 is constitutively inhibited by mouse double minute 2 homolog (MDM2) (Wade et al., 2013). Cellular stresses such as DNA damage result in phosphorylation of p53, which allows p53 to dissociate from MDM2. p53 exerts its tumor-suppressive effects via the activation of pro-apoptotic BCL-2 protein BAX and cyclin-dependent kinase (CDK) inhibitor p21. On the one hand, p53 directly activates BAX in the absence of other BH3 domain only subset of pro-apoptotic BCL-2 proteins to permeabilize mitochondria and therefore engage the apoptotic program (Chipuk et al., 2004). p53 induces p21 expression, which binds and inactivates G1/S phase CDK and S phase CDK complexes, causing cell cycle arrest in the G1 phase (Abbas and Dutta, 2009). The $TP53$ gene is mutated in approximately 22% of OS (McIntyre et al., 1994; Ta et al., 2009). The genetic disease with $TP53$ mutation is Li-Fraumeni syndrome (LFS) (Table 1-2), which is characterized by an autosomal dominant mutation of $TP53$ leading to the development of multiple cancer types, including OS (Hauben et al., 2003). Studies have shown that adenoviral-mediated expression of wild-type p53 resulted in reduced cell viability and increased sensitivity to chemotherapy in $p53$-null SaOS-2 OS cell line (Ganjavi et al., 2006). $TP53$ mutation has also been reported to be more common in high-grade OS versus low-grade OS (Park et al., 2004). Furthermore, $TP53$ mutation is significantly correlated with high levels of genomic instability in OS (Overholtzer et al., 2003).

In addition to p53, the pRb tumor suppressor has also been implicated in OS tumorigenesis (Feugeas et al., 1996). $RB1$ was the first tumor suppressor gene discovered in retinoblastoma (Knudson, 1971), which is a master regulator of many critical biological pathways involved in
cell cycle regulation, differentiation, apoptosis and senescence (Dick and Rubin, 2013; Imai et al., 2014; Khidr and Chen, 2006). Inherited loss of function RB1 mutations causes Retinoblastoma (RB) syndrome, a genetic disease that predisposes a patient to retinal tumors as well as multiple other malignancies including OS (Table 1-2). The pRb protein regulates the cell cycle by binding to the transcriptional factor E2F. E2F is inactivated by pRb until the CDK4/cyclinD complex phosphorylates pRb that leads to E2F accumulation and S-phase entry (Nevins, 2001). As a result, loss of function mutations of RB1 contribute to the continuous cell cycle and cause tumorigenesis (Alberts et al., 2008). Like germline mutations of RB1, copy number loss of the RB1 gene also increase the risk of developing OS (Longhi, 2001). LOH for RB1 has been reported to confer both more aggressive and poorer prognosis for OS patients (Benassi et al., 1999; Feugeas et al., 1996; Heinsohn et al., 2007). Studies in the SaOS-2 OS cell line further showed that cells lacking active pRb had reduced sensitivity to the growth-suppressive effect of methotrexate treatment when compared to cell lines with wild-type pRb (Iida et al., 2003).

1.1.2.6 DNA helicase

RECQ helicase-like (RECQL) belongs to the Superfamily II group of helicases, which help to maintain stability of the genome and suppress inappropriate recombination (Chu and Hickson, 2009; Croteau et al., 2014; Hanada and Hickson, 2007). Loss of function mutations in RECQ family of helicases result in aberrant genetic recombination and DNA replication activities, which lead to CIN and an overall decrease in proliferation. Dysfunctional RECQL4 gene is associated with OS tumorigenesis. Loss of RECQL4 function due to truncation mutations in autosomal recessive familial Rothmund-Thomson syndrome (RTS) (Type II) results in significantly increased risk of developing OS (Table 1-2) (Wang et al., 2003), but in sporadic
OS, the \textit{RECQL4} mutation rate is less than 5% (Nishijo et al., 2004). In contrast, frequent copy number gain and elevated gene expression of \textit{RECQL4} have been reported in sporadic OS (Maire et al., 2009). Bloom syndrome (BS) and Werner syndrome (WS) are two additional autosomal recessive syndromes that predispose affected individuals to OS as well as many other malignancies (German, 1997; Goto et al., 1996) resulting from genetic instability caused by inherited mutations of a RECQL family DNA helicase gene, \textit{BLM} (Bloom syndrome RecQ like helicase, encodes RECQL3) and \textit{WRN} (Werner syndrome RecQ like helicase, encodes RECQL2), respectively (Table 1-2) (Mohaghegh and Hickson, 2001).

1.1.2.7 Transcription Factors

The activator protein 1 (AP-1) complex is a transcription regulator that controls cell proliferation, differentiation, and bone metabolism (Shaulian and Karin, 2002; Wagner, 2002). AP-1 is a heterodimeric complex comprised of members of nuclear phosphoprotein families, including FOS, JUN, ATF and JDP that bind to specific DNA consensus sequence and modulate expression of their target genes (Shaulian and Karin, 2002). FOS and JUN has been reported to be significantly up-regulated in high-grade OS compared with low-grade OS (Wu et al., 1990) and are associated with increased risk to develop metastatic disease (Gamberi et al., 1998). c-Fos and c-Jun double-transgenic mice develop OS with a higher frequency than c-Fos only transgenic mice (Wang et al., 1995). Studies have shown that inhibition of AP-1 mediated transcription blocked migration, invasion, and metastasis in a OS mouse model (Leaner et al., 2009). Using a DNA enzyme Dz13 that cleaves human \textit{JUN} mRNA, OS growth and progression could be inhibited in a clinically relevant murine OS model (Tan et al., 2010).

\textit{MYC} onco gene is a transcription factor that regulates gene expression to stimulate cell growth and division in response to many growth-promoting signals, and contributes to the mutagenesis
of many human cancers (Dang, 2012). MYC amplification has been implicated in OS tumorigenesis as well as resistance to chemotherapies (Bogenmann et al., 1987; Hattinger et al., 2009a; Ueda et al., 1997). Over-expression of Myc in bone marrow stromal cells from a murine OS model leads to OS development accompanied by loss of adipogenesis (Shimizu et al., 2010). MYC is amplified in a human U2OS OS cell line variant resistant to doxorubicin (DX), as well as methotrexate (MTX)-resistant SaOS-2 OS cell lines (Hattinger et al., 2009b). As a result, MYC has been extensively examined as a therapeutic target for OS. Indeed, down-regulation of Myc by antisense phosphorothioate oligodeoxynucleotide (ODN) enhanced the therapeutic activity of MTX against OS cell lines (Scionti et al., 2008). Adenovirus-mediated antisense c-Myc led to increased chemotherapy sensitivity through cell cycle arrest and enhanced apoptosis in the MG-63 OS cell line (Yang et al., 2005). Using a conditional transgenic mouse model of Myc-induced OS, researchers observed that tumor cells underwent proliferative arrest and differentiated into mature bone upon Myc inactivation with doxycycline treatment (Arvanitis et al., 2008).

1.1.2.8 Tumor Angiogenesis

Tumor angiogenesis, the formation and maintenance of blood vessels, is important for sustained OS growth and metastasis, which provides necessary nutrients and oxygen for OS cell proliferation at both primary and metastatic sites. Angiogenesis is regulated by the balance between pro-angiogenic and anti-angiogenic signals (Potente et al., 2011). Neovascularization is induced by tissue hypoxia, blood acidosis, gain of function of oncogenes, and loss of function of tumor suppressor genes.

Vascular endothelial growth factor (VEGF) is the best-characterized pro-angiogenic factor, which stimulates endothelial cell proliferation, migration and blood vessel maturation (Ferrara et al., 2003). Since OS is a relatively vascular tumor, several studies have shown a correlation
between VEGF expression level and microvascular density in OS tumors (Hara et al., 2006; Kaya et al., 2000; Kreuter et al., 2004). On the other hand, anti-angiogenic factors such as Thrombospondin 1, transforming growth factor beta (TGF-β), Troponin I, pigment epithelial-derived factor (PEDF), and reversion-inducing cysteine rich protein with Kazal motifs (RECK) have shown to be down-regulated in OS (Cai et al., 2006; Clark et al., 2007; Moses et al., 1999; Ren et al., 2006).

1.1.2.9 Cell Adhesion, Migration and Invasion

OS is a highly metastatic tumor with pulmonary metastasis are the most common cause of mortality in OS patient (Ta et al., 2009). The metastasis is a multi-step process involving local tumor cell invasion through the detachment from their primary tumor site and adhesion to the extracellular matrix (ECM) followed by local migration and invasion through stromal tissue, entry into the vasculature, and finally extravasation and colonization at the distal metastatic sites (Lee and Lotze, 2009). Therefore, the ability of cancer cells to metastasize relies on complex cell-to-cell and cell-to-matrix interactions. ECM is a collection of extracellular molecules composed of various proteoglycans and protein fibrils secreted by cells that provide structural and biochemical support to the surrounding cells (Hynes, 2009). Cells adhere to ECM components via their cell-surface receptors (Berrier and Yamada, 2007). The major receptors are integrins, which bind to the matrix protein fibronectin and provide a link to the cytoskeleton inside the cell (Luo et al., 2007). The integrins also play a critical role in regulating protrusion and adhesion involved in cell migration (Moissoglu and Schwartz, 2006). Integrin-binding protein Talin associates with integrin’s cytoplasmic domain and acts via other scaffold proteins such as Vinculin, Paxillin, and α-Actin, which in turn up-regulate protein kinases including focal adhesion kinase (FAK), protein kinase C (PKC), PI3K, Src and RhoA GTPases (Berrier and
Yamada, 2007). RhoA GTPases are critical to cell migration (Nimnual et al., 2003). Two important RhoA GTPases are Rac1 and RhoA, which play distinct role in regulating cell migration. Higher Rac1 level can suppress RhoA expression and induce the formation of membrane ruffles, which facilitate cell spreading and migration. In contrast, higher RhoA expression causes low Rac1 level and leads to membrane retraction. Through the coordination of these two processes during cell migration, the leading edge of the cell demonstrates actin remodeling and lamellipoedia formation, while the trailing edge undergoes actin disassembly and retraction. Studies have shown that inhibition of RhoA pathways reduces OS cell migration and invasion (Fromigué et al., 2008).

The cell cortex organized and maintained by the Ezrin, Radixin and Moesin (ERM) proteins also provides a critical connection between the ECM/cell membrane and the underlying filamentous actin inside the cell, which can regulate signaling pathways through their ability to bind transmembrane receptors and activate downstream signals (Neisch and Fehon, 2011). As a result, Ezrin over-expression has been reported to be associated with an increase in metastatic disease (Hunter, 2004). Increased Ezrin expression in pediatric OS patients is associated with reduced disease-free intervals, and down-regulation of Ezrin in a mouse model of human OS has been reported to reduce pulmonary metastasis through the MAPK signaling pathway (Khanna et al., 2004).

OS cell invasion also involves degradation of the ECM components. Matrix metalloproteinases (MMPs) are a family of zinc-dependent endopeptidases capable of degrading a variety of ECM proteins, and are involved in a range of physiological processes including inflammation, wound healing, and fracture healing. In OS, MMPs degrade collagens in ECM, and facilitate both tumor and endothelial cell invasion (Bjørnland et al., 2005). MMPs also play important roles in tumor
angiogenesis (Emonard and Grimaud, 1990; Foda and Zucker, 2001). In addition, MMPs involve in remodeling of vessel walls, which provides a thin and leaky vascular network and promotes tumor cells entering the bloodstream (Oh et al., 2001). Furthermore, MMP-9 mediated ECM degradation releases VEGF stored within the ECM, which in turn is able to up-regulate another MMP, MMP-2 (Bergers et al., 2000). Studies have shown that a MMP-2/MMP-9 double-deficiency significantly impairs tumor angiogenesis and invasion in mouse model (Masson et al., 2005).

MMP binding partner urokinase plasminogen activator (uPA) is another regulator of OS invasion (Choong and Nadesapillai, 2003; Nozaki et al., 2006). The ligand uPA binds to its receptor uPAR in order to become active. Activated uPA cleaves full length plasminogen to release the serine protease plasmin, which in turn breaks down the ECM by degrading ECM proteins, such as fibronectin. It also activates a number of MMPs, including MMP-1 and MMP-13 (Bonnefoy and Legrand, 2000). The role of uPA-uPAR system is well studied in OS tumorigenesis. uPA level has been reported to be negatively correlated with survival (Choong et al., 1996), and down-regulation of uPAR in a OS model resulted in reduced primary tumor growth as well as less frequent metastases (Dass et al., 2005).

1.1.2.10 Osteoclast Function

Osteoclasts are the bone-resorbing cells. Increased osteoclastic activity results in the substantial osteolysis exhibited by some OS (Avnet et al., 2008). Studies have shown that during the initial OS invasion stage, certain growth factors, such as TGF-β, are released from the degraded bone matrix and act on OS cells to stimulate the release of cytokines, including PTHrP, interleukin-6 (IL-6) and interleukin-11 (IL-11) (Guise and Chirgwin, 2003; Quinn et al., 2001). These cytokines then activate osteoclasts to release pro-resorptive cytokines and proteases for resorbing
the nonmineralized components of bone, facilitating further invasion. Cathepsin K (Cat K) is a
cysteine protease produced by osteoclasts, which can breakdown collagen I, osteopontin and
osteonectin in the bone matrix (Stoch and Wagner, 2008). Low Cat K levels at the time of
diagnosis in patients with high-grade metastatic OS have been shown to correlate with a better
prognosis (Husmann et al., 2008).

PTHRP and IL-11 secreted by OS cells can also act on osteoblasts by stimulating the expression
of receptor activator of nuclear factor κB ligand (RANKL), a key mediator of osteoclast
differentiation and activity. OS cells have been shown to autonomously produce RANKL
(Kinpara et al., 2000). RANKL activates osteoclasts through binding to its receptor RANK on
the osteoclast cell surface, whose expression is under control of cytokines IL-1, IL-6, IL-8, TNF-
α, PTHrP and TGF-α, and then initiates a cascade of signals through binding to TRAF-6, and
activating downstream NFκB and MAPK pathways (Hofbauer and Heufelder, 1998).

RANK/RANKL can also activate the c-Fos component of AP-1 complex. As a result, the nuclear
factor of activated T-cells (NFATc1) is up-regulated that in turn up-regulates transcription of
genes involved in osteoclast activation and maturation (Takayanagi, 2007). Osteoprotegerin
(OPG) is a soluble decoy receptor for RANKL and has been found to strongly suppresses
osteoclast differentiation (Tanaka et al., 2005). OPG gene therapy has been applied to a mouse
model of OS and it successfully suppressed osteolytic activity, leading to reduced local OS
progression and improved survival (Lamoureux et al., 2007).

In addition, in response to hypoxic and acidotic stresses, OS cells release endothelin-1 (ET-1),
VEGF and PDGF, which also have predominantly osteoblast-stimulatory functions (Chirgwin
and Guise, 2007; Kingsley et al., 2007).
1.1.3 Metastasis of Osteosarcoma

The most common metastasis site in OS patients is the lung causing respiratory failures (Ta et al., 2009). The classical paradigm of metastasis steps discussed above also holds true in OS: including detachment of cells from the primary neoplasm, invasion into the surrounding ECM and entrance into circulatory system, survival in the circulation and followed by subsequent extravasation, and lastly, the establishment and proliferation at new metastatic sites (Hanahan and Weinberg, 2000). A major difference in metastasis process between OS and other epithelial cancers is that OS initiating cells, osteoblasts, are of mesenchymal origin. Therefore, much of the recent focus on Epithelial-Mesenchymal-Transition (EMT) involved mechanism does not apply to OS. In contrast, it appears that other factors such as cell-cell adhesion molecules, chemokines and motility-molecules play important roles in OS metastasis. Several genes have been reported to be involved in OS metastasis, including *EZRIN* (Khanna et al., 2004), *CXCR4* (Laverdiere et al., 2005; Miura et al., 2005), *ErbB2* (Ta et al., 2009) and *MMPs* (Bjørnland et al., 2005). *EZRIN*, as part of the ERM proteins, and MMPs have been discussed above. *CXCR4* is a chemokine receptor found in a variety of tumor metastases, including OS. It is associated with poor prognosis in some studies and its inhibition results in the complete destruction of metastases in a xenograft mouse disease model (Laverdiere et al., 2005; Miura et al., 2005). *ErbB2/HER2* is a well-known oncogene in breast cancer. Some studies suggested that its over-expression in OS is correlated with increased metastasis and reduced patient survival, although other studies demonstrated that the correlation of *ErbB2* with prognosis in OS is still controversial. As a result, Herceptin/trastuzumab, a monoclonal antibody that inhibits gene function of *ErbB2*, has gone through phase 2 clinical trial for metastatic OS in combination with cytotoxic chemotherapy (Ebb, 2012; Ta et al., 2009). However, the outcome for all patients remained poor with no significant difference between the *HER2*-positive and negative groups. Unfortunately,
although many other genes have also been suggested to play a role in OS metastasis, most studies are largely correlational observations and lack detailed mechanisms.

1.1.4 Models of Osteosarcoma

1.1.4.1 Mouse Models of OS

The development of an “ideal” model of OS has proven to be particularly difficult. Several OS mouse models have been developed which are summarized in Table 1-3, although there is still a lack of optimal in vivo animal models that closely recapitulates all aspects of the human condition at the temporal, physiological and histopathological level.

Our lab has previously generated a transgenic mouse model of osteosarcoma, termed the MOTO (Murine Osteocalcin driven T-antigen Osteosarcoma) mice (Molyneux et al., 2010). MOTO was created by transgenically expressing SV40 T-antigen under the OG2 (Osteocalcin) promoter. MOTO mice develop bone lesions as early as eight weeks, which then grow into identifiable tumor by twelve weeks. MOTO mice develop malignancies at all sites that affect human disease, with a peak incidence in the long bones, as observed in human OS. They further recapitulate human OS by displaying spontaneous metastases. At endpoint, lung metastases are seen in greater than 95 percent of the mice, and a much lower frequency is observed in the liver.
## Table 1-3. Mouse models of OS

<table>
<thead>
<tr>
<th>Mouse model</th>
<th>Latency</th>
<th>Tumors</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P53-/-</td>
<td>15 months</td>
<td>50% develop lymphomas, soft tissue sarcoma and OS</td>
<td>(Donehower et al., 1992; Jacks et al., 1994)</td>
</tr>
<tr>
<td>P53-/-</td>
<td>6 months</td>
<td>lymphomas, soft tissue sarcoma and OS</td>
<td>(Jacks et al., 1994)</td>
</tr>
<tr>
<td>p53R172H Arg</td>
<td>7-21 months</td>
<td>37% were OS (69% of these metastasized); 14% are sarcomas (fibro-, angio- and lypo-sarcomas)</td>
<td>(Liu et al., 2000; Olive et al., 2004)</td>
</tr>
<tr>
<td>p53R172H</td>
<td></td>
<td>Well differentiated OS characterized by an abundance of new bone with osteoclast recruitment</td>
<td>(Rauch et al., 2010)</td>
</tr>
<tr>
<td>Tax+ARF-/-</td>
<td>7 month</td>
<td>Develop a variety of cancers, including OS</td>
<td>(Damo et al., 2005)</td>
</tr>
<tr>
<td>p18(Ink4c)/p53-</td>
<td>na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIC1+/+-/p53+/+-</td>
<td>18 months</td>
<td>40% developed OS</td>
<td>(Chen et al., 2005)</td>
</tr>
<tr>
<td>RecQL4, RTS+/+</td>
<td>na</td>
<td>Lymphomas and OS</td>
<td>(Ichikawa et al., 2002)</td>
</tr>
<tr>
<td>Rb/p53 DKO:Osx-1 Cre</td>
<td>4 months</td>
<td>75% developed OS, 40% of cases metastasis</td>
<td>(Berman et al., 2008; Walkley et al., 2008)</td>
</tr>
<tr>
<td>Tet-inducible Myc transgenic mice</td>
<td>&lt;8 weeks</td>
<td>61% developed OS when Myc was turn on after weaning</td>
<td>(Arvanitis et al., 2008)</td>
</tr>
<tr>
<td>Mice Type</td>
<td>Age</td>
<td>Development</td>
<td>References</td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>c-Fos/c-Jun double transgenic mice</td>
<td>4-6 weeks</td>
<td>90-95% developed OS after 14 weeks of age with all bones of the body affected</td>
<td>(Sunters et al., 1998; Wang et al., 1995)</td>
</tr>
<tr>
<td>c-Fos transgenic mice</td>
<td></td>
<td>65-80% developed OS after 14 weeks of age</td>
<td></td>
</tr>
<tr>
<td>Apc1638N/+;twist null/+ mice</td>
<td>5 months</td>
<td>30% developed OS in cephalic region</td>
<td>(Entz-Werle et al., 2010)</td>
</tr>
<tr>
<td>α-amylase promoted- SV40 Tag (501 mice)</td>
<td>8 months</td>
<td>Developed metastatic OS with an average life span of 13 months of age</td>
<td>(Schell et al., 2000)</td>
</tr>
<tr>
<td>Osteocalcin-SV40 T/t antigen (MOTO mice)</td>
<td>8-12 weeks</td>
<td>Exhibited OS at all sites known to be affected in the human disease, with a peak incidence in the long bones</td>
<td>(Molyneux et al., 2010)</td>
</tr>
</tbody>
</table>

na: information was unavailable. Table adapted from (Kansara and Thomas, 2007).
1.1.4.2 Canine Osteosarcoma Models

Comparative oncology is the study of cancer biology and therapy in animals with naturally occurring cancers. With the strong similarity between canine and human OS, natural development of OS in dogs provides us an important opportunity to extend our knowledge of OS tumorigenesis as well as new therapy development through preclinical studies (Paoloni and Khanna, 2008; Sutter and Ostrander, 2004; Vail and MacEwen, 2000). OS in both human and dog are characterized by primary tumor growth in the appendicular skeleton with a high risk for metastasis to the lungs (Gorlick 2010). The canine OS is indistinguishable from the human disease at the histologic and gene expression levels. In addition, therapeutic treatments for both the primary tumor and the metastatic disease showed similar responses in both species (Coomer et al., 2009). However, the primary differences between human and canine OS are the age of development and the prevalence of disease. OS tends to occur in middle-aged to older dogs between 6 to 12 years of age, with a median age of 7 years (Misdorp and Hart, 1979), whereas OS in human occurs most commonly in children and adolescent with a second peak in individuals over 60 years of age.

Several reasons make canine an ideal model to study the pathogenesis of human OS. At first, there is a greater genetic homology between dogs and humans than between human and mice (Figure 1-1) (Kirkness et al., 2003; Switonski et al., 2004). Secondly, dogs represent a more outbred population than inbred laboratory animals such as mice. Naturally occurring tumors in dogs share many aspects in tumor biology and behavior with human cancers, while tumors in experimental laboratory animals are usually induced artificially (Pinho et al., 2012). The incidence of many cancer types in dogs is higher than in humans, which allows rapid
Figure 1-1. Human and canine ortholog genes

(A) Representative picture of human and canine ortholog genes. The image was generated by R program with data available in Ensembl Genes 80. Line color was arbitrarily assigned. Each line connects a gene ortholog between human genome build GRCH38.p2 (top) and canine genome build CanFam3.1 (bottom).

(B) Human $RET$ and $PTEN$ ortholog genes reside on canine chromosomes 28 and 26, respectively, as examples.
accumulation of tumor samples. Indeed, OS is the most common primary bone tumor in dogs, and accounts for approximately 85% of malignancies originating in the skeleton (Withrow and MacEwan, 2001). Finally, the progression of canine disease is usually faster, which provides a good opportunity to evaluate novel therapeutic treatments in dogs in a relatively short time period for quick validation.

1.2 Identifying cancer ‘driver’ genes using genomic tools

1.2.1 Cancer Genome

It is believed that all cancers are developed as a result of genomic alterations. Early insights into the central role of the genome in cancer development emerged in the late nineteenth. Over the past 30 years, a lot have been learned about cancer genome including the accumulated discoveries of mutated genes in human cancer, which has provided key insights into the molecular mechanisms underlying tumorigenesis, and guided the design of targeted approaches for clinical intervention of human cancer.

Cancer is an evolutionary process (Figure 1-2). Cancer development is believed to involve two constituent processes: the continuous acquisition of heritable genetic variations in individual cells by randomly acquired mutations, and the natural selection force resulting phenotypic diversity (Stratton et al., 2009). The selection may remove cells that have acquired deleterious mutations, or it may promote outgrowth of cells carrying genetic alterations that confer the capability to proliferate and survive more effectively than other neighboring cells. Within an adult human, there may be thousands of cells with limited abnormal growth potential and undetectable as common benign growths. Occasionally, a single cell may acquire certain
Figure 1-2. Timeline for somatic mutations acquired by the cancer cells

Mutations can be acquired during cell division, which including both the intrinsic mutations acquired during normal cell division and the stimulated effects of exogenous mutagens. During cancer development, other cellular processes, such as DNA repair defects, may further contribute to the additional mutation burden. Passenger mutations do not confer any growth advantage on cancer cell, whereas driver mutations can promote clonal expansion of the cancer cell. Recurrence of the disease after initially responsive treatments can be associated with pre-existing or newly acquired drug-resistant mutations under selection. Adapted from (Stratton et al., 2009).
mutations providing the cell sufficiently proliferate advantageous that allows it to outgrowth, invade and metastasize.

Like all the normal cells in the human body, through a lineage of mitotic cell divisions, a cancer cell has gradually acquired a set of genomic alterations from its progenitor fertilized egg, which are collectively termed somatic mutations to distinguish them from germline mutations that are inherited from parents and can be transmitted to offsprings. There are several different types of genetic alterations or mutations: single base substitutions (missense or nonsense), insertion or deletion of small or large DNA segments (in-frame or frameshift), DNA rearrangement in which DNA has been broken and rejoined to a DNA segment from elsewhere in the genome (gene disruption or fusion gene formation), copy number changes (increases or reductions) from the two copies present in the normal diploid genome (Stratton et al., 2009). In addition, the cancer cells can also acquire completely new exogenous DNA sequences, such as virus infection. Human papilloma virus (HPV), Epstein Barr virus (EBV), hepatitis B virus (HBV), human T lymphotrophic virus 1, and human herpes virus 8 have all been known to contribute to the tumorigenesis of certain cancer types (Talbot and Crawford, 2004).

Although cancer cells acquire many somatic changes in the genomic DNA, not all these abnormalities involve in cancer development. Therefore, somatic mutations in a cancer genome can be classified into two groups according to its consequences during cancer development (Figure 1-2) (Stratton et al., 2009). ‘Driver’ mutations confer growth advantage on the cells carrying them and have been positively selected in the microenvironment of the tissue during the evolution of the cancer. A driver mutation is causally implicated in oncogenesis, although it may not be required for the maintenance of the tumor, but it must have been positively selected at certain time point during cancer development. The other class is ‘passenger’ mutation that are
found within cancer genomes because random somatic mutations occurred during cell division and does not have functional consequences, but present in an ancestor of cancer cell by chance due to together with the driver mutations. A passenger mutation has not undergone selection, does not confer clonal growth advantage, and therefore is not contributed to cancer development. Therefore, driver mutations normally cluster in a subset of cancer related genes, whereas passenger mutations are more randomly distributed across the genome. One important subclass of driver mutations is a mutation that confers resistance to cancer therapy (Figure 1-2), which are frequently found in recurrences of cancers that have initially responded to certain treatment but then develop resistance (Holohan et al., 2013). Resistant mutations often confer limited growth advantage on the cancer cells in the absence of therapy. Some resistant mutations may exist initially as passenger mutations with limited growth advantage compared to the driver mutations within a minor sub-population of the cancer cells until the selective environment is formed by the initiation of therapy (Mullighan et al., 2008; Roche-Lestienne et al., 2002). These passenger mutations can then convert into drivers and promote resistant subclones preferentially expanding as the recurrent disease. Resistant mutations can also be de novo acquired during cancer cell division under the selection of treatment.

The central goal of cancer genome analysis is to identify cancer genes that carry driver mutations and responsible for tumorigenesis. Therefore, the key challenge will be to distinguish driver from hundreds of passenger mutations on the background. So far, at least 400 of the 22,000 protein-coding genes in the human genome have been reported to present recurrent somatic mutations in cancers with strong evidence of their contribution to cancer development or drug-resistance (Futreal et al., 2004).

Besides the multistep model for cancer development with progressive accumulation of genetic
alterations during cancer evolution mentioned above, recent whole genome sequencing studies also revealed a new phenomenon, termed chromothripsis (chromosome shattering) (Holland and Cleveland, 2012; Stephens et al., 2011; Zhang et al., 2013), in which tens to hundreds of clustered genome rearrangements present in one cancer cell, and appear to be derived from a single catastrophic event. The detailed underlying mechanism of such phenomenon is still unknown. During chromothripsis, loss of tumor suppressor function and oncogenic fusion may be caused by rejoining of chromosomal pieces by end-joining DNA repair pathways, while aberrant DNA replication may lead to oncogene amplification (Forment et al., 2012). Involvement of micronuclei from DNA damage is also proposed as a model for the generation of chromothripsis (Zhang et al., 2015). A higher frequency of chromothripsis was observed in cells with p53 mutations. In bone cancers, such as OS, chromothripsis is quite common and can involve more than one chromosomes (Stephens et al., 2011).

1.2.2 Array-based Comparative Genomic Hybridization

Many human genetic disorders including cancer result from unbalanced chromosomal abnormalities. Traditionally, cytogenetic analysis of banding patterns of the chromosomes karyotype was the primary tool for clinical assessment of patients with a variety of chromosomal abnormalities, in which as small as approximately 5Mb of aberrations can be detected under ideal conditions (Bayani and Squire, 2004). Later on, fluorescence in situ hybridization (FISH) was used as a newer cytogenetic technique, which uses fluorescently labeled probes to target and visualize specific DNA sequences on the chromosomes through hybridization (Buckle and Kearney, 1994). A wide range of probes can be used, including whole-chromosome painting probes, repetitive sequence probes and locus-specific probes (1-10kb) (Kearney, 2001). However, prior knowledge of the chromosome imbalance under investigation is required for
designing probes for FISH, and only limited targets can be tested in one assay. In addition, high-quality metaphase spreads have to be prepared, which is often difficult for solid tumors.

As a result, comparative genomic hybridization (CGH) technology was developed as a discovery tool for genome-wide screening of CNVs in solid tumor, which is a modification of quantitative two color FISH and utilizes genomic DNA to overcome the need for metaphase chromosomes (Kallioniemi et al., 1992). In CGH, the genomic DNAs from the test sample and a control sample from an individual with a normal karyotype are extracted, differentially labeled with fluorochromes, and competitively hybridized to reference metaphase chromosomes. The fluorescent signal intensity of the labeled test DNA samples relative to that of the control DNA along the chromosome can then be quantified using digital image analysis tools, allowing the identification of specific copy number changes (Speicher and Carter, 2005). However, the resolution of CGH has been limited to alterations of approximately 5-10 Mb for most clinical applications (Kirchhoff et al., 1998; Pinkel and Albertson, 2005). Another limitation of CGH is that structural rearrangements within balanced chromosome, such as translocations and inversions that can cause gain or loss of gene function, cannot be detected. In addition, whole genome copy number changes, or ploidy changes, will be missed (Speicher and Carter, 2005).

In order to overcome the aforementioned limitations associated with traditional CGH, a newer method has been developed as array-based comparative genomic hybridization (aCGH) that combines the principles of CGH with the advantage of microarray technology (Schena et al., 1995). In aCGH, metaphase chromosomes are replaced by large numbers of small segments of DNA (known as probes) as the targets that are arrayed onto a solid support, such as glass slide, in an ordered fashion (Lucito et al., 2003). This method greatly increases the resolution of screening for genomic copy number gains and losses because probes are several orders of
magnitude smaller than metaphase chromosomes. The level of resolution is usually determined by the probe size, the genomic distance between probes as well as the density of probes on the array. Current popular human genome CGH microarrays have a high resolution down to 10-50kb (Agilent Technologies and Affymetrix).

Similar to CGH, in aCGH, genomic DNAs are extracted from a test sample and a normal reference sample, which are then labeled with fluorescent dyes of different colors. These two populations of labeled DNAs are then mixed together and applied to a microarray platform in the presence of excess Cot-1 DNA as blocking reagent to prevent binding of repetitive sequences in the genome. Denatured single strand DNAs will hybridize with their complementary probes fixed on the microarray, and unbind DNAs will be washed out. Digital imaging systems are used to capture and quantify the relative fluorescence intensities of the probes that have hybridized with labeled DNA fragments. The fluorescence ratio of the test and reference signals is determined at different positions along the genome, and provides information on the relative copy number of specific sequences in the test genome as compared to the reference normal genome (Manning et al., 2007). aCGH provides us the opportunity to simultaneously detect ploidy changes, deletions and amplifications of any genomic locus represented on an chosen microarray. The flexibility of array design has allowed developing specialized arrays for different application or for any specific diseases. A further advantage is that aCGH process can be automated for high-throughput application.

Several studies have used CGH or aCGH to reveal genetic imbalances in OS, and the most frequent regions of gains and losses were identified. (Atiye et al., 2005; Forus et al., 1995; Man et al., 2004; Ozaki et al., 2002; dos Santos Aguiar et al., 2007; Squire et al., 2003a; Zielenska et al., 2004). These studies revealed that gains were more frequent than losses in OS genome. The
median number of copy number aberration events in primary high-grade OS was significantly higher than in low- or intermediate-grade OS.

1.2.3 Comparative Oncogenomics

Chromosomal abnormalities have been associated with most human malignancies, which can be detected by methods including cytogenetics and aCGH as described above. Recurrent CNVs have been found to be associated with particular cancer types (Baudis, 2007; Weir et al., 2007a). The sizes of CNVs are variable ranging from less than a single gene to entire chromosome changes (Beroukhim et al., 2010; Pinkel et al., 2005). Cancer driver genes have been successfully identified within recurrent focal CNVs across multiple cancer samples through serials of functional validation studies to evaluate all of the candidate genes within the locus of interest (Sawey et al., 2011). However, large chromosome or chromosome arm-level CNVs, usually caused by aneuploidy, are commonly observed, which include too many genes and make the one-by-one validation approach impossible. Neither improved resolution of genomic screening technology nor increased tumor sample size can fully resolve this problem because many cancer driver genes likely occur within such large CNVs (Baudis, 2007). Therefore, there is an urgent need to find an effective approach that can filter out and reduce the number of candidate driver genes in these large CNVs to a number feasible to do the one-by-one functional validation. Cross-species comparative oncogenomics is one approach to overcome this obstacle (Figure 1-3) (Mattison et al., 2010; Peeper and Berns, 2006; Tomlins and Chinnaiyan, 2006). It has been well established that the function of human cancer genes is well conserved across other mammals. Recent large-scale mouse-to-human and dog-to-human genomic comparisons confirmed that evolutionary conservation could be used as an effective filter to reduce the noise in genomic datasets (Kim et al., 2006; Maser et al., 2007; Wallace et al., 2012; Zender et al., 2006). Indeed, comparative oncogenomics approach has already been used to make important
contributions to the understanding and practice of human oncology in fields such as basic cancer biology (Pang and Argyle, 2009; Sulaimon and Kitchell, 2003), cancer immunotherapy (G.H. and D., 1982), radiation biology (Powers et al., 1987), and systemic therapies (London et al., 2003; Selting et al., 2008) for a variety of cancers.
To identify candidate oncogenes, aCGH data and expression analysis are integrated on the gene level to compare genomic abnormalities between two species. Recurrent events serve as a filter to limit the number of candidate genes. Selected genes then undergo biological validations for tumorigenic or metastatic activity in different models.
1.3 Rationale and hypothesis

OS is the most common primary malignancy of bone. The 5-year survival rate for OS is 60%–70%, with no significant improvements in prognosis since the advent of multi-agent chemotherapy. Diagnosis, staging, and surgical management of OS remain focused on our anatomical understanding of the disease. Although several genes have been identified as discussed above, the molecular pathogenesis of OS and the mechanism for its metastasis are still largely unknown. Therefore, a comprehensive understanding of the molecular pathogenesis of OS is needed in order to identify more therapeutic targets and manage the disease.

Genomic studies have demonstrated that OS has a relatively high rate of chromosomal abnormalities. However, these CNVs and SVs in osteosarcoma have made it difficult to distinguish driver mutations from passenger mutations. Naturally occurring canine OS serves as an ideal model to study the pathogenesis of human OS due to the greater homology between dogs and humans. Therefore, in this study, we have hypothesized that using high-resolution aCGH data from a large collection of human and canine OS samples combined with cross-species comparative oncogenomics will help to further shorten the driver genes candidate list, and allow prioritization the target genes for functional validation of their role in OS tumorigenesis.
Chapter 2

Materials and Methods
2.1 aCGH

7 published human OS genomic datasets on Gene Expression Omnibus (GEO) were selected, including GSE19180 (Yang et al., 2013a), GSE12830 (Sadikovic et al., 2009b), GSE14827 (Kobayashi et al., 2010), GSE9654 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9654), GSE9508 (Endo-Munoz et al., 2010a), GSE16102 (Paoloni et al., 2009), GSE19276 (Endo-Munoz et al., 2010b). These datasets were obtained on four different aCGH platforms, including Agilent-013282 Human Genome CGH Microarray 44B, Agilent-014693 Human Genome CGH Microarray 244A, Affymetrix Human Genome U133 Plus 2.0 Array, and Agilent-012391 Whole Human Genome Oligo Microarray G4112A. Probe coordinates were obtained from manufacturers' website and normalization were performed respectively as per peer-reviewed literature or manufacturers' specifications.

Canine OS samples were provided by Dr. Geoffrey Wood (University of Guelph). Genomic DNA were extracted using phenol chloroform and treated with RNaseA (Invitrogen) for 1 hour at 37°C. aCGH assay was performed and analyzed on Affymetrix Canine Genome 2.0 Array platform according to manufacturers' specifications at the University Health Network Microarray Centre. This work was a collaboration between the Wood and Khokha laboratories.

2.2 Cell culture

2.2.1 Culture medium

U2OS, JL-31A and SJSA-1 cells were obtained from ATCC and were cultured in RPMI-1640 medium (Gibco) with 10% fetal bovine serum (FBS) (Gibco) and 100U/ml penicillin/streptomycin (Gibco). Hu3N-1 and HuO9 cells were obtained from Riken Institute and cultured in DMEM medium (Gibco) with 10% FBS and 100U/ml penicillin/streptomycin. SaOS-
2 cells were obtained from ATCC and were cultured in McCoy’s 5A medium (Gibco) with 10% FBS and 100U/ml penicillin/streptomycin.

2.2.2 Lentivirus production and transduction

Virus production was performed according to TRC (The RNAi Consortium) lentiviral production and transduction protocol. Briefly, 293T packaging cells were cultured in DMEM medium with 10% fetal bovine serum until they were 70%-80% confluent, and then transfected with 10µg lentiviral construct, 9µg packaging plasmid (pCMV-dR8.91) and 1µg envelop plasmid (pCMV-VSV-G) using FuGENE HD Transfection Reagent (Promega) according to the manufacturer’s protocol. 293T cells were incubated for 16–18 hours (37°C, 5% CO₂) before medium was changed to remove the transfection reagent and replaced with 10mL of DMEM medium containing 30% fetal bovine serum for viral harvest. Virus-containing supernatant was collected at 48 hours and 72 hours post-transfection, and subsequently filtered through 0.45µm low-protein-bind filter to remove cellular debris. The lentivirus solution was used immediately, or aliquoted and stored at -80°C for later use.

Cells to be transduced were cultured until 80%-90% confluency and culture medium was changed 2 hours prior to transduction. 2mL of lentivirus solution was added per 10cm cell culture plate in the presence of 8µg/mL polybrene. Cells were incubated for 24 hours, followed by selection with puromycin for 48 hours. The puromycin concentrations used, ranging from 2-5µg/ml, were previously titered in each cell line individually, so that all un-transduced cells would die within 48 hours after puromycin addition. After puromycin selection, transduced cells were maintained in regular culture medium and used for further biochemical and functional analysis.
2.2.3 Cell proliferation assay

$10^5$ cells were seeded in each well in triplicate in a 6-well tissue culture plate. Cells were either cultured in normal serum medium or low serum (1% FBS) condition, and then trypsinized and resuspended in culture medium for viable cell quantification using Trypan Blue solution (Sigma) and hemocytometer under a phase-contrast microscope.

For cell proliferation assay in low-attachment plate, $10^5$ cells were seeded in each well in triplicate in a 6-well ultralow-attachment plate. Cells were either cultured in normal serum medium or low serum (1% FBS) condition, and then pelleted by centrifugation at 300g for 5 minutes at room temperature. The cell pellets were digested by 0.05% trypsin for 5 minutes at 37°C, and then resuspended in culture medium to stop the trypsinization. Viable cells were quantified using Trypan Blue solution and counted by hemocytometer.

2.2.4 3D Matrigel colony formation assay

400µl of Matrigel (BD Biosciences) was added to each well of a 24-well tissue culture plate and allowed to harden for 1 hour at 37°C. 2X$10^4$ cells were suspended in 1mL of culture medium and seeded onto each well in triplicate. Cells were allowed to proliferate for 2–3 weeks and to form colonies. These were imaged and quantified under a phase-contrast microscope.

2.2.5 Scratch-wound assay

Cells were cultured in 6-well plate until 90-100% confluency as a monolayer were reached. Wounds were created by gently and slowly scratching a straight line with a sterile 200µl pipet tip across the center of the well in one direction. Detached cells were carefully rinsed off with PBS, and remaining cells were replenished with fresh culture medium. Cells were cultured for 24 hours to monitor migration. Phase-contrast photos were taken under microscope with the same
configurations for different samples. The gap distance was quantitatively evaluated using Image J software (http://rsb.info.nih.gov/ij/download.html).

2.2.6 Transwell migration assay

5,000-20,000 cells were seeded on top of the 24-well transwell insert (BD Biosciences) in appropriate medium with 0.5% FBS. The transwell insert was then placed into the well of 24-well plate with 10% FBS in appropriate medium as an attractant. Cells were then cultured for 24 hours. GFP imaging was captured under an Olympus fluorescent microscope. JL-31A cells were first stained with crystal violet reagent and then pictures were taken using a light microscope. Migrated cells were then quantified.

2.3 Mice

All animal protocols were approved by the University Health Network Animal Care Committee (Toronto, ON, Canada) and performed in accordance with the standards of the Canadian Council on Animal Care. Mice were housed in the Princess Margaret Hospital Animal Care Facility. User protocol AUP 849.1 was strictly followed.

2.3.1 Establishment of tumor cell line xenografts in immunodeficient mice

Cells grown in monolayer culture were harvested during the exponential growth phase using trypsin digestion, resuspended in culture medium and then the number of cells was quantified using a hemocytometer. Cells were pelleted by centrifugation at 225g for 5 min at RT, and re-suspended in PBS at a concentration of 2-5×10^5 cells/ml depending on the expected cell number to be injected.

8-12 weeks-old recipient nonobese diabetic/severe combined immunodeficiency (NOD-SCID)
mice were placed in the anesthesia chamber. After a mouse reached sufficient anesthetic depth, it was removed from the chamber and placed in a properly sized nose cone with the dorsal side facing upwards. The cell suspension was agitated to prevent the cells from settling, and withdrawn from the sterile tube into a 1-cc TB syringe (BD) without the needle. The skin of the mouse was lifted to separate it from the underlying muscle and inject 0.1 ml of the cell suspension (20,000 SJSA-1 cells or 50,000 JL-31A cells), subcutaneously with a 21 G needle per flank. Injection was visually checked for leakage. The animal was placed in a clean cage and observed for 10–15 min to ensure recovery from the anesthetic.

SJSA-1 cells were transduced and selected with luciferase-GFP expression vector. SJSA-1 cell generated mouse tumor xenografts were monitored weekly for 4 weeks by injecting 150mg/Kg D-luciferin intra-peritoneally 15-20min prior to bioluminescence imaging using Xenogen IVIS Imaging System 100 (STTARR facility, UHN).

JL-31A cell generated mouse tumor xenografts were monitored weekly for 4 weeks by palpation. Mice were terminated at 4 weeks post tumor cell injection, and tumors were dissected for further analysis.

2.3.2 Establish of Osteoblast-specific Dlg2 null mice

Dlg2 conditional null mouse (Dlg2\textsuperscript{fl/+}) on C57BL/6 background (Dlg2tm1a\textsuperscript{(EUCOMM)Wtsi}) was obtained from Wellcome Trust Sanger Institute, created by International Mouse Phenotyping Consortium (IMPC) (Figure 2-1). The mice can be genotyped through a combination of separate PCR reactions that detect LacZ, the gene-specific wild type (WT) allele, and a mutant allele-specific short range PCR (Table 2-1).
Figure 2-1. Targeting strategy for conditional Dlg2 null mouse model

(A) Targeting vector. Construct map for conditional Dlg2 null mouse. Exon 11 of Dlg2 was flanked by two loxP sites in the same direction.

(B) Strategy for conditional gene knock out using either Flp and Cre or Cre alone. Cre alone was used in this study; exon 11 of Dlg2 will be deleted upon Cre induced recombination, creating a Dlg2 null mouse model with LacZ expressing as a reporter.
### Table 2-1. PCR primers for genotyping

<table>
<thead>
<tr>
<th>Target genes</th>
<th>PCR type</th>
<th>Primer name 1</th>
<th>Primer sequence 1</th>
<th>Expected band size 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dlg2</td>
<td>Mutant</td>
<td>Dlg2_42053_F</td>
<td>CCAGAATGTACTTCAGCACCA</td>
<td>222bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAS_R1_Term</td>
<td>TCGTGGTATCGTTATGCACC</td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td></td>
<td>Dlg2_42053_F</td>
<td>CCAGAATGTACTTCAGCACCA</td>
<td>312bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dlg2_42053_R</td>
<td>TGTGTGTATGTTGCTGTGTTT</td>
<td></td>
</tr>
<tr>
<td>LacZ</td>
<td></td>
<td>LacZ_2_small_F</td>
<td>ATCACGACGCGCTGTAT</td>
<td>108bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LacZ_2_small_R</td>
<td>ACATCGGGCAAAATATACG</td>
<td></td>
</tr>
<tr>
<td>Tp53</td>
<td>floxed</td>
<td>p53–int10–fwd</td>
<td>AAGGGGTATGAGGGGACAAGG</td>
<td>584bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p53–int10–rev</td>
<td>GAAGACAGAAAAGGGAGGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>null</td>
<td>p53–int1–fwd</td>
<td>CACACACACAGGTAAACCC</td>
<td>612bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p53–int10–rev</td>
<td>GAAGACAGAAAAGGGAGGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WT</td>
<td>p53–int10–fwd</td>
<td>AAGGGGTATGAGGGGACAAGG</td>
<td>431bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p53–int10–rev</td>
<td>GAAGACAGAAAAGGGAGGG</td>
<td></td>
</tr>
<tr>
<td>Rb</td>
<td>floxed</td>
<td>Rb19E</td>
<td>CTCAAGAGGCTGAGACTCATG</td>
<td>283bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rb18</td>
<td>GCCGTGTGCCATCAATG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>null</td>
<td>Rb212</td>
<td>GAAAGGAAAGTCGAGGACATTG</td>
<td>260bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rb18</td>
<td>GCCGTGTGCCATCAATG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WT</td>
<td>Rb19E</td>
<td>CTCAAGAGGCTGAGACTCATG</td>
<td>235bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rb18</td>
<td>GCCGTGTGCCATCAATG</td>
<td></td>
</tr>
<tr>
<td>Cre</td>
<td></td>
<td>Cre5’</td>
<td>CCTGGAAATGCTTCTGTCGATCC</td>
<td>654bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cre3’</td>
<td>GAGTTGATAGGCTGCTGGTGGCAGATG</td>
<td></td>
</tr>
</tbody>
</table>
Mice expressing \(Rb^{fl/fl}\): \(Tp53^{fl/fl}\) (Marino et al., 2000) and mice harboring an osteoblast-specific Cre recombinase under the promoter of collagen type 1a (Col-1a Cre) (Dacquin et al., 2002) were backcrossed into FVB genetic background. Col-1a Cre mice were crossed with \(Rb^{fl/fl}\): \(Tp53^{fl/fl}\) mice to obtain combined deletion of \(Rb\) and \(Tp53\) in the osteoblast lineage (\(Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}\); \(ΔOB\), deletion in osteoblast lineage).

\(Dlg2^{fl/fl}\) mice were then crossed to \(Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}\) mice in order to generate \(Dlg2^{ΔOB/+}\): \(Rb^{ΔOB/+}: Tp53^{ΔOB/+}\) mice. These mice were then intercrossed to generate \(Dlg2^{ΔOB/ΔOB}: Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}\) mice as well as their littermate controls. For genotyping, genomic DNA was prepared from mouse tails, and then subjected to PCR with corresponding genotyping primers (Table 2-1).

2.4 Cloning

Total RNA from human fibroblast was extracted using the Qiagen RNeasy kit (Qiagen) with on-column DNase I treatment. cDNA was generated by reverse transcription from total RNA using SMARTer PCR cDNA synthesis kit (Clontech) according to the manufacturer's protocol. There are six isoforms of \(DLG2\) in humans (Figure 2-2A) (NM_001142699, NM_001364, NM_001142700, NM_001142702, NM_001206769, NM_001300983). Isoform 1 is the longest isoform with a characteristic L27 domain. Isoform 1 and 2 are believed to be the dominant isoforms. \(DLG2\) also exhibits high level of similarities across mammalian species (Figure 2-2B). Therefore, for our validation experiment, cDNA from human \(DLG2\) isoform 1 was used and cloned into a lentiviral vector, pLKO.1 (Figure 2-2B).
Figure 2-2. Construction of DLG2 expression vector

(A) Human DLG2 isoforms. PDZ, PSD95/DLG/ZO-1 domain; SH3, Src homology 3 domain; GMPK, guanosine monophosphate kinase domain.

(B) Construction of DLG2 expression vector. Isoform 1 of Human DLG2 was cloned into a lentiviral vector. Expression of DLG2 is driven by CMV promoter while eGFP and Puro are used as identification markers under SV40 promoter. Percentage of similarity between human and murine or canine are shown in the middle table.
2.5 RNA-seq

Total RNA was extracted using the Qiagen RNeasy kit (Qiagen) with on-column DNase I treatment. Total mRNA from 1 µg of total RNA for each sample was enriched with the NEBNext PolyA mRNA Magnetic Isolation Module (NEB) and further processed with the NEBNext Ultra Directional RNA Library Prep Kit (NEB) according to the manufacturer's protocol. RNA libraries were QC and then sequenced on an Illumina HiSeq2500 platform (125bp paired-end) at the Donnelly Sequencing Centre, University of Toronto. Sequence reads were aligned using the TopHat pipeline (https://ccb.jhu.edu/software/tophat/index.shtml) and differential gene expression was called using the Cufflinks pipeline (http://cole-trapnell-lab.github.io/cufflinks/) (Figure 2-3).

2.6 Biochemical analysis

Total protein from cells or tissues were extracted by lysing or homogenizing the sample in RIPA buffer (50mM Tris-HCl, pH7.5, 150mM NaCl, 2mM EDTA, 1% NP40, 0.5% Na deoxycholate, 0.1% SDS) containing a protease and phosphatase inhibitor cocktail (Roche). Cell lysate or tissue homogenate was then centrifuged at 16,100 x g for 15 min at 4 °C, and the resulted supernatant were collected. Cell lysates (10-25µg protein) were resolved by SDS-PAGE and specific protein of interest was analyzed by immunoblotting.

Several DLG2 antibodies were tested, AB5168 (rabbit-based polyclonal Ab) from Millipore were determined to be the most appropriate antibody and used in downstream experiments. Beta-actin (Santa Cruz) was used for sample loading control. Primary antibody binding was visualized with IRDye infrared secondary antibodies using the Odyssey Infrared Imaging System and Odyssey V3.0 software (LI-COR Biosciences).
Total RNA was extracted from tumor samples followed by mRNA enrichment using oligo dT magnetic beads. Directional RNA libraries were prepared, QC and then sequenced on Illumina sequencing platform. Sequence reads were QC and aligned to determine the Reads Per Kilobase of transcript per Million mapped reads (RPKM) for each gene. Genes with less than 1 RPKM were filtered out. Differential gene expression was called. Top candidate genes were compiled for further validation.

Figure 2-3. Schematics of RNA-seq work flow on tumor samples.
2.7 X-ray and MicroCT scan

For X-ray imaging (Molyneux et al., 2010), mice were anesthetized with 2% isoflurane, and imaged using a Faxitron MX-20 Radiography system (Faxitron X-ray Corporation) with a 24 kV, 4 second exposure time. Multiple images per mouse in 2 orientations were analyzed using Specimen DR version 3.2.2 software (Faxitron X-ray Corporation) to generate length ($l$), width ($w$), and the height ($h$) measurements (in mm) for the diameter of each lesion, the tumor volume was estimated as the volume of an ellipsoid ($\frac{4}{3}\pi l w h$).

For MicroCT scan (Molyneux et al., 2010), isoflurane-anesthetized mice were imaged using a Locus Ultra microCT scanner (GE Healthcare) subjected to a 16-second Anatomical Scan Protocol (total of 68 images were generated for each mouse) at 80kV, 70mA, using a 0.150mm Cu Filter to achieve a resolution of 150µm. 3D micro-CT data were reconstructed at 58-µm resolution using the Feldkamp algorithm.

2.8 Histology and immunohistochemistry

Organs or tumor samples were dissected, washed thoroughly with PBS, fixed with 4% paraformaldehyde (PFA) overnight at 4°C, decalcified in 20% EDTA (pH 7.4), and then embedded in paraffin. Sections (5µm) were prepared and stained with hematoxylin & eosin (H&E) or Masson-Trichrome by Toronto Centre of Phenogenomics (TCP) pathology laboratory via routine protocols (http://www.phenogenomics.ca/cmhd/index.html?v=5).

2.9 Cell cycle analysis by flow cytometry

Cells were cultured in 6-well ultra-low attachment plate (Nunc). BrdU was diluted to a 1mM concentration with 1X DPBS according to the BD Pharmingen BrdU Flow Kits Instruction
Manual. BrdU-pulse was done by adding BrdU to each well at a final concentration of 10µM for 1-3 hours depending on type of cell lines, except to the unpulsed control well. After pulse, cells were washed twice with 1X HBSS to remove unincorporated BrdU, fixed using BD Cytofix/Cytoperm buffer, washed and stained according to the BD Pharmingen BrdU Flow Kits Instruction Manual. Flow cytometry preparation was performed according to manufacturer's recommendation (BD). Cell collection was done on BD Fortessa II. Analysis was done using FlowJo Version 7.65 or Version VX. The gating pathway used for cell-cycle analysis was shown in Figure 2-4.

2.10 Quantitative real time PCR

For gene expression assay, total RNA from cell lines or tumor tissues was prepared using the RNeasy mini-kit (QIAGEN). RNA purity was confirmed using a Nanodrop 2000 Spectrophotometer (Thermo Scientific). 2 µg of total RNA was reverse transcribed using SuperScript III first-strand synthesis system (Life Technologies) and then subjected to quantitative real time PCR (qPCR). For qPCR-based copy number analysis, genomic DNA from tumor tissues was extracted using Qiagen DNeasy Blood and Tissue Kit (Qiagen), and then analyzed by qPCR. qPCR was performed with specific PCR primers for human and canine DLG2 genes in SYBR Green PCR master mix (Applied Biosystems) on ABI PRISM 7900HT system (Applied Biosystems). GAPDH gene was used as a control for gene expression. RnaseP gene was used as a control for copy number analysis, and 0.25-fold cutoff was used. All raw data
BD Biosciences protocol was followed for cell cycle analysis on BD flow cytometer. In brief, the data was acquired using the BrdU FITC assay by creating a worklist and running samples automatically on the Loader with acquisition criteria of 10,000 events or 3 minutes for each tube. During acquisition preview, gates for cells were adjusted in the FSC-A vs SSC-A plot, and the DNA 7-AAD-A voltage was adjusted to place the mean of the singlet peak (G0/G1) at 50,000 in the histogram. In addition, cell cycle gates were adjusted as needed to encompass the G0/G1, S, and G2/M populations. The percentage of events in the cell gate (% total cells), and the percentages of cells in G0/G1, S, and G2/M phases were calculated.

Figure 2-4. Gating pathway used for cell cycle analysis by flow cytometry
were analyzed using Sequence Detection System software (SDS version 2.1; Applied Biosystems).

2.11 Gene expression microarray analysis

Total RNA from canine OS tumors was prepared using the RNeasy mini-kit (QIAGEN). RNA purity was confirmed using a Nanodrop 2000 Spectrophotometer (Thermo Scientific). Gene expression microarray on GeneChip Canine Genome 2.0 Array (Affymetrix) were performed at the University Health Network Microarray Centre. Gene identifiers were downloaded from manufacturer's website (Affymetrix). Multi-array average (RMA) (Irizarry et al., 2003) normalization was used across the sample sets examined. Batch correction was performed using distance weighted discrimination (DWD) (Benito et al., 2004). Differentially expressed genes were identified using T-test with FDR correction (P<0.05 and FDR<0.1) or Samtools. Analyses were performed using R (version 2.82), Agilent Genespring GX (version 11) or Microsoft Excel 2007. Appropriate software packages were obtained from https://www.bioconductor.org and http://linus.nci.nih.gov/download_BRBArryTools.html. Gene Ontology analysis was performed on https://david.ncifcrf.gov/. Total Ensemble genes were used as background. Pathways with P <0.1 were analyzed.

2.12 Statistics

Statistical significance was determined using appropriate statistical methods based on the experimental designs. Two-tailed Student's t-test was used to determine if two sets of normally distributed data are significantly different from each other, and were performed using PRISM 5.0. software. Biological replicates were performed independently at least 3 times. Log rank test was performed for survival curves using PRISM 5.0. software. For all studies, p<0.05 was
considered significant. Specialized analyses for bioinformatic data, including aCGH analysis, RNA-seq analysis and gene expression microarray analysis, were performed using dedicated tools mentioned elsewhere in the methods.
Chapter 3

Results
3.1 Cross-species Oncogenomics Identifies $DLG2$ as a Tumor Suppressor Candidate in Osteosarcoma

3.1.1 Optimization of aCGH analysis

In this study, we aimed to further investigate the mutated genes for OS pathogenesis by using a cross-species comparison of aCGH data between human and canine OS genomes. To this end, we required to adopt a robust aCGH analysis bioinformatics pipeline in order to obtain reliable CNV results across multiple samples and datasets.

The data obtained from aCGH experiments are the intensities for each marker in test and reference sample. If markers with normalized intensities in test sample are significantly greater than the reference intensities, these positions are considered copy number gains. Similarly, significantly lower intensities in the test sample compared to those in reference sample are indicative of copy number losses. The statistical methods for analyzing CNV data are therefore aimed at identifying locations with significant copy number gains or losses. Several detection methods are available for identifying CNVs from aCGH data. The statistical models underlying these approaches include hidden Markov models (HMMs) (Colella et al., 2007; Wang et al., 2007), segmentation algorithms (Hupé et al., 2004; Olshen et al., 2004), $t$-tests and standard deviations (SDs) of the log $R$ ratio (Fiegler et al., 2006). Although all these programs are available for detecting CNVs from SNP arrays, a thorough comparison of these methods has not been conducted. Lai and his group tested 11 methods using receiver operating characteristic (ROC) curves and found that segmentation algorithms performed consistently well (Lai et al., 2005). Therefore, we adopted and compared the two most common segmentation algorithms: circular binary segmentation (CBS) (Olshen et al., 2004) and gain and loss of DNA (GLAD) (Hupé et al., 2004).
CBS model is a modification of binary segmentation (Sen and Srivastava, 1975), which segments a chromosome into contiguous regions and bypasses parametric modeling of the data with its use of a permutation reference distribution (Olshen et al., 2004). GLAD is a methodology for the automatic detection of breakpoints from aCGH profile, and the assignment of a status to each chromosomal region (Hupé et al., 2004). The breakpoint detection step is based on the Adaptive Weights Smoothing (AWS) procedure (Polzehl and Spokoiny, 2000).

The recurrent copy-number alterations (CNAs) were then determined by Genomic Identification of Significant Targets in Cancer (GISTIC), which is a tool to identify targeted genes by focal somatic CNAs (Mermel et al., 2011). Genes targeted by somatic CNAs have been postulated to play important roles in tumorigenesis and cancer therapy (Santarius et al., 2010). However, the discovery of new cancer genes being targeted by CNAs is complicated. Firstly, somatic alterations are acquired randomly during cell division, and only driver alterations can confer growth advantage and promote cancer development (Greenman et al., 2007). Neutral or weakly deleterious passenger alterations may still maintain within the subclone carrying driver alterations (Merlo et al., 2006). Secondly, large CNA may simultaneously affect up to thousands of genes, but the selective benefits are likely to be mediated by only one or a few of driver genes within the region. Therefore, additional experimentation and analysis are required to distinguish the driver from the passenger alterations, and to shortlist the genes for functional validation. A commonly used approach to identify drivers is to study a large cohort of cancer samples, with the concept that regions containing driver alterations would appear at higher frequencies than regions containing only passengers. For such purpose, GISTIC algorithm was developed to identify likely driver CNAs by evaluating the frequency and amplitude of observed CNA events (Mermel et al., 2011). This algorithm has been applied to to identify new targets of amplifications (Bass et al., 2009; Chiang et al., 2008; Firestein et al., 2008a; Weir et al., 2007b).
and deletions (Northcott et al., 2009) in multiple cancer types (Beroukhim et al., 2007; Firestein et al., 2008b; Lin et al., 2008; Weir et al., 2007c). Several additional algorithms for identifying driver somatic CNAs have also been described (Shah, 2009). However, none of these methods can provide a priori statistical confidence to interpret copy-number analysis like GISTIC method, which can help the end-users to prioritize candidate genes for time-consuming validation experiments.

As a result, our aCGH analysis pipeline was optimized as shown in Figure 3-1. In brief, normalized aCGH log ratios from each sample were analyzed by two different segmentation algorithms, CBS or GLAD, to accurately define the copy number profile in the sample. Segmented data was then analyzed by GISTIC algorithm. Targeted genes in recurrent CNAs across the whole dataset were reported.

We first used the aCGH dataset (GSE19180) (Yang et al., 2013b) from 10 frozen human OS specimens obtained on Agilent-013282 Human Genome CGH Microarray 44B as a training dataset to compare these two different segmentation methods. Using either CBS or GLAD segmentation method, we received similar GISTIC output (Figure 3-2A and B), indicating the high concordance between CBS and GLAD algorithms. 700 significant genes were identified using CBS method, while 698 genes were identified using GLAD method where 589 (84%) overlapped between the two methods (Figure 3-2C). Therefore, genes/CNAs identified by both CBS and GLAD were chosen for downstream analysis.
Figure 3-1. aCGH analysis pipeline

Potential algorithms and analysis paths were shown. Paths labeled with red line were taken in this study.
Figure 3-2. Test of different segmentation methods using a training dataset
Deletion GISTIC plot using CBS segmentation method (A) and GLAD segmentation method (B) are shown with G-scores on the top and Q-value at the bottom (FDR-corrected P-value). A higher peak indicates greater statistical significance. Y-axis is chromosomal locations from Chr1 to Chr22, which is aggregated vertically. (C) Venn diagram showing the numbers of genes deemed to be significantly altered using CBS and GLAD.
3.1.2 Cross-species oncogenomics analysis identify DLG2 gene

Naturally occurring canine OS serves as an ideal model to study the pathogenesis of human OS. Due to the greater homology between dogs and humans than between human and mice, canine OS can be used to further narrow down mutated genes in human OS pathogenesis through cross-species comparative oncogenomics. To this end, 98 human OS aCGH analysis from 7 published human OS genomic datasets were selected from a total of 34 datasets on Gene Expression Omnibus (GEO) (see Chapter 2 Materials and Methods for more details). Selection criteria used were: (a) genomic resolution finer than 50kb; (b) the dataset contains more than 5 unique OS samples; (c) the datasets were accompanied by interpretable sample keys and full sample information. In parallel, aCGH data were generated from 9 canine OS specimens and analyzed.

Figure 3-3 shows that both human and canine OS contain high levels of chromosomal instabilities as indicated by frequent whole chromosomal arms gains and losses. When pooling the CNA levels from all human OS samples, recurrent CNAs across all samples were identified. The “usual suspect” oncogenes (MYC) and tumor suppressors (PTEN and RB1) were identified through this analysis, which provide proof to the principle of our approach (Figure 3-4, top panel). A novel Chr11:q14 region showing focal deletion across multiple samples was also noted, with as yet unknown potential tumor suppressor genes (Figure 3-4, bottom panel). The segmented CNA profiles were further subjected to GISTIC algorithm to identify targeted genes by focal somatic CNAs. Our analysis revealed that top mutated genes in human OS include MYC, PTEN, CDKN2A/B and RB1, which were the same in canine OS. This indicated the high level of disease similarity between humans and dogs (Figure 3-5).
Figure 3-3. Whole genome landscape of CNAs in human and canine OS

Segmented CNA profiles of 98 Human OS genome (left panel) and 9 canine OS genome (right panel) were mapped. Red indicates amplification and blue indicates deletion. X-axis presents individual samples. Y-axis lists chromosomal locations from Chr1 to Chr22 for human, or Chr1 to Chr38 for dog. Red indicates amplification and blue indicates deletion.
Figure 3-4. Identification of a putative tumor suppressor gene locus

Selected zoomed-in Integrative Genomics Viewer (IGV) view of the human OS CNA profiles shows the known oncogenes and tumor suppressors including *MYC*, *PTEN* and *RB1* (Top; red indicates amplification and blue indicates deletion) as well as a focal deletion at Chr11:q14 region with an unknown gene(s) (Bottom).
Examination of the deletion landscape of human and canine OS genomes revealed that Disks large homolog 2 (DLG2) is one of the top altered gene within both human and canine OS besides known tumor suppressor genes PTEN and RB1 (Figure 3-6A and B). We next validated the CNV status of DLG2 by qPCR analysis in 9 additional canine OS samples, and concluded that 28.9% of human OS and 44.4% of canine OS carry a DLG2 deletion (Figure 3-6C).

An independent cohort of 31 human OS samples was then obtained from Dr. David Malkin (SickKids Hospital, Toronto, Canada) and Dr. Ana-Patino Garcia (University Hospital of Navarra, Pamplona, Spain). Of these, 9 samples exhibited DLG2 deletion as determined by qPCR-based copy number analysis (Figure 3-7A). The mutational frequency of DLG2 in this new cohort of patients is 29.0%, which is similar to the discovery cohort shown in (Figure 3-6C).

We examined the Catalogue of Somatic Mutations in Cancer (COSMIC) database and found that DLG2 is highly mutated in other human cancers (Figure 3-7B and C), with most of the mutations resulting in protein sequence changes (Figure 3-7B). Several mutational hotspots are found within the DLG2 protein, among which are V161 and V191 within the first PDZ domain, and S624 located in the connector sequence following SH3 domain (Figure 3-7C). In total, we observed 11 DLG2 mutations in 101 human tumors (11%) (Table 3-1a). Of these, 7 mutations would result in drastic amino acid property changes as listed in Table 3-1b.

Taken together, our analysis of a large cohort of human OS aCGH datasets and its comparison to canine OS aCGH datasets pointed to a focal deletion within Chr11:q14 region in human OS, which contains the DLG2 gene. DLG2 deletion was further validated in an independent cohort of human OS, which confirmed a detection frequency of 29% in human OS cases. COSMIC database further revealed DLG2 mutation in 11% of multiple human cancers, suggesting that DLG2 gene is a potential tumor suppressor, and may play an important role in OS tumorigenesis.
Figure 3-5. Top five mutated oncogenes and tumor suppressors in OS

Grid of top mutated oncogenes and tumor suppressors in both human (left) and canine (right). Samples are ranked based on DLG2 status, which is shown in the last row. Red indicates amplification and blue indicates deletion.
Figure 3-6. Identification of *DLG2* as a tumor suppressor in OS

(A) Deletion GISTIC output using CBS segmentation method in human (left) and canine (right) is shown with G-scores on the top and Q-value at the bottom. Y-axis is vertically aggregated chromosomal locations. Genes in the highest peaks are labeled.

(B) Top altered genes in human and canine from our OS aCGH analysis. *DLG2* is ranked among the top 10 in both species.

(C) Cases and frequencies of *DLG2* deletion in both human and canine OS tumors.
Figure 3-7. DLG2 in OS and other human cancers

(A) 9 out of 31 human OS samples exhibited DLG2 deletion in an independent cohort.

(B) Types and fractions of DLG2 mutations from COSMIC database.

(C) Mutational hotspots within the DLG2 protein. DLG2 domain structure is shown. Cases of amino acid substitutions along the protein is labeled as black bar, and the position of Ins/del identified indicated by blue arrow. Several mutational hot spots are labeled below the domain structure of the protein.
Table 3-1. *DLG2* mutations in human cancer

a. *DLG2* mutations frequency in cancer

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th># of Samples Analyzed</th>
<th># of Mutation observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovary</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Pancreas</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>Breast</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Pancreas</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>Skin</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Biliary tract</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Upper AG tract</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Colon</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>CNS</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>101</strong></td>
<td><strong>11</strong></td>
</tr>
</tbody>
</table>

b. *DLG2* nonsynonymous mutations in cancer

<table>
<thead>
<tr>
<th>mRNA</th>
<th>Protein</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.2075G&gt;A</td>
<td>p.R692Q</td>
<td>Positive</td>
<td>Polar</td>
</tr>
<tr>
<td>c.2158G&gt;A</td>
<td>p.E720K</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>c.962C&gt;T</td>
<td>p.P321L</td>
<td>Special</td>
<td>Hyrophobic</td>
</tr>
<tr>
<td>c.2287A&gt;G</td>
<td>p.R763G</td>
<td>Positive</td>
<td>Polar</td>
</tr>
<tr>
<td>c.481G&gt;T</td>
<td>p.V161F</td>
<td>Hyrophobic</td>
<td>Aromatic</td>
</tr>
<tr>
<td>c.491G&gt;T</td>
<td>p.S164I</td>
<td>Polar</td>
<td>Hydrophobic</td>
</tr>
<tr>
<td>c.1589_1590CC&gt;TT</td>
<td>p.S530F</td>
<td>Polar</td>
<td>Aromatic</td>
</tr>
</tbody>
</table>
3.2 Functional Validation of DLG2 as a Tumor Suppressor in vitro and in vivo

3.2.1 Functional validation of DLG2 in OS cell lines

From our cross-species comparative oncogenomics study, we identified DLG2 as a potential tumor suppressor gene, which may play an important role in OS pathogenesis. DLG2 is also known as channel-associated protein of synapse-110 (chapsyn-110) or postsynaptic density protein 93 (PSD-93), which is a multiple PDZ domain-containing protein that belongs to the family of molecular scaffolding proteins known as membrane-associated guanylate kinases (MAGUKs) (Kim et al., 1996; Mazoyer et al., 1995). Members of this family are defined by a basic core of three different protein domains: the PDZ domain, an SH3 domain, and the guanylate kinase (GK) domain (Oliva et al., 2012a). The organization of these domains enables MAGUKs to function as scaffold proteins that orchestrate the assembly of multiple signal transduction networks.

In Drosophila, there is only one family member, DLG, which has tumor suppressor properties. Specifically, the loss of DLG function is associated with excessive proliferation and neoplastic transformation of epithelial cells within the imaginal discs (Subbaiah et al., 2012a; Woods and Bryant, 1989). In C. elegans the single homologue of DLG, DLG1, is required for proper adherens junction formation (Bossinger et al., 2001; Firestein and Rongo, 2001). However, whether mammalian homologues of DLG (DLG1, DLG2, DLG3, DLG4 and DLG5) also possess tumor suppressor function is not well understood. Several studies have focused on mammalian DLG1, and a number of its binding partners have been identified, including APC (Zhang et al., 2011), beta-catenin (Subbaiah et al., 2012a), MEK2 (Gaudet et al., 2011), PTEN (Cotter et al., 2010), NET1 (Carr et al., 2009) and SGEF (Subbaiah et al., 2012b). These vary depending on the
location of DLG1 protein in the cell as the function of DLG proteins can be regulated by their subcellular localization. DLG1 has shown tumor suppressor properties in several studies (Fuja et al., 2004; Gardiol et al., 1999; Sotelo et al., 2012), while it has also been demonstrated to exhibit oncogenic activities in some contexts. As a major target of the HPV E6 oncoprotein, DLG1 is responsible for the invasive capacity of HPV transformed cells through its interaction with RhoG-specific Src homology 3 domain-containing guanine nucleotide exchange factor (SGEF) (Subbaiah et al., 2012b). The biological function of DLG2 was well studied in neuronal system as a regulator of synaptic function (Guo et al., 2012; McGee et al., 2001a; Sun and Turrigiano, 2011). It is enriched at postsynaptic sites, and forms complexes with various synaptic proteins to construct postsynaptic signaling networks and control synaptic transmission (Brenman et al., 1996). However, the role of DLG2 in cancer remains poorly investigated.

If a cell contained an alteration of a driver gene, then ablating the function of this gene should have an effect on its tumorigenic properties. For functional validation of DLG2, we set out to perform gain-of-function and loss-of-function studies using OS cancer cell lines. First, we sought to use cell lines with pre-existing DLG2 deletion for downstream validation. All available OS cell lines from multiple sources (including ATCC, RIKEN and collaborating laboratories) were screened for DLG2 status based on their already available genomic profiling (SNP array or aCGH). We found four human OS cell lines and one canine OS cell line, which harbored DLG2 losses (Figure 3-8 and data not shown). These cell lines were pilot-tested for their ability to form colonies in soft agar assays and to form xenograft tumors in NOD-SCID mice. These assays are widely used to examine tumorigenic characteristics of cell lines. Among these cell lines, only one human cell line (SJSA-1) and one canine cell line (JL-31A) were able to develop colonies on soft agar or xenograft tumors in NOD-SCID mice in our laboratory (Table 3-2). However, human SaOS-2 line has been demonstrated to be xenograftable in the literature (Jia et al., 2002;
Yuan et al., 2009), suggesting that the line in our laboratory might lose its original xenograft property.

There are six isoforms of $DLG2$ in humans (http://www.ncbi.nlm.nih.gov/gene/1740). Isoform 1 and 2 are believed to be the dominant isoforms, in which isoform 1 is the longest isoform with a characteristic L27 domain. For our validation experiment, human $DLG2$ isoform 1 cDNA was cloned into lentiviral vector (see Chapter 2 Material and Methods), and lentivirus-based delivery was able to enable $DLG2$ re-expression in all of above 4 cell lines (Figure 3-9).
Table 3-2. Properties of each cell lines in various *in vitro* and *in vivo* assays

<table>
<thead>
<tr>
<th>Species</th>
<th>Name</th>
<th>Source</th>
<th>DLG2 Status</th>
<th>Xenograftable</th>
<th>Soft Agar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>SAOS-2</td>
<td>OS</td>
<td>Loss</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SJSA-1</td>
<td>OS</td>
<td>Loss</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Hu3N1</td>
<td>OS</td>
<td>Primary</td>
<td>Loss</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HuO9</td>
<td>OS</td>
<td>Primary</td>
<td>Loss</td>
<td></td>
</tr>
<tr>
<td>Canine</td>
<td>JL-31A</td>
<td>OS</td>
<td>Loss</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Figure 3-8. Selection of OS cell lines with *DLG2* deletion

4 cell lines exhibiting *DLG2* losses were selected for experiments towards functional validation. Red lines indicate the genomic location of *DLG2* gene.
Figure 3-9. Lentivirus-based delivery of DLG2 re-expression in various cell lines

Re-expression of DLG2 in multiple OS cell lines detected by western blotting with specific DLG2 antibody. Beta-actin (β-actin) was used as a loading control. For each cell line, left lane is the uninfected parental cell line, and right lane is the lentiviral-transduced cell line for DLG2 re-expression.
Functional validation experiments were then carried out in these cell lines. Overall, DLG2 re-expression had little or no effect on proliferation in normal or low serum conditions, although SJSA-1 cells showed a significant reduction when grown in low FBS conditions for 8 days (Figure 3-10).

We also performed 3D-matrigel assay, which showed that re-expression of DLG2 in SJSA-1 or JL-31A cells significantly reduced their colony formation ability (Figure 3-11). This suggested that DLG2 may play a role in anchorage-independent growth of OS cells. Additional examination in soft agar assay showed that re-expression of DLG2 in these two cell lines reduced their capacity to form colonies (Figure 3-12).

To test the effect of DLG2 expression on tumorigenesis, parental SJSA-1 and JL-31A cells and their corresponding lentiviral induced DLG2 expressing cell lines were injected subcutaneously into NOD-SCID mice. The SJSA-1 cells were also co-transduced with luciferase-GFP expression virus for in vivo imaging. Tumor growth was monitored by either bioluminescence imaging for SJSA-1 cell derived tumors (Figure 3-13A) or palpation for JL-31A cell derived tumors (Figure 3-13B). We found that re-expression of DLG2 in these cell lines could significantly influence xenograft tumor growth, resulting in smaller tumors (Figure 3-13, lower panel).

To further investigate the molecular effect of DLG2 in OS cells, gene expression array data was generated from 9 canine tumors and analyzed (see Chapter 2 Material and Methods). In combination of the corresponding aCGH data generated earlier, we stratified tumors into DLG2-loss vs. DLG2-normal groups, and found 167 genes that to be differentiated expressed within the two groups (Figure 3-14A). These genes were identified using either T-test with FDR correction (Figure 3-14B) or SamTools (Figure 3-14C), and both methods yielded consistent results.
2,000-10,000 Cells were seeded per well, depending on the cell line, and cultured in 6-well plates in triplicate under 10% FBS or 1% FBS conditions. Their growth was monitored for up to 8 days. Cell numbers were determined using either hemocytomer or Celltiter-Glo assay, and then plotted with day “0” value as 100. Three independent experiments were performed. Results for each time point are shown as mean±SD. Two-tailed student’s t test was performed between DLG2 re-expression cells and control cells. *, p<0.05.
Figure 3-11. DLG2 re-expression reduces colony formation under 3D-matrigel conditions.

Representative pictures of colonies in 3D-matrigel. DLG2 re-expression induced significant reduction in colony numbers. Quantification of colony area is shown on the right. The experiment was done in triplicate. 5 fields for each sample were randomly selected and quantified using Image J software. Data are shown as mean±SD. Two-tailed student’s t test was performed between DLG2 re-expression cells and control cells. *, p<0.05.
**Figure 3-12. DLG2 re-expression reduces colony number in soft agar assay.**

Representative soft agar images are shown on the left. Colonies were stained with crystal violet and colony number was counted manually. The experiment was done in triplicate, and three independent times. Data were shown as mean+SD. Two-tailed student’s t test was performed between DLG2 re-expression cells and control cells. *, p<0.05.
Figure 3-13. DLG2 re-expression reduces the tumor xenograft growth in vivo

(A) 20,000 SJSA-1 cells were injected subcutaneously into 8-12 weeks-old NOD-SCID mice. Mice were monitored weekly for 4 weeks by injecting D-luciferin intra-peritoneally. A representative image shown at 3 weeks post-injection (top). Experiment were performed two times, n=10 per group. Colored contours represent Photon influx, which is a surrogate marker for tumor size. The intensity of Photon influx in each individual mouse is shown at the bottom, mean±SD. Two-tailed student’s t test was performed for testing the significance.

(B) 50,000 JL-31A cells were injected subcutaneously into 8-12 weeks-old NOD-SCID mice. Mice were monitored weekly for 4 weeks by palpation. Mice were sacrificed at 4 weeks, and tumors were dissected. Shown are tumors from 5 mice/group (top). Experiment were performed two times, n=10 per group. Tumor volume was calculated and is shown at the bottom, mean±SD. Two-sided student’s t test was performed for testing the significance.
Figure 3-14. Intergrative analysis of aCGH and gene expression from canine tumors

(A) Differentially expressed genes in DLG2-null vs DLG2-wildtype groups, which were identified using T-test with FDR correction (B, highlight in blue) or SamTools (C, highlight in red and green). Both methods yielded identical results.

(D) Top altered pathways were analyzed using Gene Ontology pathways.
Differentially expressed genes were then analyzed using Gene Ontology pathways (GO Enrichment Analysis, Gene Ontology Consortium, http://geneontology.org) to identify top altered pathways as shown in Figure 3-14D.

RAN/RHO activity modifying proteins, GTPase activating Rap/RanGAP domain-like 3 (GARNL3) (Wiemann et al., 2001), were among the top differentially expressed genes between DLG2-null and DLG2-wildtype groups (Figure 3-15). Tumors with DLG2-null had a much lower expression of GARNL3. All 3 probes corresponding to GARNL3 were among the top differentially expressed list, with an average fold change of greater than 5 folds between the two groups. GARNL3 is highly conserved across mammalian species and encodes a protein that exhibits GTPase activator activity. However, the function of GARNL3 has not yet been well studied so far.

The acquisition of a motile and invasive phenotype is an important step in the development of tumors, and ultimately the metastasis. This step requires the abrogation of cell-cell contacts, the remodeling of the extracellular matrix and of cell-matrix interactions, and finally the movement of the cell mediated by the actin cytoskeleton (Friedl and Wolf, 2003). The function of small GTPases signaling pathways has been implicated in migration and invasion (Bravo-Cordero et al., 2012; Schmitz et al., 2000). Inhibition of RhoA has been shown to reduce OS cell migration and invasion (Fromigué et al., 2008). We therefore performed a scratch-wound cell migration assay. DLG2 re-expression cell lines showed significantly larger blank area 24 hours after the scratch (21% vs 4% for Hu-09 cells, and 41% vs 5% for JL-31A cells) (Figure 3-16). We also observed a similar effect of DLG2 in transwell migration assay (Figure 3-17). Here, re-
<table>
<thead>
<tr>
<th>Parametric p-value</th>
<th>FDR</th>
<th>Geom mean of intensities in class 1</th>
<th>Geom mean of intensities in class 2</th>
<th>Fold-change</th>
<th>UniqueID</th>
<th>Symbol</th>
<th>Name</th>
<th>ENTREZ id</th>
<th>Accession</th>
<th>UGCluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.60E-06</td>
<td>0.0112</td>
<td>10</td>
<td>140.92</td>
<td>0.071</td>
<td>A_11_P120626</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.70E-06</td>
<td>0.0112</td>
<td>10.59</td>
<td>84.14</td>
<td>0.13</td>
<td>A_11_P000039966</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.80E-06</td>
<td>0.0112</td>
<td>152.03</td>
<td>19.39</td>
<td>7.84</td>
<td>A_11_P000038842</td>
<td>GZMA</td>
<td>granzyme A (granzyme 1, cytotoxic T-lymphocyte-associated serine esterase 3)</td>
<td>487207</td>
<td>XM_544335</td>
<td></td>
</tr>
<tr>
<td>2.50E-06</td>
<td>0.0117</td>
<td>10</td>
<td>313.68</td>
<td>0.032</td>
<td>A_11_P164793</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.10E-06</td>
<td>0.0283</td>
<td>618.46</td>
<td>31.8</td>
<td>19.45</td>
<td>A_11_P060756</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1.8E-05</td>
<td>0.0283</td>
<td>49.72</td>
<td>1960.49</td>
<td>0.025</td>
<td>A_11_P039441</td>
<td>NPY</td>
<td>neuropeptide Y</td>
<td>475257</td>
<td>XM_532492</td>
<td>Cf1 20447</td>
</tr>
<tr>
<td>7.1.25E-05</td>
<td>0.0283</td>
<td>10.94</td>
<td>62.91</td>
<td>0.17</td>
<td>A_11_P112196</td>
<td>TTYH1</td>
<td>twenty homolog 1 (Drosophila)</td>
<td>476386</td>
<td>XM_533589</td>
<td>Cf1 10146</td>
</tr>
<tr>
<td>8.1.34E-05</td>
<td>0.0283</td>
<td>212.8</td>
<td>1241.76</td>
<td>0.17</td>
<td>A_11_P119391</td>
<td>GARNL3</td>
<td>GTPase activating Rap/RanGDP domain-like 3</td>
<td>480718</td>
<td>XM_857841</td>
<td>Cf1 5502</td>
</tr>
<tr>
<td>9.1.36E-05</td>
<td>0.0283</td>
<td>10</td>
<td>81.37</td>
<td>0.12</td>
<td>A_11_P207533</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.1.92E-05</td>
<td>0.0316</td>
<td>37.13</td>
<td>547.16</td>
<td>0.068</td>
<td>A_11_P126441</td>
<td>FLRT3</td>
<td>fibronectin leucine rich transmembrane protein 3</td>
<td>477149</td>
<td>XM_534339</td>
<td>Cf1 23396</td>
</tr>
<tr>
<td>11.1.99E-05</td>
<td>0.0316</td>
<td>22.5</td>
<td>211.35</td>
<td>0.11</td>
<td>A_11_P000022534</td>
<td>FLRT3</td>
<td>fibronectin leucine rich transmembrane protein 3</td>
<td>477149</td>
<td>XM_534339</td>
<td>Cf1 23396</td>
</tr>
<tr>
<td>12.2.20E-05</td>
<td>0.0316</td>
<td>15.51</td>
<td>88.33</td>
<td>0.18</td>
<td>A_11_P083571</td>
<td>GARNL3</td>
<td>GTPase activating Rap/RanGDP domain-like 3</td>
<td>480718</td>
<td>XM_857841</td>
<td>Cf1 5502</td>
</tr>
<tr>
<td>13.2.21E-05</td>
<td>0.0316</td>
<td>128.41</td>
<td>767.83</td>
<td>0.17</td>
<td>A_11_P000016845</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.2.36E-05</td>
<td>0.0316</td>
<td>10.85</td>
<td>49.83</td>
<td>0.22</td>
<td>A_11_P126166</td>
<td>LOC484934</td>
<td>zinc finger protein 709-like</td>
<td>484934</td>
<td>XM_542050</td>
<td>Cf1 22459</td>
</tr>
<tr>
<td>15.2.85E-05</td>
<td>0.0356</td>
<td>34.38</td>
<td>266.77</td>
<td>0.13</td>
<td>A_11_P189798</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.3.15E-05</td>
<td>0.0369</td>
<td>12.36</td>
<td>55.3</td>
<td>0.22</td>
<td>A_11_P183013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.4.06E-05</td>
<td>0.0447</td>
<td>12.88</td>
<td>122.82</td>
<td>0.1</td>
<td>A_11_P091981</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.4.57E-05</td>
<td>0.0476</td>
<td>20.55</td>
<td>112.49</td>
<td>0.18</td>
<td>A_11_P137766</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.5.10E-05</td>
<td>0.0489</td>
<td>649.06</td>
<td>16.94</td>
<td>38.31</td>
<td>A_11_P060716</td>
<td>LOC507200</td>
<td>Ig lambda chain V-1 region BL2-like</td>
<td>607200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.5.22E-05</td>
<td>0.0489</td>
<td>14.93</td>
<td>64.19</td>
<td>0.23</td>
<td>A_11_P100896</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.6.69E-05</td>
<td>0.0579</td>
<td>10.56</td>
<td>34.83</td>
<td>0.3</td>
<td>A_11_P107231</td>
<td>N1A</td>
<td>nuclear factor 1/A</td>
<td>479552</td>
<td>XM_536691</td>
<td>Cf1 37070</td>
</tr>
<tr>
<td>22.6.93E-05</td>
<td>0.0579</td>
<td>12.31</td>
<td>79.21</td>
<td>0.16</td>
<td>A_11_P183668</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.7.11E-05</td>
<td>0.0579</td>
<td>26.93</td>
<td>188.74</td>
<td>0.14</td>
<td>A_11_P102221</td>
<td>PCDHB14</td>
<td>protocadherin beta 14</td>
<td>487175</td>
<td>XM_844094</td>
<td></td>
</tr>
<tr>
<td>24.8.51E-05</td>
<td>0.0638</td>
<td>14.69</td>
<td>114.93</td>
<td>0.13</td>
<td>A_11_P134231</td>
<td>BVE5</td>
<td>blood vessel epicardial substance</td>
<td>481944</td>
<td>XM_539065</td>
<td>Cf1 37928</td>
</tr>
<tr>
<td>25.9.13E-05</td>
<td>0.0638</td>
<td>34833.8</td>
<td>1114.44</td>
<td>31.26</td>
<td>A_11_P053591</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.9.17E-05</td>
<td>0.0638</td>
<td>152.31</td>
<td>512.06</td>
<td>0.3</td>
<td>A_11_P000016593</td>
<td>TTYH1</td>
<td>twenty homolog 1 (Drosophila)</td>
<td>476386</td>
<td>XM_533589</td>
<td>Cf1 10146</td>
</tr>
<tr>
<td>27.9.20E-05</td>
<td>0.0638</td>
<td>20.16</td>
<td>189.17</td>
<td>0.11</td>
<td>A_11_P000042014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28.1.00E-04</td>
<td>0.0641</td>
<td>12.17</td>
<td>41.73</td>
<td>0.29</td>
<td>A_11_P000013025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29.0.0001019</td>
<td>0.0641</td>
<td>16.66</td>
<td>83.57</td>
<td>0.2</td>
<td>A_11_P083566</td>
<td>GARNL3</td>
<td>GTPase activating Rap/RanGDP domain-like 3</td>
<td>480718</td>
<td>XM_857841</td>
<td>Cf1 5502</td>
</tr>
<tr>
<td>30.0.0001039</td>
<td>0.0641</td>
<td>167.53</td>
<td>812.49</td>
<td>0.21</td>
<td>A_11_P000029082</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3-15. List of top differentially expressed genes between DLG2-null and DLG2-wildtype group**

GARNL3 is highlighted in red boxes.
Figure 3-16. DLG2 re-expression cells show reduced motility in scratch-wound assay.

Representative images of HuO9 and JL-31A cells 24 hrs after scratch-wound (left), and blank area quantification by Image J (right). Experiments were performed twice, with triplicate wells each time. Data were shown as mean±SD (right). Two-tailed student’s t test was performed between DLG2 re-expression cells and control cells. *, p<0.05. **, p<0.01.
Figure 3-17. DLG2 re-expression cells show reduced motility in transwell migration assay.

Representative images of HU-O9 and JL-31A cells 24 hrs after seeding in transwells (left). Both 8µM and 12µM pore-size transwell insert were used. HU-O9 cells were co-infected with GFP expression plasmid and imaged using fluorescence microscope. JL-31A cells were imaged by crystal violet cell staining due to unachievable GFP plasmid infection. Cell quantification is shown on the right. Experiments were performed twice in triplicates each time. Data are shown as mean+SD. Two-tailed student’s t test was performed between DLG2 re-expression cells and control cells. *, p<0.05.
expression of DLG2 in either Hu-O9 or JL-31A cells significantly reduced the number of cells that migrated across the transwell. These phenotypes are in agreement with our gene expression microarray data indicating that DLG2 affects GTPase signaling pathway, which in turn may influence cell migration.

In addition to a role in cell migration, small GTPase proteins such as Rho/Ran, like Ras, are known to promote uncontrolled growth by regulating cell cycle progression (Coleman et al., 2004; Pruitt and Der, 2001). Rho GTPases promote cell cycle by modulating the activity of both positive (promote cyclinD1 gene expression) (Westwick et al., 1997) as well as negative (the cyclin-dependent kinase inhibitors p21CIP1 and p27KIP1) regulators of the Rb tumor suppressor (Adnane et al., 1998; Olson et al., 1998), which functions to regulate G1 progression, and also play an important role in OS tumorigenesis. The other family of small GTPase, Ran, orchestrates multiple stages of the cell cycle for coordinating the onset of the mitosis with S-phase completion, from nucleocytoplasmic transport in interphase, to spindle assembly and nuclear envelope assembly during the M-phase (Clarke and Zhang, 2008; Moore, 2001). We therefore performed cell cycle analysis using flow cytometry (see Chapter 2 Material and Methods). DLG2 re-expressing cells showed significantly more cells in G2/M and/or S phases when grown under low-attachment condition (Figure 3-18). These results combined with slower proliferation in 3D culture conditions (Figure 3-11) suggest that DLG2 re-expressing cells take longer to go through the G2/M or S phase. This also agrees with our observation of reduced colony forming capacity of DLG2 re-expressing cells in soft agar assay (Figure 3-12) and tumor growth in xenograft assay (Figure 3-13). In combination with our gene expression data (Figure 3-14 and Figure 3-15), it is highly likely that the DLG2 tumor suppressor function is mediated at least in part through modulating the small GTPase signaling pathway, such as Rho and Ran.
Figure 3-18. Effect of DLG2 re-expression on cell cycle progression.

JL-31A (A) or Hu-O9 (B) cells were cultured in ultra-low attachment plates. BrdU-incorporation and DNA content were used to separate cells into different cell cycle stages. Representative flow cytometry plots are shown in the left with percentage of each stage is shown in the right. Experiments were performed in triplicates, and at least 50,000 events were collected for every sample. Data are shown as mean±SD. Two-tailed student’s t test was performed between DLG2 re-expression cells and control cells. *, p<0.05.
3.2.2  Conditional Dlg2-null mouse for studying of OS

To validate the role of DLG2 as an OS tumor suppressor in vivo, we set out to examine if Dlg2 deletion would accelerate tumor development in genetically engineered mouse models (GEMM) of this cancer. We selected to create Tp53, Rb and Dlg2 triple-deletion in the osteoblast lineage ($\text{Dlg2}^{\text{ΔOB/ΔOB}}$; $\text{Rb}^{\text{ΔOB/ΔOB}}$; $\text{Tp53}^{\text{ΔOB/ΔOB}}$) for studying OS tumorigenesis.

Dlg2 conditional knock out mice ($\text{Dlg2}^{\text{fl/fl}}$) were acquired from Wellcome Trust Sanger Institute. This particular strain also harbors a lacZ reporter for Dlg2 deletion (see Chapter 2 Material and Methods). This mouse was bred with Col-1a Cre transgenic mice, in which Cre expression is driven under an osteoblast-specific promoter, commonly-used for creating OS (Liu et al., 2004) combined with loss of Rb and p53 (Berman et al., 2008) ($\text{Rb}^{\text{ΔOB/ΔOB}}$; $\text{Tp53}^{\text{ΔOB/ΔOB}}$). In summary, $\text{Dlg2}^{\text{fl/fl}}$ mice were crossed to $\text{Rb}^{\text{ΔOB/ΔOB}}$: $\text{Tp53}^{\text{ΔOB/ΔOB}}$ mice in order to generate $\text{Dlg2}^{\text{ΔOB/+}}$: $\text{Rb}^{\text{ΔOB/+}}$: $\text{Tp53}^{\text{ΔOB/+}}$ mice. These mice were then intercrossed to produce $\text{Dlg2}^{\text{ΔOB/ΔOB}}$: $\text{Rb}^{\text{ΔOB/ΔOB}}$: $\text{Tp53}^{\text{ΔOB/ΔOB}}$ mice as well as their littermate control cohort (Figure 3-19A). LacZ staining of growth plate confirmed that Dlg2 deletion is specific (Figure 3-20A), as $\text{Dlg2}^{\text{ΔOB/ΔOB}}$: $\text{Rb}^{\text{ΔOB/ΔOB}}$: $\text{Tp53}^{\text{ΔOB/ΔOB}}$ growth plate exhibited strong blue staining. Kidney and Large intestine are non-specific artifacts commonly seen in Dlg2 staining (Figure 3-20B). The resulting $\text{Dlg2}^{\text{ΔOB/ΔOB}}$: $\text{Rb}^{\text{ΔOB/ΔOB}}$: $\text{Tp53}^{\text{ΔOB/ΔOB}}$ mice are viable, but develop early onset OS with complete penetrance. The tumors that developed spontaneously display many of the characteristics of human OS, including a high degree of metastatic.

As shown in Figure 3-19B, $\text{Dlg2}^{+/+}$: $\text{Rb}^{\text{ΔOB/ΔOB}}$: $\text{Tp53}^{\text{ΔOB/ΔOB}}$ mice started to die at around 200 days of age, whereas additional loss of Dlg2 further shorten the overall survival in the $\text{Dlg2}^{\text{ΔOB/ΔOB}}$: $\text{Rb}^{\text{ΔOB/ΔOB}}$: $\text{Tp53}^{\text{ΔOB/ΔOB}}$ mice (Figure 3-19B).
Figure 3-19. Osteoblast-specific Dlg2 KO accelerates OS tumor growth in mice.

(A) Breeding strategy for generating osteoblast-specific Dlg2 deletion in the Tp53- and Rb1-null background.

(B) Dlg2 deletion further shortens overall survival of RbΔOB/ΔOB, Tp53ΔOB/ΔOB mice (log rank test; P-value = 0.0003).

(C) Representative X-ray and (D) MicroCT images for 22-week-old mice are shown.

(E) Representative images of tumors on rib (left), tibia (right top), and femur (right bottom).

(F) H&E staining of representative tumors.
Figure 3-20. Whole body LacZ profiling of Dlg2 deleted, LacZ-expressing mice under ColIa-Cre promoter.

(A) LacZ staining of femoral bone close to growth plate. $\text{Dlg}^{2+/+}$: $\text{Rb}^{\alpha\text{OB}/\alpha\text{OB}}$: $\text{Tp53}^{\alpha\text{OB}/\alpha\text{OB}}$ on top panel and $\text{Dlg}^{\alpha\text{OB}/\alpha\text{OB}}$: $\text{Rb}^{\alpha\text{OB}/\alpha\text{OB}}$: $\text{Tp53}^{\alpha\text{OB}/\alpha\text{OB}}$ on bottom panel.

(B) LacZ staining and scoring of various body parts evaluated by pathologist. -, no lacZ staining. +, positive lacZ staining.
X-ray (Figure 3-19C) and MicroCT (Figure 3-19D) images of $Dlg2^{ΔOB/ΔOB}: Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}$ mice at 22-weeks showed significant primary OS as well as metastasis to the lung. Abundant osteoid deposition shown in the H&E stained tumor confirmed that the tumor was of osteoblastic nature (Figure 3-19E and F).

$Dlg2^{ΔOB/ΔOB}: Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}$ tumors had significantly more Ki-67 positive cells, indicating greater proliferative activity (Figure 3-21). This agreed with accelerated tumor growth and shorter survival time. Also, this is in concordance with in vitro proliferation and flow cytometry data where DLG2 re-expression reduced proliferation and prolonged cell-cycle progression.

To further investigate the effect of $Dlg2$ deletion, expression profile in 6 $Dlg2^{ΔOB/ΔOB}: Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}$ tumors and 4 $Dlg2^{+/+}: Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}$ tumors was analyzed using RNA-Seq and differentially expressed genes were determined in these two groups (Figure 2-3). Many significantly differentially expressed genes were noted between the two groups (Table 3-3), including some known oncogenes and tumors suppressors. These included LSAMP (Barøy et al., 2014; Kresse et al., 2009) and RECK (Noda et al., 2003; Xu et al., 2010), small GTPase pathway related genes, and SEMA family members, which is a protein family thought to regulate small GTPase signaling pathway (Neufeld and Kessler, 2008; Yazdani and Terman, 2006). Top altered pathways indicated by Gene Ontology analysis also showed that small GTPase signaling pathways are altered in $Dlg2^{ΔOB/ΔOB}: Rb^{ΔOB/ΔOB}: Tp53^{ΔOB/ΔOB}$ tumors (Table 3-4), again indicating that Dlg2 could exert its tumor suppressive effects through the small GTPase pathway.

Taken together, through cross-species oncogenomics, we identified $DLG2$ to be deleted in both spontaneously human and canine OS. Multiple in vitro and in vivo assays using human and
canine cell lines showed that DLG2 re-expression reduced tumoriogenicity and motility. Using GEMM, we found that Dlg2 deletion accelerates tumor development and reduces survival in mice. RNA-seq data in $Dlg2^{\Delta OB/\Delta OB}$, $Rb^{\Delta OB/\Delta OB}$, $Tp53^{\Delta OB/\Delta OB}$ tumors suggested that Dlg2 likely exerts its tumor suppressive effects through the small GTPase signaling.
Figure 3-21. Ki-67 staining of mouse OS

Representative images of invasive front of OS tumors stained with Ki-67.
Table 3-3. Top differentially expressed genes by RNA-seq.

<table>
<thead>
<tr>
<th>#Gene</th>
<th>Locus</th>
<th>Sample1</th>
<th>Sample2</th>
<th>FPKM1</th>
<th>FPKM2</th>
<th>Fold_change</th>
<th>log2(fold_change)</th>
<th>STAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSAMP</td>
<td>chr3:115521209-116164385</td>
<td>++</td>
<td>FF</td>
<td>0.622</td>
<td>0.108</td>
<td>0.17</td>
<td>-2.52</td>
<td>DOWN</td>
</tr>
<tr>
<td>RECK</td>
<td>chr9:36036909-36124452</td>
<td>++</td>
<td>FF</td>
<td>1.487</td>
<td>0.300</td>
<td>0.20</td>
<td>-2.31</td>
<td>DOWN</td>
</tr>
<tr>
<td>Ran</td>
<td>chr5:129020155-129024321</td>
<td>++</td>
<td>FF</td>
<td>111.426</td>
<td>170.398</td>
<td>1.53</td>
<td>0.61</td>
<td>UP</td>
</tr>
<tr>
<td>Rhod</td>
<td>chr19:4425456-4477143</td>
<td>++</td>
<td>FF</td>
<td>8.966</td>
<td>14.119</td>
<td>1.57</td>
<td>0.66</td>
<td>UP</td>
</tr>
<tr>
<td>Rhof</td>
<td>chr5:123118179-123132629</td>
<td>++</td>
<td>FF</td>
<td>2.154</td>
<td>3.559</td>
<td>1.65</td>
<td>0.72</td>
<td>UP</td>
</tr>
<tr>
<td>Rhog</td>
<td>chr7:102239122-102250118</td>
<td>++</td>
<td>FF</td>
<td>37.601</td>
<td>59.426</td>
<td>1.58</td>
<td>0.66</td>
<td>UP</td>
</tr>
<tr>
<td>Rhq</td>
<td>chr17:86963110-87025401</td>
<td>++</td>
<td>FF</td>
<td>27.962</td>
<td>13.624</td>
<td>0.49</td>
<td>-1.04</td>
<td>DOWN</td>
</tr>
<tr>
<td>Rou</td>
<td>chr8:123653938-123663880</td>
<td>++</td>
<td>FF</td>
<td>11.601</td>
<td>5.861</td>
<td>0.51</td>
<td>-0.98</td>
<td>DOWN</td>
</tr>
<tr>
<td>Rhov</td>
<td>chr2:11926920-119271926</td>
<td>++</td>
<td>FF</td>
<td>0.661</td>
<td>1.452</td>
<td>2.20</td>
<td>1.14</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap1</td>
<td>chr2:91650117-91672320</td>
<td>++</td>
<td>FF</td>
<td>13.348</td>
<td>20.311</td>
<td>1.52</td>
<td>0.61</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap11a</td>
<td>chr2:113831491-113848661</td>
<td>++</td>
<td>FF</td>
<td>5.649</td>
<td>9.557</td>
<td>1.69</td>
<td>0.76</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap15</td>
<td>chr2:43748823-43987623</td>
<td>++</td>
<td>FF</td>
<td>3.002</td>
<td>6.919</td>
<td>2.30</td>
<td>1.20</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap19</td>
<td>chr19:41766587-41802084</td>
<td>++</td>
<td>FF</td>
<td>2.980</td>
<td>8.110</td>
<td>2.72</td>
<td>1.44</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap23</td>
<td>chr11:97450159-97502400</td>
<td>++</td>
<td>FF</td>
<td>9.785</td>
<td>17.128</td>
<td>1.75</td>
<td>0.81</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap30</td>
<td>chr1:7138859-171410239</td>
<td>++</td>
<td>FF</td>
<td>4.454</td>
<td>9.832</td>
<td>2.21</td>
<td>1.14</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap33</td>
<td>chr7:39022225-39035009</td>
<td>++</td>
<td>FF</td>
<td>0.611</td>
<td>1.390</td>
<td>2.27</td>
<td>1.18</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap4</td>
<td>chr7:3892101-73911298</td>
<td>++</td>
<td>FF</td>
<td>2.876</td>
<td>6.015</td>
<td>2.09</td>
<td>1.06</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgap44</td>
<td>chr11:6479034-65162961</td>
<td>++</td>
<td>FF</td>
<td>3.213</td>
<td>1.853</td>
<td>0.58</td>
<td>-0.79</td>
<td>DOWN</td>
</tr>
<tr>
<td>Arhgap9</td>
<td>chr10:127323736-127341579</td>
<td>++</td>
<td>FF</td>
<td>6.047</td>
<td>15.870</td>
<td>2.62</td>
<td>1.39</td>
<td>UP</td>
</tr>
<tr>
<td>Arhdib</td>
<td>chr6:136923708-136941717</td>
<td>++</td>
<td>FF</td>
<td>29.639</td>
<td>140.289</td>
<td>4.73</td>
<td>2.24</td>
<td>UP</td>
</tr>
<tr>
<td>Arhdig</td>
<td>chr17:26199182-26201350</td>
<td>++</td>
<td>FF</td>
<td>3.260</td>
<td>15.484</td>
<td>4.75</td>
<td>2.25</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgef18</td>
<td>chr8:3393007-3456600</td>
<td>++</td>
<td>FF</td>
<td>3.863</td>
<td>6.674</td>
<td>1.73</td>
<td>0.79</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgef19</td>
<td>chr4:141242883-141257562</td>
<td>++</td>
<td>FF</td>
<td>10.593</td>
<td>17.400</td>
<td>1.64</td>
<td>0.72</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgef39</td>
<td>chr4:43496143-43499660</td>
<td>++</td>
<td>FF</td>
<td>4.694</td>
<td>12.467</td>
<td>2.66</td>
<td>1.41</td>
<td>UP</td>
</tr>
<tr>
<td>Arhgef9</td>
<td>chr9:95048943-95166486</td>
<td>++</td>
<td>FF</td>
<td>1.240</td>
<td>0.485</td>
<td>0.39</td>
<td>-1.35</td>
<td>DOWN</td>
</tr>
<tr>
<td>Ranbp1</td>
<td>chr16:18239978-18248694</td>
<td>++</td>
<td>FF</td>
<td>102.378</td>
<td>182.105</td>
<td>1.78</td>
<td>0.83</td>
<td>UP</td>
</tr>
<tr>
<td>Ranbp10</td>
<td>chr8:105768307-105827350</td>
<td>++</td>
<td>FF</td>
<td>7.855</td>
<td>15.959</td>
<td>2.03</td>
<td>1.02</td>
<td>UP</td>
</tr>
<tr>
<td>Ranbp3l</td>
<td>chr15:9055681-9067333</td>
<td>++</td>
<td>FF</td>
<td>0.131</td>
<td>1.399</td>
<td>10.67</td>
<td>3.42</td>
<td>UP</td>
</tr>
<tr>
<td>Ranbp6</td>
<td>chr19:29808107-29812974</td>
<td>++</td>
<td>FF</td>
<td>2.739</td>
<td>1.484</td>
<td>0.54</td>
<td>-0.88</td>
<td>DOWN</td>
</tr>
<tr>
<td>Ranbp9</td>
<td>chr13:43938375-43480973</td>
<td>++</td>
<td>FF</td>
<td>20.747</td>
<td>13.543</td>
<td>0.65</td>
<td>-0.62</td>
<td>DOWN</td>
</tr>
<tr>
<td>Rangap1</td>
<td>chr15:81704247-81729919</td>
<td>++</td>
<td>FF</td>
<td>18.653</td>
<td>33.646</td>
<td>1.80</td>
<td>0.85</td>
<td>UP</td>
</tr>
<tr>
<td>Rap1gap2</td>
<td>chr11:74383482-74590158</td>
<td>++</td>
<td>FF</td>
<td>4.117</td>
<td>1.496</td>
<td>0.36</td>
<td>-1.46</td>
<td>DOWN</td>
</tr>
<tr>
<td>Rapgef1</td>
<td>chr2:29619719-29740363</td>
<td>++</td>
<td>FF</td>
<td>33.270</td>
<td>18.390</td>
<td>0.55</td>
<td>-0.86</td>
<td>DOWN</td>
</tr>
<tr>
<td>Rapgef15</td>
<td>chr12:117516478-117756978</td>
<td>++</td>
<td>FF</td>
<td>1.962</td>
<td>2.980</td>
<td>1.52</td>
<td>0.60</td>
<td>UP</td>
</tr>
<tr>
<td>SEMA3D</td>
<td>chr7:84624871-84751247</td>
<td>++</td>
<td>FF</td>
<td>3.773</td>
<td>1.619</td>
<td>0.43</td>
<td>-1.22</td>
<td>DOWN</td>
</tr>
<tr>
<td>SEMA3G</td>
<td>chr5:25467267-25479043</td>
<td>++</td>
<td>FF</td>
<td>1.283</td>
<td>0.241</td>
<td>0.19</td>
<td>-2.41</td>
<td>DOWN</td>
</tr>
<tr>
<td>SEMA6D</td>
<td>chr15:47476402-48066420</td>
<td>++</td>
<td>FF</td>
<td>24.603</td>
<td>9.212</td>
<td>0.37</td>
<td>-1.42</td>
<td>DOWN</td>
</tr>
<tr>
<td>SEMA7A</td>
<td>chr15:74701629-74726299</td>
<td>++</td>
<td>FF</td>
<td>96.657</td>
<td>54.686</td>
<td>0.57</td>
<td>-0.82</td>
<td>DOWN</td>
</tr>
</tbody>
</table>

Orange region shows known oncogenes and tumors suppressors to be differentially expressed in
\[ \text{Dlg2}^{\Delta\text{OB}/\Delta\text{OB}} : \text{Rb}^{\Delta\text{OB}/\Delta\text{OB}} : \text{Tp53}^{\Delta\text{OB}/\Delta\text{OB}} \]
vs
\[ \text{Dlg2}^{+/+} : \text{Rb}^{\Delta\text{OB}/\Delta\text{OB}} : \text{Tp53}^{\Delta\text{OB}/\Delta\text{OB}} \]
tumors. Green region shows the change of small GTPase pathway related genes the two group. Blue region lists SEMA family members found to be under-expressed in
\[ \text{Dlg2}^{\Delta\text{OB}/\Delta\text{OB}} : \text{Rb}^{\Delta\text{OB}/\Delta\text{OB}} : \text{Tp53}^{\Delta\text{OB}/\Delta\text{OB}} \]
Table 3-4. Top altered gene ontology pathways identified by RNA-seq in Dlg2-/− tumor.

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>RT</th>
<th>Genes count</th>
<th>%</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Protein dimerization activity</td>
<td>RT</td>
<td>14</td>
<td>7.2</td>
<td>4.2E-3</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Small GTPase regulator activity</td>
<td>RT</td>
<td>9</td>
<td>4.6</td>
<td>7.9E-3</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>GTPase regulator activity</td>
<td>RT</td>
<td>11</td>
<td>5.7</td>
<td>9.5E-3</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Nucleoside-triphosphatase regulator activity</td>
<td>RT</td>
<td>11</td>
<td>5.7</td>
<td>1.1E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Ras GTPase binding</td>
<td>RT</td>
<td>5</td>
<td>2.6</td>
<td>1.5E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Transcription repressor activity</td>
<td>RT</td>
<td>9</td>
<td>4.6</td>
<td>1.7E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Small GTPase binding</td>
<td>RT</td>
<td>5</td>
<td>2.6</td>
<td>2.1E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Protein homodimerization activity</td>
<td>RT</td>
<td>9</td>
<td>4.6</td>
<td>2.3E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>GTPase binding</td>
<td>RT</td>
<td>5</td>
<td>2.6</td>
<td>2.7E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Transcription activator activity</td>
<td>RT</td>
<td>10</td>
<td>5.2</td>
<td>2.7E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Single-stranded RNA binding</td>
<td>RT</td>
<td>3</td>
<td>1.5</td>
<td>3.2E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Transmembrane receptor protein tyrosine kinase activity</td>
<td>RT</td>
<td>4</td>
<td>2.1</td>
<td>3.3E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Transcription regulator activity</td>
<td>RT</td>
<td>24</td>
<td>12.4</td>
<td>3.9E-2</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>Enzyme binding</td>
<td>RT</td>
<td>11</td>
<td>5.7</td>
<td>4.6E-2</td>
</tr>
</tbody>
</table>
Chapter 4

Discussion
4.1 Oncogenomics

Copy number alteration is one of the many ways in which gene expression and function can be modified. Genomic studies have demonstrated that OS has a relatively high rate of CIN including CNVs and SVs. aCGH is a method that combines the principles of CGH with the use of microarrays. This method greatly increases the resolution of screening for genomic CNAs. The human OS genomic datasets we selected from GEO were obtained from different aCGH platforms, including Agilent-013282 Human Genome CGH Microarray 44B (approximately 43,000 probes spanning both coding and noncoding sequences with an 43kb overall median probe spacing), Agilent-014693 Human Genome CGH Microarray 244A (approximately 240,000 probes spanning both coding and noncoding sequences with an 8.9kb overall median probe spacing), Affymetrix Human Genome U133 Plus 2.0 Array (more than 54,000 probes sets and 1.3 million distinct oligonucleotide covering over 47,000 transcripts and variants), and Agilent-012391 Whole Human Genome Oligo Microarray G4112A (44K microarray queries ~41,000 human genes and transcripts) (see detailed information on Agilent and Affymetrix websites). The aCGH analysis for canine OS tumor samples was generated in Affymetrix Canine Genome 2.0 array platform, which contains ~43,000 canine probe sets and covers over 18,000 canine transcripts as well as over ~20,000 non-redundant canine gene predictions. Therefore, the datasets analyzed in this study represent high resolution and high sensitivity for CNA analysis.

However, identification of driver genes in regions of CNA is challenging. If narrow regions with highly elevated copy number or total deletion that contain previously known cancer-related genes, high-probability candidate genes could be immediately identified. As expected, in our aCGH analysis, we successfully identified CNV regions with known oncogene (MYC) and tumor suppressors (PTEN and RB1). However, in many cases, even minimally defined CNA regions
resulting from combining data from many specimens may contain several or none potential candidates, or the genomic structure of CNA may be complex, making it difficult to determine how many different loci are under selection.

Examination of gene expression at the RNA or protein level is essential for candidate validation. If a gene is a target in a region with copy number gain, it should be overexpressed with elevated mRNA or protein level. However, if the region contains several genes, they may all be overexpressed, which makes it difficult to separate driver gene from other putative passenger genes. Evaluations of genes in regions with copy number losses are also complicated. Deletion of copies of a genomic region is easily detected in more homogenous population such as cell lines, but their reliable detection in tumor specimens is depended on the percentage of inclusion of normal cells, or tumor content. aCGH can adequately detect focal homozygous deletion of the whole gene. In our study, DLG2 gene was identified in a focal deletion region, and its level was confirmed to be diminished by qPCR method. However, the function of some tumor suppressor genes can be totally abrogated by deletion of only one copy, with point mutation or epigenetic alteration of the other copy resulting in LOH, and no further copy number alteration, a change not detectable by aCGH method. From COSMIC database, DLG2 was mutated in 11% of multiple human tumors by point mutations. Therefore, we may miss OS tumor samples with DLG2 point mutations by analyzing the CNA data only, and under estimate the importance of DLG2 in OS tumorigenesis. As next generation sequencing (NGS) datasets, such as exome or whole genome sequencing, for OS tumors are becoming publicly available, the combination of NGS data for point mutations and aCGH data my help us to identify more genes involved in OS tumorigenesis.
The type of microarray plateforms used in our aCGH analysis as well as the analysis pipeline have a bias on genes, or coding regions, expect the Agilent Human Genome CGH Microarray chip 44B and 244A, which contains probes spanning both coding and noncoding sequences. Studies have shown that CNAs in regulatory elements of the gene, such as distant-acting gene enhancer sequences in the non-coding regions can contribute to disease (Heintzman and Ren, 2009; Kloppocki and Mundlos, 2011; Savarese et al., 2012). Therefore, it will be interesting to determine whether recurrent CNAs in such regulatory elements can also contribute to OS pathogenesis.

Overall, aCGH is a useful top-down approach. On the other hand, the more traditional studies focused on individual functional units also uncover important details that are not revealed by global analytical approaches. Improving the ability to integrate genome-wide study with detailed selective analysis is essential for discovery research.

4.2 Cross-species comparative oncogenomics

Another approach used in this thesis to identify the potential dysfunctional focal CNAs is the cross-species comparative genomic analysis. The genome of the domesticated dog, a close evolutionary relation to human, is a powerful new tool for understanding the human genome (Lindblad-Toh et al., 2005). As discussed in the Introduction section, naturally occurring canine OS serves as an idea system in many respects for understanding the pathogenesis of human OS. Indeed, this study uncovers the same top altered genes between human and canine OS samples. DLG2 deletion was subsequently validated in OS samples for both species. DLG2 functional validation through a number of assays showed similar effects in both human and canine OS cell lines, and finally in the mouse model.
The primary difference between human and canine OS is the age of development. Human OS occurs most commonly in children and adolescents, whereas canine OS tends to occur in middle-aged to older dogs. Another current drawback is that canine genome annotation is not as good as the human genome, although it has improved considerably in recent years. This poor annotation results in losing important information while performing comparative oncogenomics.

4.3 Tumor suppressor function of DLG2 in human cancer

In this study, we identified and validated DLG2 as a novel tumor suppressor in osteosarcoma. Recent study analyzed recurrent somatic structural variations in pediatric OS and pointed to DLG2 gene showing recurrent somatic alterations in 34 human pediatric OS tumors by whole-genome sequencing (Chen et al., 2014b). No functional analysis was undertaken.

The protein encoded by DLG2 gene is a multiple PDZ domain-containing protein that belongs to the family of molecular scaffolding proteins known as MAGUKs (Kim et al., 1996; Mazoyer et al., 1995). Members of the MAGUK family are defined by a basic core of three distinct interaction domains: the PSD95/DLG/ZO-1 (PDZ) domain, an SH3 domain, and the C-terminal enzymatically inert GK domain (Oliva et al., 2012b). PDZ domains typically recognize specific signature sequence at the extreme C-termini of target proteins, but some also recognize the internal polypeptides of target proteins, lipids, or heterodimerize with themselves (Harris and Lim, 2001). SH3 domains usually bind to proline rich sequences (PXXP) in target proteins (Kaneko et al., 2008), but binding partners for MAGUK’s SH3 domains have not been reported. Instead, it has been shown that the SH3 and GK domains of MAGUKs interact with one another in an intramolecular fashion, which blocks the recognition and prevents SH3 domain from interacting with PXXP motif in other proteins (McGee and Bredt, 1999; McGee et al., 2001b). Some members of the DLG subfamily also contain a positively charged HOOK domain that
mediates interaction with proteins of the ERM family that are linked to the actin cytoskeleton (Lue et al., 1996). Different MAGUK family members can also form intermolecular interaction through these domains. Through such protein-protein interactions, MAGUKs play important role in the organization of cell-cell adhesion sites including neurological synapses and tight junctions of epithelial cells (González-Mariscal et al., 2000; Oliva et al., 2012b). Therefore, loss of MAGUK family members in osteoblast may influence its cytoskeleton structure as well as the organization of cell-cell adhesion, which in turn affect its growth and migration.

In Drosophila, DLG is a tumor suppressor, although a clear tumor suppressor function has yet to be confirmed for $DLG2$ or other DLG family members in human cancer. Majority of studies on $DLG2$ or PSD-93 have focused on its role in synapse function (Kruger et al., 2013; Sun and Turrigiano, 2011). In discrete neuronal populations, DLG2 forms a heterodimer with a related family member, such as DLG4/PSD-95, that may interact at postsynaptic sites to form a multimeric scaffold for the clustering of receptors, ion channels, and associated signaling proteins. Although COSMIC database reports several DLG2 mutations in a variety of human cancers, our studies provides the first evidence that DLG2 can function as a tumor suppressor in a human cancer. More detailed studies are required to investigate the mechanism for DLG2 function in normal development and in cancer.

In our study, $DLG2$ loss was identified in 29% of human OS patient samples, but not the rest of OS tumors. One possibility may be that this deletion involves in tumor progression rather than initiation. Additional analysis of patient samples being suggested by conveying an inferior prognosis or being associated with metastases would provide evidence on whether this is the case. Another possibility is that OS is a highly heterogeneous disease and different subsets of patients have distinct initiating events, which need to be further investigated.
4.4 Conditional *Dlg2* KO mouse model

The role of *DLG2* has not been previously studied in osteosarcoma, and genetically engineered mouse models provide a unique opportunity to investigate its importance in osteosarcoma tumorigenesis *in vivo*.

Studies have shown that if *Dlg2* was deleted in the whole body (*Dlg2<sup><i>tm1Ds</i></sup>*) (Gerdin, 2010), homozygous mutant mice exhibit several abnormalities (McGee et al., 2001a). Both sexes had atypical indirect calorimetry and DEXA parameters. Females also had decreased body weight, decreased circulating HDL cholesterol levels, and increased susceptibility to bacterial infection. No cancer related phenotypes were reported in these mice.

In order to study the function of DLG2 in OS tumorigenesis, we generated an osteoblastic-specific *Dlg2* deletion under the *Col-1a* promoters, and combined its deletion with that of *p53* and *Rb* in a tissue specific manner. We observed that *Dlg2* loss significantly shortens the overall survival of *Rb<sup>ΔOB/ΔOB</sup>*: *Tp53<sup>ΔOB/ΔOB</sup>* mice. Considering that *DLG2* deletion leads to promotion of OS cell migration and invasion *in vitro*, *Dlg2* knockout may promote tumor metastasis and generate more aggressive tumors in our mouse model.

However, the *Dlg2* deletion was restricted to osteoblasts in our mouse model, which precludes assessments of whether DLG2 can function as tumor suppressor in other cancer types. Considering that *Rb* deletion is embryonic lethal (Clarke et al., 1992; Jacks et al., 1992; Lee et al., 1992), a *p53* and *Dlg2* double knockout mouse could be generated to evaluate Dlg2's function in other cancers.
Chapter 5

Summary and Future Directions
5.1 Summary and Key Findings

OS is the most common primary malignancy of bone. Although several mutated genes have been identified, the molecular pathogenesis of OS and its metastasis mechanism are still largely unknown. Therefore, a comprehensive understanding of the molecular pathogenesis of OS is needed in order to identify more therapeutic targets and manage the disease. Genomic studies have demonstrated that OS has a relatively high rate of chromosomal abnormalities. However, these CNVs and SVs in osteosarcoma have made it difficult to distinguish driver mutations from passenger mutations. In this thesis, these issues were addressed by utilizing cross-species oncogenomics between human OS tumors and naturally occurring canine OS tumors, as well as performing a series of functional validation studies in both in vitro and in vivo systems.

We firstly optimized an aCGH analysis pipeline for datasets, and then carefully selected 7 published human OS genomic datasets from GEO for cross-species genomics analysis. Canine OS tumor genomic datasets generated by the Wood and Khokha laboratories collaboration were incorporated. Recurrent CNAs across all samples were identified including in regions that cover known oncogenes and tumor suppressors, as well as in novel regions. We identified a new focal deletion in multiple human and canine OS samples which contains the gene DLG2. This observation was confirmed in an independent cohort of human OS samples, which indicated a similar deletion frequency, suggesting that DLG2 is a potential tumor suppressor gene in OS.

We then undertook validation of DLG2 as a tumor suppressor in biological systems. Using multiple human and canine cell lines that naturally harbored DLG2 deletion, we observed that DLG2 re-expression reduced tumorigenic characteristics in 3D-culture, soft agar assays, as well as in mouse xenograft assays. Dlg2 re-expression also reduced cell motility as tested by scratch-
wound and transwell migration assay. In addition, Dlg2 re-expression altered cell cycle kinetics and possibly lengthened the G2/M phase. Through integrative-genomics in canine OS samples (aCGH and expression microarrays), we found GARNL3, a small GTPase regulator to be differentiated-expressed in Dlg2-null vs Dlg2-wildtype tumors. Using GEMM, we found that Dlg2 deletion accelerated tumor development and reduced survival in mice. Finally, through RNA-seq, we found that $D_{lg2}^{\Delta OB/\Delta OB} \colon R_{b}^{\Delta OB/\Delta OB} \colon T_{p53}^{\Delta OB/\Delta OB}$ mouse tumors could represent a distinct group encompassing aberrant expression of a variety of small GTPase regulators, including Ran and Rho family members. It is possible that Dlg2 exerts its tumor suppressive effect through the activation of small GTPase pathway.

Taken together, in this thesis, we identify and functionally validate a novel tumor suppressor called DLG2 in OS. The principles of this study are widely applicable to other human cancers, offering a way to not only short potential new candidate cancer driver genes residing within chromosomal alterations, but also to fast track their biological investigation.

5.2 Future Directions

5.2.1 Role of DLG2 as a scaffold protein in bone cells

As a multi-domain scaffold protein, it will be interesting to uncover the binding partners of DLG2 in osteoblasts, and their downstream signaling pathways. We will then have a better understanding of the molecular pathways that are disrupted upon $DLG2$ loss during OS tumorigenesis. To this end, osteoblasts can be prepared from mice, and subjected to co-immunoprecipitation against DLG2. Some of currently known DLG family binding partners could be tested by western blotting with specific antibodies. Potential candidates are APC (Matsumine et al., 1996), beta-Catenin (Subbaiah et al., 2012c), MAPK12 (Sabio et al., 2010),
MEK2 (Gaudet et al., 2011), NET1 (Carr et al., 2009), NEDD4 (Van Campenhout et al., 2011), PTEN (Cotter et al., 2010) and SGEF (Subbaiah et al., 2012b), among which NET1 and SGEF are known GEFs and regulate the activation of Rho GTPase. Alternatively, mass spectrometry analysis could offer an approach for co-immunoprecipitation based identification of unknown osteoclast-specific DLG2 binding partners.

5.2.2 The role of small GTPs in DLG2-deficiency derived osteosarcoma

Through integrative-genomics in canine OS samples and RNA-seq in our $Dl g_2^{AB/OB}$, $Rb^{AB/OB}$, $Tp 53^{AB/OB}$ mouse tumors, we identified aberrant expression of a variety of small GTPase regulators, including regulators for Ran and Rho family members. Therefore, it will be worthy to follow-up on the function of these small GTPases in OS, especially Ran and Rho. Our preliminary results showed that Ran or Rho could be effectively knocked out in osteoblast cells from $Dl g_2^{+/+}$, $Rb^{AB/OB}$, $Tp 53^{AB/OB}$ and $Dl g_2^{AB/OB}$, $Rb^{AB/OB}$, $Tp 53^{AB/OB}$ mouse models by CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats/ CRISPR-associated protein-9 nuclease) technology (Sander and Joung, 2014). These models can further facilitate examination of the role of small GTPs in normal and DLG2-deficient OS.

5.2.3 Other candidate genes in significant changed CNVs identified by this study

Using large datasets of aCGH and cross-species comparative genomics of human and canine OS, we have identified a number of recurrent focal amplifications and deletions. Besides the regions containing known oncogenes and tumor suppressors, and the DLG2 gene identified in the novel focal deletion in the Chr11:q14 region, we shortlisted a number of recurrent CNAs for further
study. This could lead to discovery of new genes involved in OS tumorigenesis and enhance our understanding of the pathogenesis of OS, possibly forming the basis of new therapy.

5.3 Concluding Remarks

In summary, we find that DLG2 is frequently deleted in spontaneous human and canine OS through cross-species genomics analysis. Using osteoblast-specific knock-out of Dlg2 in mouse model and a variety of biochemical and molecular approaches, we demonstrate DLG2 function as a tumor suppressor in OS. Further in-depth molecular studies are required to understand the signaling pathways downstream of DLG2, and its precise effect on osteoblast transformation.
References


constitutional mutations in the p53 gene among 21 Li-Fraumeni families. Cancer Res. 54, 1298–1304.


variation in the human genome. Genome Res. 16, 1566–1574.


Scionti, I., Michelacci, F., Pasello, M., Hattinger, C.M., Alberghini, M., Manara, M.C., Bacci,


Wallace, M.D., Pfefferle, A.D., Shen, L., McNairn, A.J., Cerami, E.G., Fallon, B.L., Rinaldi,


genome evolution from complex chromosomal rearrangements. Genes Dev. 27, 2513–2530.
