THE ROLE OF LONG NON-CODING RNAS IN PEDIATRIC CANCERS: NEUROBLASTOMA AND OSTEOSARCOMA

ARABA PANYIN SAGOE-WAGNER
Bachelor of Science, Biological Sciences, University of Lethbridge, 2017

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

BIOLOGICAL SCIENCES

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

THE ROLE OF LONG NON-CODING RNAS IN PEDIATRIC CANCERS: NEUROBLASTOMA AND OSTEOSARCOMA

ARABA PANYIN SAGOE-WAGNER

Date of Defense: April 23, 2020

Dr. Olga Kovalchuk Supervisor	Professor	Ph.D, M.D.
Dr. Robbin Gibb Thesis Examination Committee Member	Professor	Ph.D.
Dr. Aru Narendran Thesis Examination Committee Member	Professor	Ph.D., M.D.
Dr. Athanasios Zovoilis Chair, Thesis Examination Committee	Assistant Professor	Ph.D., M.D.

DEDICATION

To my parents, who believed in seeking the best opportunities for us,

My twin, who supported me in more ways than she could have figured out,

and

To my husband, who kept me motivated.

ABSTRACT

Emerging studies into long non-coding RNAs in cancers continue to implicate these molecules in roles crucial to tumourigenesis and progression of cancers. Long non-coding RNAs are transcripts greater than 200bp that do not code for protein but have been found to regulate gene expression.

Given the extensive knowledge of mutations in the genome that could account for pediatric cancers in children, there is a need for a greater understanding of epigenetic regulation and the epigenome of osteosarcoma and neuroblastoma. LncRNAs contribute to epigenetics, thus these transcripts could also influence the cancers via epigenetic mechanisms. This thesis therefore aims to identify differentially expressed long non-coding RNA in neuroblastoma and osteosarcoma that may have roles which could serve as potential targets in therapeutics.

Our findings reveal long non-coding RNAs *LINC00261, LINC01133, LINC01268* and *LINC01139* that are differentially expressed and could have potential roles in the progression of the disease.

ACKNOWLEDGEMENT

Getting this document together has been the longest task yet I have achieved, and though it has taken me a while, I look forward to the next challenge.

I am grateful to my parents Mr. and Mrs. Sagoe, for their support and encouragement over the years. Being far from home is always tough, but they made it seem like a breeze with their consistent chats of encouragements and making the long trips to spend a few days with me. Without their encouragement and prayers, things would have taken a lot longer. To my twin, Araba Kakraba Sagoe (some moments call for the full name), you have been invaluable in finishing this paper and my experiments. You may not have contributed physically, but your emotional support was tremendous, and for that I am forever grateful. And to my dear friend, Suzanne, thanks for being there and lending a hand.

To my supervisor, Dr. Olga Kovalchuk, I am glad you took me under your wings to do this research. To my committee members Drs. Gibb and Narendran, thank you for answering my questions, and being available to reach out to.

I would also like to express gratitude to Drs. Bo Wang and Dongping Li – you have been the best resources while I was undertaking my project. Rommy and Rocio, thank you for being available for me to bounce my ideas off of, and being encouraging in times when things did not seem to be working out. Dr. Slava Ilyntskyy – thanks for all your help with the bioinformatics and taking the time to answer my many questions.

I am also grateful to the collaborators from POETIC (Pediatric Oncology Experimental Therapeutics Consortium) for providing the samples for my project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANRIL – antisense non-coding RNA INK4 locus
CCAT1 – colon cancer associated transcript
CNS – central nervous system
DNA – deoxyribonucleic acid
EMT – epithelial-mesenchymal-transition
EZH2 – enhancer of zeste homolog 2
FA – fanconi anemia
FAP – familial adenomatous polyposis
HAT – histone acetyltransferase
HDAC – histone deacetylase
HMT – histone methylase
LFS – li-fraumeni syndrome
Linc – long intergenic non-coding RNA
LncRNA – long non-coding RNA
LOH – loss of heterozygosity
MALAT1– metastasis-associated lung adenocarcinoma transcript 1
NB – neuroblastoma
NcRNA – non-coding RNA
NF1– neurofibromatosis 1
NFRs – nucleosome free regions
OS – osteosarcoma
PRC2 – polycomb repressor complex 2
RNA – ribonucleic acid
SOC – serious ovarian cancer
SRA – steroid receptor activator

CHAPTER 1: GENERAL INTRODUCTION

1.0. Pediatric Cancers

Pediatric cancers are one of the world's leading causes of death in children. In developing countries, pediatric cancers are the leading cause of death by disease among children over one year of age. Most pediatric tumours are likely to be embryonic or hematopoietic in origin; they arise within the developing tissues, with the most common pediatric cancers being leukemia, lymphoma, soft tissue sarcomas, neuroblastoma, other peripheral nervous tumours, and cancers of the central nervous system (CNS) (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2019). Cancer classification in children uses the morphology of the cancers rather than the primary site of origin due to the diverse histological nature of the cancers and their occurrence at different sites. Overall, the cancers are more aggressive and more invasive than in adults (Canadian Cancer Society, 2017; Steliarova-Foucher et al., 2005).

In Canada, there was an average of 943 new cancer cases per year between 2009 and 2013 in children 0–14 years of age, and an estimated 1,000 children will be diagnosed with cancer in 2019 (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2017; Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2019). There are 12 main diagnostic groups and subgroups into which childhood cancers are classified according to the International Classification of Childhood Cancer, 3rd ed. (ICCC-3). Table 1 shows the main classifications and subgroups of pediatric cancers (Steliarova-Foucher et al., 2005).

The majority of cancers are caused by spontaneous mutations, with genetic factors contributing to approximately 5–10% of cancer cases (Lynch et al., 1995; Narod et al., 1991). A large proportion of germline mutations that contribute to these genetic factors have been found in a group of genes known as “cancer predisposition genes”. These genes are inherited in families and increase an individual’s susceptibility to cancer in families, leading to hereditary cancers (Lynch et al., 1995; Zhang et al., 2015).

Warthin (1913) was the first to describe hereditary cancers. In his study on a group of families with established pedigrees, he found a marked susceptibility to certain cancers in descendants of an individual with cancer. He also found that there is a higher tendency of progeny to cancer where both maternal and paternal lineages presented with cancer. This suggests an increased susceptibility to the cancers in families with cancer. Macklin's (1932) later study on tumours and inheritance also described heredity as playing a role in cancer, finding that individuals not born with a certain cancer had a predisposition to developing it. Knudson's (1971) study of retinoblastoma in later years explained the molecular biology behind cancer predisposition genes. He came up with the “two-hit” theory; it suggests that most tumor suppressor genes require mutations in both of the alleles contained in the gene, one of which may be inherited (germline mutation) to cause a phenotypic change. The germline mutation may therefore lead to carcinogenesis when the second allele is mutated via a somatic mutation. Heterozygote germline mutations that are inherited are not lethal, as the normal gene partially compensates for function. Individuals that inherit the mutation have a higher risk of developing cancer than those who do not (Knudson, 1971).

Cancer predisposition in the human genome is due to a mutation in one of three groups of genes: oncogenes, tumour suppressor genes, and DNA repair/stability genes (Strahm & Malkin, 2006; Vogelstein & Kinzler, 2004). About 8.5% of children and adolescents with cancer have been identified to harbor germline mutations in cancer predisposition genes (Zhang et al., 2015). Sixty clinically relevant genes associated with autosomal dominant cancer predisposition genes were analyzed and studied for the presence of single nucleotide variants, small insertions and deletions and germline mosaicism (Zhang et al., 2015). The *TP53, APC, BRCA2, NF1, PMS2, RB1* and *RUNX1* genes had the most variants and were deemed pathogenic or probably pathogenic (non-silent coding variants). The prevalence of germline mutations deemed as pathogenic or probably pathogenic was highest in non-central nervous system (CNS) tumours, including osteosarcomas (Zhang et al., 2015). Some of these cancer predisposition genes are also associated with syndromes known as cancer predisposition syndromes (CPSs), further increasing the susceptibility of individuals with the germline mutation to various cancers (Monsalve et al., 2011).

Cancer predisposition syndromes (CPSs) include a multitude of familial cancers in which a clear mode of inheritance can be or has been established (Monsalve et al., 2011). In children, CPSs may increase the risk of developing cancer. Li-Fraumeni Syndrome (LFS), caused by a mutation in the *TP53* gene, is the most inherited cancer predisposition syndrome. About 75% of individuals with LFS harbor a loss-of-function mutation in the *TP53* gene. LFS has been associated with many cancers, including stomach, colon, pancreas, esophagus or lung tumours as well as an increased risk of early onset melanomas. About 10% of children with osteosarcoma have germline mutations in

their *TP53* gene (Monsalve et al., 2011; Strahm & Malkin, 2006). Neurofibromatosis1 (*NF1*) affects the function of the *RAS* oncogene. The expression of protein neurofibromin produced by the *NF1* gene, which acts as a tumour suppressor, is dysregulated when the gene undergoes a mutation, disrupting the negative regulation of the *RAS* oncogene (Monsalve et al., 2011). Mutations in the *APC* gene have been associated with familial adenomatous polyposis (FAP), a CPS syndrome that increases the risk of cancer in children by almost 100% (Monsalve et al., 2011). Mutations in the *RB1* gene are associated with retinoblastomas, with a 90% chance of developing an eye tumours when a mutation in one of the *RB1* alleles is inherited (Knudson, 1971; Monsalve et al., 2011). Finally, mutations in *FANCD1*, also known as *BRCA2*, have been linked to fanconi anemia (FA), a CPS syndrome that is caused by the dysregulation of DNA repair genes. It is well-known for being associated with an increased risk of breast and ovarian cancers (Monsalve et al., 2011; Strahm & Malkin, 2006)

Spontaneous somatic mutations, which constitute the majority of cancers, are random sporadic changes that occur in an individual's genome and may lead to carcinogenesis. Downing et al.'s (2012) study identified certain somatic mutations that drive the major subtypes of pediatric cancer. Major somatic mutations identified include single nucleotide variations, insertions and deletions (indels) and structural variations of genes. These mutations were found after the whole genome sequencing of 260 pediatric tumours and matching the germlines of 15 specific tumour types, including neuroblastomas and osteosarcomas. A common type of mutation in pediatric cancers is structural alterations, which involve inter- and intra-chromosomal rearrangements (Downing et al., 2012; Downing et al., 2012). These alterations or variations include

inversions, deletions duplications and complex re-arrangements in the genome (Colnaghi et al., 2011). They are especially seen in pediatric leukemias and solid tumours. The spectrum of mutations that have been discovered in children differ from adults with similar histology, and there are variations within specific cancers dependent on child's age. Study looking at mutations in children with stage 4 NB found that the samples from the adolescents and young adults had *ATRX* mutations whereas samples taken from infants had no *ATRX* mutations (Cheung, 2012; Downing et al., 2012; Downing et al., 2012).

Although most mutations are completely random, some risk factors have been associated with certain pediatric cancers (Spector et al., 2015). These risk factors can be demographic, environmental or intrinsic in nature. Cancer incidence has been found to be highest in infants, after which it declines before increasing again in children aged 15–19. There is also a higher incidence of the majority of childhood cancers in males than females (Spector et al., 2015). This could be due to greater inactivation of some tumor suppressor genes like *RB* in males compared to females (Sun et al., 2014), and perhaps a stronger immune system in females compared to males (Chiaroni-Clarke et al., 2016; Hayter & Cook, 2012).

Radiation therapy (radiotherapy) has been confirmed to significantly increase the risk of neural tumours after exposure in childhood, and ionization radiation has been found to induce leukemia, as well as increase the risk of cancers of the stomach, bladder, kidney, and bone (Ron, 2002; Ron et al., 1988; Spector et al., 2015). A study by Cheng et al. (2014) found that coffee consumption by pregnant women may increase an unborn child's risk of acute leukemia. One explanation for this is because caffeine may act as a

topoisomerase (II) inhibitor, which is a DNA repair inhibitor or a carcinogen metabolism inhibitor (Cheng et al., 2014; Spector et al., 2015). No other known environmental factors have been found to be a precise causal factor in pediatric cancers (Ron et al., 1988; Spector et al., 2015). Unlike sporadic cancer cases, children with germline mutations have a higher risk for cancer because only one other mutation event is required. However, the frequency of such mutations remains unknown (Wang, 2016).

Although there has been an increase in the cure rate for pediatric cancers over the years, the current treatment modalities—cytotoxic chemotherapy, radiotherapy, and surgery—have some major side effects on children that negatively affect their quality of life (Downing et al., 2012). More research is thus warranted to explore the biology of the diseases in order to target the genetic changes involved in them and find novel treatment options. This thesis focuses on two pediatric cancers: neuroblastomas and osteosarcomas.

These cancers have been found to have disappointing survival rates in children. In neuroblastoma, high-risk patients have overall survival rate of less than 40% despite receiving intensive multimodal treatment (Mueller & Matthay, 2009). In osteosarcoma, children with metastasis or recurrence of the disease have a poor prognosis (Tang et al., 2008)

1.1. Non-Coding Ribonucleic Acids (ncRNAs)

The human genome comprises both protein-coding genes and RNA genes that do not code for proteins (Lander et al., 2001). Non-coding RNAs are transcripts that are not used as a template in protein synthesis (Mattick, 2001) —a large fraction of those found in humans originate as antisense transcripts from protein coding genes (Szymanski et al.,

2005; Yelin et al., 2003). Results from the Human Genome Project revealed that only 30,000–40,000 genes (~2%) coded for proteins despite the complexity of the human genome (International Human Genome Sequencing Consortium, 2004; Lander et al., 2001). NcRNAs are classified into two groups based on transcript size: small ncRNAs (less than 200 bp) and long ncRNAs (longer than 200bp) (Takahashi et al., 2014).

Some small non-coding RNAs include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), small nucleolar RNAs (snoRNAs), microRNAs (miRNAs), piwi-interacting RNAs (piRNAs), and small nuclear RNAs (snRNAs) (Costa, 2010; Lander et al., 2001). Transfer RNAs (tRNAs) translate the nucleic acid code of the mRNA into the amino acid protein sequence. Ribosomal RNAs convert mRNA into proteins via peptide bond formation catalysis and are relevant to translation machinery. snoRNAs are needed for base modification in the nucleolus and rRNA processing. snRNAs are critical parts of spliceosomes (large ribonucleic complexes that are responsible for removing introns from the pre-mRNAs found in the nucleus). MicroRNAs (miRNAs) play pertinent regulatory roles via post-translational gene silencing. Piwi-interacting RNAs (piRNAs) are associated with Piwi proteins and are implicated in gametogenesis (Costa, 2010; Lander et al., 2001).

Previously regarded as “junk DNA”, non-coding RNAs now play crucial and pertinent roles in regulating mechanisms that underlie development and differentiation in humans by controlling protein expression (Szymanski et al., 2005). Non-coding RNAs have been found to play various transcriptional and post-transcriptional regulatory roles via epigenetic changes such as chromatin modifications (Costa, 2010; Taft et al., 2010); they have been implicated in diseases such as cancer.

1.1.1. Long Non-Coding RNAs

Long non-coding ribonucleic acids (lncRNAs) are non-coding RNAs that have more than 200 base pairs (bp) that do not code for proteins (Brosnan & Voinnet, 2009; Kapranov et al., 2007); they have been found to influence gene expression regulation (Brosnan & Voinnet, 2009; Rinn & Chang, 2012). They are mostly transcribed by RNA polymerase II, with a few transcribed by RNA polymerase III—many do not undergo processing, like in the case of messenger RNA (mRNA) (Brosnan & Voinnet, 2009). LncRNAs have been found to be involved in a number of regulatory processes (Costa, 2010), show tissue-specific expression, have high expression variability (Derrien et al., 2012; Takahashi et al., 2014), and are retained in various subcellular compartments after transcription. They are generally more enriched in the nucleus, unlike mRNAs, which are exported into the cytoplasm. This suggests that the lncRNAs could have a potential function where they are localized (Derrien et al., 2012; Saxena & Carninci, 2011).

LncRNAs have been classified into many categories, some of which have been found to overlap. They may be broadly classified into five categories based on origin of transcription in relation to protein coding genes: (1) sense, (2) antisense, (3) bidirectional, (4) intronic, and (5) intergenic (illustrated in Figure 1a, Takahashi et al., 2014). Sense lncRNAs are transcribed when they overlap one or more exons of another transcript on the same strand, antisense lncRNAs overlap on the exons of the opposite strand, bidirectional lncRNAs are expressed when the lncRNA is transcribed in close proximity to a neighbouring coding transcript on the opposite strand, intronic lncRNAs are derived from within an intron, and intergenic lncRNAs are transcribed within the genomic interval between two genes (Ponting et al., 2009; Takahashi et al., 2014).

LncRNAs are also categorized based on transcription length as well as their function or association. This includes association with annotated protein-coding genes, with other DNA elements of known function, with repeats, with biochemical pathway or stability, and with subcellular structures and localisation. LncRNAs associated with annotated protein-coding genes are transcribed from overlapping non-coding and coding genes. Other categories include; sequence or structure conservation, expression in different biological states, protein coding resemblance among a few others. LncRNAs with a spliced structure, a polyA tail and a conserved sequence are said to have protein coding or mRNA resemblance (St. Laurent et al., 2015).

LncRNAs have again been broadly classified based on their mechanisms of action; epigenetic regulation, transcriptional regulation, post-transcriptional regulation, and non-regulatory roles (Signal et al., 2016). Epigenetic regulation functional roles include maintaining chromatin structure by recruiting chromatin modifying complexes to DNA and targeting chromatin regulators, positive and negative chromatin modification, gene silencing and chromatin remodeling. Transcriptional regulation by lncRNAs involves altering or affecting transcription patterns by targeting proteins to specific genomic loci, transcriptional interference, and acting as a co-factor for transcriptional factors. Post-transcriptional regulation consists of RNA processing, serving as precursors for functional small RNAs, miRNA binding and miRNA sponging, RNA splicing, RNA localisation, and influencing the stability of other RNAs and proteins. LncRNAs sponge miRNA by binding to them via base-pairing with miRNA-recognition elements, thus sequestering them and reducing their effect on target mRNAs. Non-regulatory functions constitute lncRNAs serving as a protein scaffold, acting as a mimic or decoy and binding

to proteins to influence their localisation (Brosnan & Voinnet, 2009; Derrien et al., 2012; Paci et al., 2014; Rinn & Chang, 2012; Signal et al., 2016; Wilusz et al., 2009). Figure 1b provides a summary of some of the functions of lncRNAs in a cell (Wilusz et al., 2009).

Here, some functions of lncRNAs are briefly discussed, and the mechanisms involved are illustrated. LncRNAs *XIST, AIR,* and *HOTAIR* exhibit lncRNAs' influence in transcriptional gene regulation. *XIST* is a 17kb lncRNA found in mammals that is expressed from the nucleus and is an element in the '"X-inactivation centre". LncRNA *XIST* has been extensively studied and found to be involved in gene silencing by physically coating the future inactive X chromosome, recruiting silencing factors, and condensing the X-chromatin (Erwin & Lee, 2008). *XIST* also binds to the polycomb repressor complex 2 (PRC2) that is involved in histone H3K27 trimethylation, subsequently leading to transcriptionally inactivating the X chromosome (Erwin & Lee, 2008). Another lncRNA, *AIR,* is also involved in silencing genes by accumulating at the promoter gene to coat it, and recruiting a histone methyltransferase G9a, which leads to H3K9 methylation (Brosnan & Voinnet, 2009; Nagano et al., 2008). LncRNAs like *HOTAIR* are also involved in silencing genes. *HOTAIR* recruits the PRC2 complex to genes, forming inactive domains (Brosnan & Voinnet, 2009).

Recent discoveries implicate lncRNAs in many cellular processes, including cell cycle regulation, embryonic stem cell pluripotency, and functioning as either oncogenes or tumour suppressors in some cancers (Rinn & Chang, 2012). *LincRNA-21* belongs to a group of lncRNAs that are strongly associated with P53, a tumour suppressor that plays a relevant role in regulating cell cycles. P53 triggers apoptosis or initiates cell cycle arrest

when DNA damage is detected in cells. *LincRNA-21* mediates transcriptional regression by acting as a downstream target repressor when p53 is activated (Huarte et al., 2010).

1.1.2. Long Non-Coding RNAs in Cancer

Many lncRNAs have been implicated in various diseases such as Alzheimer's disease, schizophrenia, hypertension, and cancer (Harries, 2012), and more research is being done into the roles that these lncRNAs as well as the effects novel lncRNAs could play in cancer tumourigenesis. The dysregulation of lncRNAs as well as changes in their interaction with the other biological molecules they target affect gene expression and the regulation of varying pathways, contributing to cancer.

LncRNAs are associated with cancers involving the following systems: digestive, respiratory, reproductive, urinary, skeletal, and CNS (Zhang et al., 2016). These lncRNAs have been implicated in different cancers. Further, some cancers have more than one lncRNA that contributes to their tumourigenesis or affects their progression.

The lncRNA *HOTAIR* has been implicated in the progression of human cervical cancer by upregulating the vascular epithelial growth factor (VEGF), matrix metallopeptidase 9 (MMP-9), and epithelial-to-mesenchymal transition (EMT)-related genes that contribute to tumourigenesis by promoting migration and invasion (Kim et al., 2015). *HOTAIR* overexpression has also been found in primary breast cancer tumours, but it is most frequently found in metastatic tumours (Gupta et al., 2010). *HOTAIR* has also been implicated in other cancers, with studies reporting that its overexpression promotes the genomic relocation of polycomb repressor complex 2 (PRC2) and H3 lysine 27 trimethylation, leading to metastatic progression (Gupta et al., 2010).

ANRIL (an antisense non-coding RNA in the INK4 locus) is overexpressed in serious ovarian cancer (SOC) tissues; elevated levels of *ANRIL* have been found to be associated with lymph node metastasis and poor prognosis in SOC (Qiu et al., 2015). The increased expression of lncRNA *SRA* (steroid receptor activator) is strongly associated with breast cancer; further, polymorphisms in lncRNA *SRA* have been suggested to affect the estrogen receptors in breast development (Yan et al., 2016). Estrogen-stimulated cell growth in breast cancer cells has also been found to be affected by lncRNA *H19*, with increased expression of *H19* promoting growth in MCF-7 breast cancer cell lines (Sun et al., 2015).

Various studies have found lncRNA *MALAT1* (metastasis-associated lung adenocarcinoma transcript 1) expression to be elevated in lung cancers (Feng et al., 2019; Gutschner et al., 2013; Li et al., 2018); it is associated with disease progression by regulating metastatic gene expression (Gutschner et al., 2013), modulating miR-124 expression (Li et al., 2018), and gefitinib resistance via sponging miR-200a (Feng et al., 2019). *MALAT1* expression is also increased in gall bladder cancers, playing an oncogenic role that may promote the cancer by sponging miR-206, which plays a tumour-suppressive role (Wang et al., 2016).

H19 is also upregulated in bladder cancers, and may promote metastasis via interaction with EZH2 (an enhancer of zeste homolog 2) and inhibiting E-cadherin (a key molecule in cell-to-cell adhesion) (Luo et al., 2013). *XIST* lncRNA has been found to be involved in bladder cancers, with upregulated expression found in bladder cancer tissues along with increased levels of androgen receptors. Increased levels of *XIST* and androgen receptors positively correlated with a poorer tumour, node, metastasis (TNM) stage of

bladder cancer. Further, the knockdown of *XIST* resulted in reduced cell proliferation, migration and invasion (Xiong et al., 2017). *XIST* lncRNA also progresses glioblastoma (a common and aggressive brain tumour), and it has been found to be upregulated in glioblastoma stem cells and tissues (Yao et al., 2015).

The overexpression of lncRNA *MEG3* (maternally expressed gene 3) reduces cell proliferation in hepatocellular carcinoma (HCC) cell lines, and it has been found to be downregulated in HCC cells and tissues (He et al., 2017). He et al. (2017) suggested that *MEG3* sponges miR-664 when overexpressed, which relieves the transcription and translation of the ADH4 gene that inhibits the proliferation of HCC cells, acting as a tumour suppressor in HCC.

LncRNA *CCAT1* (colon cancer-associated transcript 1) expression is increased in acute myeloid leukemia (AML) (M4 and M5 subtypes) as well as in colon cancers (Alaiyan et al., 2013; Chen et al., 2016). In AML, the increased expression of *CCAT1* promotes cell proliferation and inhibits the differentiation of the myeloid cells by binding miR-155—which functions as a tumour suppressor and is down-regulated in AML—and controls c-Myc expression (Chen et al., 2016). In colon cancers, the overexpression of *CCAT1* promotes cell proliferation and the invasion of colon cancer cells; it shows a significant association with stage, lymph node metastasis, and survival time in patient tissues. c-Myc has also been found to promote *CCAT1* transcription and expression in colon cancers (He et al., 2014). *CCAT1* has also been reported to be upregulated in gastric carcinomas, with its enhanced expression promoting the proliferation and migration of cancer cells. *CCAT1* overexpression also correlates with the growth of

primary tumours, lymph node metastasis, and metastatic disease, with c-Myc influencing its expression in gastric carcinoma (Yang et al., 2013).

Many other lncRNAs have been identified and continue to be identified in cancers, outlining possible associations and contributions either in oncogenic or tumour-suppressive capacities. This scope is too broad for this thesis to cover in its entirety.

1.2. Epigenetics

The chromatin structure within a cell influences which genes are activated or silenced—it therefore affects gene expression within the cell. Waddington (1942, 2011) first described epigenetics as an interaction between genes and their products leading to the expression of a particular phenotype. Further studies into epigenetics report that these changes are heritable and affect gene expression without changing the DNA sequence in the genome (Bird, 2002; Jones & Martienssen, 2005; Sharma et al., 2010). These changes are believed to occur during differentiation in the cells, after which the changes are maintained—allowing the cells to contain the same information following multiple replications. These epigenetic changes that occur within the human genome led to the coining of the term "epigenome" — a complement to the modification of the human genome that allows for cellular diversity by regulating cellular machinery's access to genetic information (Sharma et al., 2010).

Epigenetic modifications are changes within a cell that lead to a change in the chromatin structure and mediate the heritability of gene expression. These modifications include the methylation of cytosine bases in DNA (DNA methylation), post-translational modifications of histone proteins (covalent histone modifications), changes in nucleosome positioning along DNA and histone variants (non-covalent mechanisms), and non-coding RNA interactions. (Sharma et al., 2010). These modifications help to regulate cell expression by altering chromatin structure and therefore regulating the compactness of the DNA and the cellular machinery's access to the DNA.

DNA methylation covalently modifies the cytosine residues of DNA found in CpG dinucleotides by transferring methyl groups from the S-adenosyl methionine (SAM) to the cytosine bases (Sharma et al., 2010; Wongtrakoongate, 2015). DNA methylation occurs mostly on CpG islands—small stretches of concentrated CpG repetitive sequences that occupy about 60% of the human genome. The majority of these CpG islands are unmethylated during cell development and in undifferentiated tissues. However, some CpG island regions found in promoters become methylated, leading to long-term transcriptional silencing and therefore affecting the expression of the genes that would be found in the region. DNA methylation works by preventing or promoting the regulatory proteins' recruitment to DNA, silencing the genes in either the coding or non-coding regions of the genome (Sharma et al., 2010).

DNA methyltransferases (DNMTs) are involved in DNA methylation and either generate methylation patterns and heritability or maintain the methylation patterns over multiple cell replications. Although five DNMT family proteins have been identified (DNMT1, DNMT2, DNMT3A, DNMT3B, DNMT3L), only three are functional DNA methyltransferases (DNMT1, DNMT3A and DNMT3B). DNMT2 is an RNA methyltransferase, and DNMT3L does not have a catalase domain for methyltransferase (Wongtrakoongate, 2015). DNMT1 is a methyltransferase with a preference for already methylated and hemimethylated DNA that is active during cell replication. Finally, DNMT 3A and 3B are *de novo* methyltransferases that are involved in generating methylation patterns and maintaining heritability. These DNMTs tend to act independently of replication and have no preference for either unmethylated or

hemimethylated DNA. All three DNMTS can maintain DNA methylation patterns from the DNA template to the daughter strand (Sharma et al., 2010; Wongtrakoongate, 2015).

Histone modifications alter histone proteins, which affect chromatin state. DNA is organized in the cell as chromatin; the nucleosome is a fundamental unit chromatin that consists of an octamer of four core histones (H3, H4, H2A, and H2B) with ~147bp DNA wrapped around it (Kouzarides, 2007; Sharma et al., 2010). These histone proteins have two domains: N-terminal and C-terminal domains, and the N-terminal domain can undergo post-translational covalent modifications (Sharma et al., 2010). The N-terminal tail of histone proteins is unstructured; it possesses a large number and type of modified residues and undergoes modifications that include methylation, acetylation, sumoylation, ubiquitylation, and phosphorylation (Kouzarides, 2007; Sharma et al., 2010). These modifications help to regulate key cellular processes like transcription, replication, and repair (Sharma et al., 2010).

Histone modifications influence gene expression by altering higher-order chromatin structure, therefore affecting accessibility to the genes and leading to genes' activation or repression (Kouzarides, 2007; Sharma et al., 2010). Modifications act either by unraveling chromatin or by recruiting non-histone proteins, or both. Where non-histone proteins are recruited, they either bind to or are occluded from chromatin. These non-histone proteins may also carry enzymes whose activities could further modify the chromatin structure (Kouzarides, 2007). The enzymes involved in modifying histones include histone acetyltransferases (HATs) and histone deacetylases (HDACs), which add and remove acetyl groups, respectively; histone methylases (HMTs) and histone demethylases (HDMs), which add and remove methyl groups, respectively; and

phosphorylases and ubiquitilases, which add phosphoryl groups and ubiquitin, respectively, to the histones (Kouzarides, 2007; Sharma et al., 2010).

Non-covalent mechanisms including nucleosome positioning and histone variants in the genome may also alter chromatin structure and therefore influence gene activity and regulation. The positioning of nucleosomes alters the transcription factors' access to regulatory DNA sequences. Some regions are present at the ends of the 5' and 3' genes and are thought to be nucleosome-free regions (NFRs). Sharma et al. (2010) suggest that these areas are the sites of transcription machinery assembly and disassembly. NFRs present at promoter regions have been found to correlate with rapid gene activation, nucleosomes present in NFRs have been associated with gene repression, and the loss of a nucleosome directly upstream of a transcription start site has been found to correlate with gene activation (Sharma et al., 2010). NFR modulation occurs in an adenosine triphosphate (ATP)-dependent manner and includes chromatin-remodeling complexes. This disrupts the DNA–histone interactions and grants access to the DNA regulatory sites by facilitating the sliding and ejection of the nucleosomes (Lund, 2004; Sharma et al., 2010).

Histone variants include specialized histone variants like H3.3 and H2A.z replacing canonical histone proteins. Histone variants may also incorporate into the nucleosome, influencing the nucleosome's capacity—this in turn affects gene activity. Unlike the major histone types, the synthesis and incorporation of which are restricted to the S phase in the DNA replication cycle, histone variants are involved throughout the cell cycle in a dynamic manner, further influencing the regulation of gene expression (Lund, 2004; Sharma et al., 2010). H2A.X (a H2A variant) has been found to mark

double strand breaks in DNA and is required for the repair of the breaks. Cells that are deficient in this histone variant have an increased correlation with higher frequencies of genomic instability, and the impaired formation of certain DNA repair proteins (e.g., BRCA) (Lund, 2004). H2A.Z in mammals is needed for chromosome segregation, and its incorporation may also contribute to gene activation by protecting against DNA methylation (Sharma et al., 2010). Histone variants may also undergo other histone modifications, including acetylation and ubiquitylation, which could further influence the regulation of gene activity (Sharma et al., 2010).

Early ncRNAs that were found to play a role in epigenetics were microRNAs (miRNAs). These ncRNAs have been found to post-transcriptionally regulate gene expression by silencing target genes. The miRNA post-transcriptional mechanism that silences target genes conducts sequence-specific pairings with messenger RNAs (mRNAs) (which code for protein), presents within the RNA-induced silencing complex (RISC), and inhibits the translation or degradation of the mRNAs (Lund, 2004; Sharma et al., 2010). MiRNAs help to regulate a number of cell processes, including cell proliferation, apoptosis, and differentiation—further, their expression is tissue specific. Although miRNAs are involved in epigenetic regulation, they may themselves also be epigenetically regulated, as the cell transcribes them. Some miRNAs may also target the enzymes that are responsible for DNA methylation and histone modification (Sharma et al., 2010).

Downing et al.'s (2012) study mentions the importance of integrating the influence of epigenetics in pediatric cancers. When they analyzed integrated genome—level data with epigenetic expression and RNA expression data, they found that

retinoblastomas, which had few mutations across the genome, had an aberrant SYK (a cytoplasmic tyrosine kinase) expression. Subsequent tests on SYK inhibition revealed apoptosis of the retinoblastoma tumour cells, revealing the benefit of combining transcriptome sequencing with whole-genome sequencing and hence the pertinent role that epigenetics could play in cancer expression.

1.2.1. Long non-coding RNAs in epigenetics

LncRNAs have been involved in epigenetically silencing several genes through lncRNA-chromatin remodeling complexes and other epigenetic mechanisms. LncRNA involvement was suggested when the enzymatic members of chromatin-remodeling complexes were found to have RNA binding domains and not DNA binding domains (Saxena & Carninci, 2011; Y. Sun & Zhang, 2005). However, it is still unclear whether the domains present on the members of the chromatin remodeling complexes bind with specific ncRNAs or general ncRNAs if they conform to a certain secondary structure (Saxena & Carninci, 2011).

LncRNA *KCNQ1OT1* has been found to interact with H3K9 and H3K27 histone methyltransferases as well as chromatin, resulting in the regulation of genes found in the kncq1 domain (including bidirectional silencing) (Pandey et al., 2008; Saxena & Carninci, 2011). LncRNA *H19*, transcribed from the H19-imprinted gene has been found to be involved in the regulation of genes in the imprinted gene network (IGN). *H19* lncRNA has been found to epigenetically interact with the target genes in the network by binding with the methyl-CpG-binding domain protein 1 (MBD1). MBD1—a relevant protein in the maintenance of H3K9me3 (trimethylation) by recruiting lysine methyl

transferases (KMTs)—and the resulting H19-MBD1 complex that is formed represses the expression of the target genes by inducing histone tail modifications in their differentially methylated regions (DMRs) (Monnier et al., 2013). *HOTAIR* has been found to physically interact with polycomb repressor complex 2 (PRC2), recruiting the complex to silence specific genes, including genes at the *HOXD* locus via methylating specific histone tails (Rinn et al., 2007; Saxena & Carninci, 2011). *HOTAIR* has also been found to interact with REST/CoREST complexes, which are involved in the demethylation of specific histones and are therefore capable of binding and linking both methylase and demethylase by acting as a modular scaffold (Tsai et al., 2010). *XIST* lncRNA has been found to be involved in X-chromosome inactivation by coating the chromatin at the inactive X chromosome, silencing its expression by restricting access to transcriptional mechanisms (Erwin & Lee, 2008; Saxena & Carninci, 2011). *XIST* also directly binds PRC2 complex, affecting the methylation of the histone tails and affecting the gene expression. X-chromosome inactivation is a well-studied developmental regulatory mechanism affecting the expression of the genes on the entire X-chromosome. It reaches dosage equivalence between males and females by randomly inactivating an X-chromosome in females (Brown et al., 1991; Lee, 2012).

Several lncRNAs have been found to be involved in epigenetic mechanisms; this results in differential regulation of the cell either by silencing or activating target genes. Some target genes may fail to be silenced in the absence of the lncRNA involved, suggesting the pertinent role these ncRNAs play in regulating cell processes.

OBJECTIVE

This thesis focuses on the influence of long non-coding RNAs in neuroblastoma and osteosarcoma. The aim is to identify differential expression of lncRNAs in the pediatric cancers which can serve as potential diagnostic and therapeutic markers.

HYPOTHESIS

I hypothesize that long non-coding RNAs are differentially expressed in neuroblastoma and/or osteosarcoma tissues and cell lines and that these long non-coding RNAs may play regulatory roles and have potential functions in the progression of these cancers.

CHAPTER 2: THE ROLE OF LONG NON-CODING RNAS IN NEUROBLASTOMA

2.0. Neuroblastoma

Neuroblastoma (NB) is identified as one of the most common malignant solid tumours in children; it can occur anywhere along the sympathetic nervous system and arises when neural crest cells are improperly differentiated. It is a heterogeneous tumour with the tendency to spontaneously mature and regress (D'Angio et al., 1971; Gestblom et al., 1997; Kamijo & Nakagawara, 2012; Pandey et al., 2014) or to show an aggressive therapy-resistant phenotype, sometimes in association with the mutations involved (Maris et al., 2007; Monclair et al., 2009). It is commonly diagnosed within the first year of a child's life with the majority arising in the adrenal medulla (Kamijo & Nakagawara, 2012; Pandey et al., 2014).

The International Neuroblastoma Risk Group Staging System (INRGSS) classifies the disease into stages prior to surgical resection or any treatment, which differs from the previous staging system, the International Neuroblastoma Staging System (INSS) (Monclair et al., 2009). The INRGSS classification is based on tumour imaging, the understanding being that diagnostic images are more likely to be reproducible than operative findings, and it is dependent on clinical criteria, such as the age of the patient, the presence or absence of specific image-defined risk factors (IDRFs), and metastasis (Monclair et al., 2009).

The INRGSS has defined four stages in classifying NB at the time of diagnosis: L1, L2, M, and MS, with the descriptions summarized in Table 2 (Monclair et al., 2009). The classifications follow the previous INSS classifications, where the disease was categorized into stages 1, 2A, 2B, 3, 4, and 4S. INSS stage 4 is identical to INRGSS stage M, and INSS stage 4S is highly similar to INRGSS stage MS (Monclair et al., 2009).

Changes to the classification helped address difficult situations in which the tumour seemed to have a different classification based on surgical excision, the localized tumour was anticipated to regress, or the assessment of lymph node involvement differed based on the surgeon involved (Monclair et al., 2009). The INSS system is still however used in NB classification. A brief description of the INRGSS and INSS stages is provided in Table 2a and 2b, respectively (adapted from Monclair et al., 2009; Mueller & Matthay, 2009).

Mutations that lead to NB may be familial or spontaneous. Familial or hereditary NB encompasses a small percentage of NB cases, with less than 5% being through autosomal dominant inheritance (Maris et al., 1997). Spontaneous mutations are the cause of most NB cases, and the alterations in the individual's genome may include non-random chromosomal alterations such as *MYCN* proto-oncogene (*MYCN*) amplification; mutations in the anaplastic lymphoma kinase (*ALK*) gene; a gain on chromosome 17q, loss of heterozygosity (LOH) on chromosome 1p; and the allelic loss of chromosome 11q (Attiyeh et al., 2005; Kamijo & Nakagawara, 2012; Pandey et al., 2014). More than one of the mutations may be found in primary NBs.

MYCN is frequently amplified in NB and occurs in about 20% of primary NB cases (Suenaga et al., 2014). Amplified *MYCN* in tumours shows a strong correlation with an advanced disease stage or more aggressive tumours and poor prognosis (Maris et al., 2007; Weiss, 1997). *MYCN* amplification between 50- and 100-fold is usually seen in tumours; however, different amplified *MYCN* values can be found in NBs—between 5-fold and more than 500-fold (Kamijo & Nakagawara, 2012).

MYCN is part of the *MYC* family of proto-oncogenes that generally play a positive role in the cell cycle and proliferation in the presence of growth factors, and they act as transcription factors to regulate gene expression (Galderisi et al., 1999). The N-Myc expression is most often limited to brain cells, kidney cells, and lymphocytes in the early stages of differentiation, as well as in neuronal transformation (Galderisi et al., 1999; Weiss, 1997). *MYCN* inhibition was found to induce differentiation and decrease the proliferation of most NB cell lines (Galderisi et al., 1999).

Mutations in the *ALK* gene have been found to be a frequent source of gene alterations in NB cases, occurring in approximately 11% of advanced NB cases. The *ALK* gene is a repeated target of copy number gain and gene amplification in primary NB samples (Chen et al., 2008; George et al., 2008). Mutated *ALK* has been found to show an increase in kinase activity, allowing for the proliferation of the disease, and an interference of mutated *ALK* in NB has been found to suppress the disease growth (Chen et al., 2008). *ALK* mutations correlate highly with *MYCN* amplification in neuroblastoma. However, unlike *MYCN* amplification, *ALK* mutations are unlikely to have a profound effect on the prognosis and survival of the disease (Chen et al., 2008).

ALK mutations have also been found to be a familial NB, as well as somatic gene. The mutation occurs in about 1 to 2% of familial cases, and even within families, there is much heterogeneity in the disease (Maris et al., 2007; Mossé et al., 2008). NB is often diagnosed earlier in the case of familial inheritance (Maris et al., 2007; Mossé et al., 2008), and it sometimes presents with other primary tumours and other disorders related to the abnormal development of neural crest-derived tissues, such as Hirschsprung's disease (Mossé et al., 2008).

Chromosome loss (or allelic loss) at 1p is one of the marked regions of LOH in NB, and it has been found to occur frequently in NB cases (Attiyeh et al., 2005; Mueller & Matthay, 2009; White et al., 1995) . It occurs in about 36% of all primary tumours and has also been found to be associated with *MYCN* amplification, and advanced disease stage (Maris et al., 2007; Mueller & Matthay, 2009). Allelic loss at the 1p region has been found to also predict an increased risk of disease relapse in localized tumours (Maris et al., 2007). Deletions in this region have been mapped to large areas, as well as specific locations within the genome encompassing 1p36 (1p36.1–1p36.3) (White et al., 1995).

Chromosome loss of 11q is another marked region of LOH in NB that occurs in more than 30% of primary tumours. The shortest region of overlap has been mapped to 11q23. Unlike the deletions in 1p, LOH in 11q is rarely associated with *MYCN* amplification (Maris et al., 2007; Mueller & Matthay, 2009). A putative NB suppressor gene has been suggested to map between bands 11q14–23 (Maris et al., 2001).

Allelic loss of 3p has also been found to occur in NB cases, with a 13% to 25% occurrence, and it is frequently associated with LOH in 11q (Mueller & Matthay, 2009;

Spitz, Hero, Ernestus, & Berthold, 2003). Allelic losses of 11q and 3p have been found more often in older children, suggesting that one of the aberrations may occur at a later event (Mueller & Matthay, 2009; Spitz et al., 2003)

Gain at 17q is a frequent chromosomal change that occurs in 70–80% of primary NBs (Janoueix-Lerosey et al., 2000). Chromosome gain at 17q has been found to correlate with a more aggressive phenotype of the disease and poor survival (Janoueix-Lerosey et al., 2000; Maris et al., 2007). The chromosome 17q arm gains from unbalanced translocations between chromosome 17 and most often chromosome 1, but it may also gain from a translocation with other chromosomes. The gain region is between 17q22–23 and 17qter (Janoueix-Lerosey et al., 2000). Chromosome 17q gain has been found to be closely associated with deletion in 1p36 and *MYCN* amplification (Janoueix-Lerosey et al., 2000). Current treatment modalities for NB include surgery, chemotherapy, radiotherapy, biotherapy, and observation in some selected cases (Maris et al., 2007).

2.1. Long Non-coding RNAs in Neuroblastoma

Different long non-coding RNAs have been reported to be involved in NB, with some influencing the disease based on its genetic abnormality. A study by Prajapati et al. (2019) identified three lncRNAs that are significantly altered using an available RNAseq dataset on the disease. The lncRNAs *CASC15*, *PPP1R26-AS1*, and *USP3-AS1* have been identified to be significantly altered. Meanwhile, the lncRNA *CASC15* (cancer susceptibility candidate 15) found in the genomic loci region 6p22.3 that is usually altered in NB was found to be significantly upregulated, along with *PPP1R26-AS1*, while

USP3-AS1 was significantly downregulated. The lncRNAs *PPP1R26-AS1* and *USP3-AS1* were associated with overall survival probability within the samples (Prajapati et al., 2019).

Other lncRNAs have been found to play oncogenic roles in NB, affecting cell fate and influencing cell proliferation, apoptosis, migration and invasion. LncRNA *MIAT* downregulation in NB cell lines is associated with reducing long-term survival of NB cells and may decrease migration of the disease (Bountali et al., 2019). Small nucleolar RNA host gene 16 (*SNHG16*) is aberrantly expressed in NB with expression levels correlating to the INSS staging system, a high expression associated with cell proliferation, migration invasion and worse overall survival (Wen et al., 2020). Increased *MYCNOS-01* expression contributes to NB cell growth in MYCN-amplified cells (O'Brien et al., 2018). *LNCUSMYCN* is highly expressed in MYCN-amplified cell lines and tissues, and it upregulates the N-Myc expression through post-transcriptional mechanisms. A high expression of *LNCUSMYCN* has been reported to induce cell proliferation in NB, and the lncRNA upregulates the N-Myc expression by binding NonO protein (Liu et al., 2014a, 2014b).

XIST`s involvement in NB has also been recorded. The *XIST* expression has been found to be higher in NB cell lines and is reported to modulate the H3 histone methylation of DKK1 (Dickkopf wnt signaling pathway inhibitor 1) via enhancer of zeste homolog 2 (EZH2) to influence the cell growth of NB (Zhang et al., 2019). *MALAT1* has been found to show higher expression levels in NB and suggested to mediate the Axl expression that promotes invasion and migration. Axl is a part of the Tyro-3-Axl-Mer

(TAM) family of receptor tyrosine kinases (RTKs), and its inhibition is reported to lead to cell death and enhanced chemosensitivity in NB (Bi et al., 2017).

Neuronal differentiation lincRNA hosting miR-125 (*LINC-NED125*) has been found to play a tumour-suppressive role by increasing during neuronal differentiation with overexpression in cells and decreasing cell proliferation and arresting cell cycle progression. The *LINC-NED125* gene is suggested to host miR-125, which promotes neuronal differentiation and is down-regulated in NB (Bevilacqua et al., 2015). NB-associated transcript 1 (*NBAT1*) has been found to be expressed at low levels in NB tissues and cell lines, and the overexpression of the lncRNA induces neuronal differentiation. A high expression of *NBAT1* is associated with good prognosis in NB cases, and it has been reported to interact with *EZH2*. The *NBAT1*- EZH2 interaction suppresses target genes implicated in cell proliferation and cell migration via chromatin-mediated changes, therefore implicating *NBAT1* in a tumour suppressor role (Pandey et al., 2014). More lncRNAs continue to be discovered that contribute to tumourigenesis in NB in hopes to serve as novel markers for prognosis and therapeutic targets.

2.2. Materials and Methods

2.2.1. Cell Lines and Cell Culture

The NB cell lines IMR5, SK-N-AS, SK-N-BE(2), and SK-N-MC; normal cell line WI38; and the glioblastoma cell lines A172, MO59J, MO59K were cultured by supplying appropriate media supplemented with 10% fetal bovine serum (FBS) and maintained in a humidified incubator at 37°C with 5% carbon dioxide (CO_2). A172 and SK-N-AS were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Multicell, Wisent Inc.), SK-N-MC was cultured in Eagle's Minimal Essential Medium (EMEM) (Multicell, Wisent Inc.), SK-N-BE2, M059J and M059K were cultured in Dulbecco's modified Eagle's medium and Ham's F-12 nutrient mixture (DME/F12) (Multicell, Wisent Inc.), and IMR5 was cultured in Roswell Park Memorial Institute 1640 medium (RPMI-1640) (Multicell, Wisent Inc.).

2.2.2. RNA Extraction and Sequencing

Total RNA from NB tumour samples and normal tissue were stored at -80°C until RNA isolation (Table 3 lists all samples used in the study). A section was cut from each of the samples provided to be used in the study. Total RNA was extracted using the Direct-zol™ RNA Miniprep Plus Kit (Zymo) adhering to the manufacturer's protocol. The RNA quality and quantity were determined using a NanoDrop 2000c spectrophotometer (v. 1.5). The RNA was sequenced using Illumina® NextSeq 500/550 kits following the manufacturer's instructions. Tissue samples were provided as a part of POETIC Genomics Consortium collaboration. The study was approved by the U of L Human Subject Committee, protocol number 2016-064.

2.2.3. In-Depth Transcriptome Profiling and Bioinformatics

Sequencing data were analyzed by bioinformatics to determine the differential expression of long non-coding RNAs between normal and tumour samples to identify significant differences in expression. The data were compared with a range of human RNA libraries to find significant changes in lncRNA levels using the CASAVA software (v.1.8).

To confirm the lncRNAs of interest did not randomly associate with other genes, the long non-coding RNAs identified were subjected to correlation analysis with protein coding genes in NB tissue samples using Pearson's correlation method, and adjustments for multiple comparisons were done using the Benjamini–Hochberg (BH) method to determine significant interactions. Significant genes were determined using adjusted p-value (padj) < 0.05.

Gene ontology (GO) enrichment and the KEGG pathway analysis were used to determine the enrichment of the biological themes of protein coding genes significantly correlated with the lncRNAs using geneSCF v.1.1-p2. Genes with an absolute Pearson's correlation coefficient >0.75 and a padj <0.05 were considered significant. Enrichments were calculated using Fisher's exact test, after which there were multiple comparison adjustments made using several alternative methods.

Prediction of the subcellular localisation of lncRNAs of interest was done using the lncATLAS database (Mas-Ponte et al., 2017). RIblast software was used to predict RNA-RNA interactions between the lncRNAs of interest and target genes. The top 25

target genes were compared to genes differentially expressed between the NB tumour and comparative normal samples.

2.2.4. Protein Extraction and Quantification

Cells were harvested by dislodging using trypsin/EDTA and rinsing with cold 1× phosphate buffer sulfate (PBS). The mixture was centrifuged at 2,000 rpm for five min. The supernatant was discarded, and the pellets were solubilized in laboratory-prepared 1% sodium dodecyl sulfate (SDS) lysis buffer (BioUltraPure, Bioshop) by sonication using a Braunsonic model 1510 sonicator (B. Braun Germany) operating at 80% sonication intensity. Lysates were centrifuged at 15,000rpm for 10 min and the supernatant was decanted for use. Protein concentrations were determined using the Bradford protein assay with bovine serum albumin as the standard protein using the Nanodrop 2000c spectrophotometer (v. 1.5).

2.2.5. SDS-PAGE and Western Blotting Analysis

Total proteins (50–100 μg) were separated by electrophoresis on 8-10% sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE). The protein was then electro-transferred onto activated polyvinyl difluoride (PVDF) membranes (Amersham Hybond P 0.45, GE Healthcare). The membranes were incubated for one hour in a blocking solution (5% dry skimmed milk in PBS, 0.5% Tween 20) at room temperature, and incubated with specific primary antibodies at 4°C overnight. The primary antibodies anti-Gsk3α/β monoclonal antibody (1:200, Santa Cruz Biotechnology- sc-7291), anti-EZH2 polyclonal antibody (1:1000, Cell Signaling Technologies, 4905), anti-ENX-1 (D-8)

monoclonal antibody (1:200, Santa Cruz Biotechnology, sc-137255) and anti-Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) monoclonal antibody (1:1000, Santa Cruz biotechnology, sc-47724) were used. Blots of the primary antibodies tested were developed with peroxidase-labelled secondary antibodies specific to the primaries. The membranes were washed (5 times of 5 min of washing with PBS-Tween) before and after adding secondary antibodies. The membranes were run in duplicates.

2.2.6. Immunofluorescence and Chemiluminescence Detection

The protein bands of interest were detected using an enhanced chemiluminescence (ECL) system by incubating for 5 min in ECL detection reagents (GE Healthcare, Amersham Biosciences) and visualized using the FluorChem HD2 ALC detection system (software v.3.2.2.0805, Cell Biosciences).

2.2.7. Quantitative Real-Time PCR

The expressions of the long non-coding RNAs *LINC00261* and *LINC01133* in the NB cell lines were determined using SsoFast EvaGreen Supermix (Bio-Rad, Foster City, CA, USA) and the Bio-Rad CFX96^{TM} Real-Time System (Bio-Rad, Foster City, CA, USA). The total reaction volume was 20 µL and contained 10 µL Ssofast Evagreen Supermix, one µL cDNA template, five µM of each primer (forward and reverse), and nuclease-free water adjusted to the total volume. Primer sequences used are listed in Table 4. Conditions for amplification were 95 °C for 30 sec, followed by 49 cycles of 95 °C for 5 sec and 60 °C for 10 sec. The melting curves were obtained by slow heating (0.5°C/s) at temperatures in the range of 65 to 95 °C. All samples were run in triplicate,

and GAPDH was used as an internal control. Both the *LINC00261* and 01133 levels were standardized to the internal control using the ΔCT (relative cycle), and fold changes were calculated using the $-2^{\Delta\Delta Ct}$ method.

2.2.8. Data Analyses

Results from western blots were analyzed using the Image J software (1.4.3.67) to assess the levels of protein expressed in the cell lines and were normalized using GAPDH.

Differences between the mean lincRNA expression and mean GAPDH in qRT-PCR were determined using the student's *t*-test (Microsoft excel software, version 1908). All p-values were two-sided, and $p<0.05$ was considered statistically significant. Results presented are means ± standard deviation (SD) of the experiments.

2.3. RESULTS

2.3.1. Identification of Differentially Expressed Long Non-Coding RNAs in NB Tissues

A bioinformatics analysis was done comparing NB tumour samples with adjacent normal where available. An analysis done on the samples showed clustering of most NB tumour and normal samples on one end of the cluster diagram. One osteosarcoma tumour and an adjacent normal pair were found within the NB cluster (Figure 2a). The principal component analysis revealed large variances between the NB tumours compared to the normal samples as seen in Figure 2b.

A heat map was derived for differential gene expression in the NB and the OS samples compared to their respective normal tissues. Figure 3 shows the top 500 differentially expressed genes within the samples. Results show variations in expression between normal and cancer tissues as well as within normal tissue types seen in samples s37NC and s10NC. Figure 4 shows a heat map showing many differentially expressed lincRNAs in NB tissues compared to normal tissues, as derived after analysis. The derived lincRNAs were sorted based on the magnitude of differential fold expression of either up- or down-regulation in cancer tissues compared to normal tissues using DE foldlog2 values. The red colour on the heat maps indicates an up-regulation in the gene or lincRNA expression, and the blue colour indicates a down-regulation.

To elucidate lncRNAs that may potentially affect both pediatric cancers, literature research was done to identify lncRNAs that have physical interactions within the cancers. *LINC01133* and *LINC00261* were found to physically interact with EZH2 and GSK3β respectively in other cancers and showed differential expressions in NB tissues compared to the normal tissues.

2.3.2. *LINC01133, LINC00261* and *LINC01268* Show Little to no Expression in NB Tissues

The expression profiles of *LINC01133* and *LINC00261* were analyzed in NB tissue samples compared to the adjacent normal tissue. *LINC00261* was expressed in some adjacent normal tissues from the NB samples, with differential expression between two normal tissues (s10NC, s37NC), as seen in Figure 7a. Other tissue samples showed no expression of *LINC00261*

LINC01133 showed a high expression in one normal tissue (s37NC), low expression in one tumour sample, and no other tissues were found to express *LINC01133* (Figure 7b). *LINC01268* was downregulated in all neuroblastoma tissues compared to the normal tissues (Figure 9).

2.3.3. *LINC00261* and *LINC01133* are Variably Expressed in NB Cancer Cell Lines

The expression profiles of the *LINCRNA00261* and *LINC01133* were examined in NB cell lines to find a similarity to the samples provided and ascertain a potential relevance. The expression pattern of *LINC00261* was determined in NB cell lines by qRT-PCR using WI-38 as normal. There was a significant upregulation of the expression of *LINC00261* in the pediatric cell line IMR5 ($p<0.05$), and other cell lines showed varying levels of the *LINC00261* expression. Most cell lines expressed about the same amount of *LINC00261* compared to the normal cell line, with some showing some upregulation *LINC00261* expression, although it was not significant (Figure 11a).

The *LINC01133* expression pattern in NB cell lines compared to the WI-38 normal cell line showed significant upregulation in the GBM cell lines A172, M059K and in the NB pediatric line IMR5. There was significant downregulation in the pediatric cell lines SK-N-BE (2) and SK-N-MC (Figure 11b).

2.3.4. Subcellular Localisation Prediction and GO and KEGG Pathway Analyses

From the lncATLAS prediction, *LINC01133* was mostly expressed in the cytoplasm, and *LINC00261* was expressed in both the nuclear and cytoplasm cell compartments, as seen in Figure 6a and 6b, respectively.

A GO analysis determined which differentially expressed lncRNAs of interest were enriched in terms of biological themes. The top 20 genes enriched in the GO analysis for *LINC00261* and *LINC01133* are listed in Tables 4 and 5, respectively.

The analysis from the top 25 predicted target genes of the lncRNAs of interest were compared to the list of genes differentially expressed in the neuroblastoma tissues. The list from the comparison can be seen in Table 7.

2.3.5. GSK3β and ENX-1 expression in Pediatric NB Cell Lines

To determine a probable function of *LINC00261* and *LINC01133* in NB cell lines, the expressions of the proteins GSK3β and ENX-1 (EZH2), respectively, were determined and compared to GAPDH. All cell lines showed the expression of GSK3β, with SK-N-AS and SK-N-MC having a slightly weaker expression compared to the others. None of the pediatric cancers (IMR5, SK-N-AS, SK-N-BE [2]) expressed ENX-1; however, the other NB cell lines expressed this protein (Figure 13a, b).

CHAPTER 3: THE ROLE OF LONG NON-CODING RNAS IN OSTEOSARCOMA

3.0. Osteosarcoma

Osteosarcoma is the most common pediatric primary bone malignancy, most frequently occurring in children and adolescents (Endicott et al., 2017; Ritter & Bielack, 2010). The majority of osteosarcoma cases in children and adolescents are high-grade conventional osteosarcoma arising from mesenchymal cells that can acquire the capacity to produce osteoid and/or immature bones or that have the capacity to do so beginning in the intramedullary space of metaphyseal locations in long bones of the lower extremities (Gorlick & Khanna, 2010; Ritter & Bielack, 2010).

Osteosarcoma has a bimodal age distribution. The first peak incidence is within the first decade of life (most often during the adolescent growth spurt), and the second peak incidence is in older adults in which it presents as Paget's disease (Marina, 2004). The incidence of osteosarcoma is more common around puberty, occurring at the metaphysis of long bones and correlating with periods of rapid bone growth. Although it does occur in younger children, it is rare before age five (Endicott et al., 2017; Marina, 2004; Ritter & Bielack, 2010).

The disease has been found to occur about 1.4 times more frequently in males than in females, but it peaks earlier in females than in males. This suggests a correlation with increased rate and/or duration of rate of bone growth in males than in females as well as an earlier age of growth spurts in girls, respectively (Endicott et al., 2017; Marina, 2004; Ritter & Bielack, 2010).

Osteosarcoma is characterised by many complex somatic chromosomal abnormalities, including a high percentage of marker chromosomes (structurally abnormal chromosomes in which no part is identified), and it has a high heterogeneity. The ploidy number of the disease ranges from haploidy to near hexaploidy, and the multitude of alterations have made the disease difficult to study. Additional research is therefore warranted to understand the biology of the disease (Gorlick & Khanna, 2010; Marina, 2004; Romeo et al., 2010; Tang et al., 2008). Numerical and structural abnormalities, such as copy number alterations (CNAs) and structural variations (SVs), have been found in many osteosarcoma samples. Gains in chromosome 1, losses of chromosome 9, 10, 13 and/or 17 and partial and complete loss of the long arm of chromosome 6 have been found in osteosarcoma (OS) samples. Structural rearrangements of chromosomes 11, 19 and 20 have been found to also contribute to osteosarcoma. The most common include 1p11-13, 1q10-12, 1q21-22, 11p15, 14p11, 17p and 19q13 (Bridge et al., 1997; Chen et al., 2014; Marina, 2004; Tang et al., 2008). Loss of heterozygosity in chromosomes 3q, 13q, 17p and 18q were also observed in some OS cases, with losses of whole chromosomes being more common than gains (Bridge et al., 1997; Marina, 2004).

About 3% of sporadic OS tumours have been found to harbor genetic mutations in the *TP53* gene with translocations of the first intron of *TP53* identified to be unique to pediatric osteosarcoma (Chen et al., 2014), and data show that most OS tumours also show alterations in either the *TP53* or *RB1* gene. Survivors of retinoblastoma due to mutations in the *RB1* gene are also more likely to have a higher incidence of a second malignancy, the majority being osteosarcoma, and LFS predisposes affected individuals

to an increased risk of developing osteosarcoma (Bridge et al., 1997; Marina, 2004). Other hereditary disorders including Diamond-Black fan anemia, Rothmund-Thompson syndrome and hereditary neuroblastoma, have all been linked to a higher incidence of the disease (Endicott et al., 2017; Ritter & Bielack, 2010).

Like other cancers, changes to tumour suppressor genes and oncogenes may lead to the development of osteosarcoma. *INK4A* deletion and an amplification of the chromosome 12q13 region, which contains the murine double minute 2 (*MDM2*) gene product and the cyclin dependent kinase 4 (*CDK4*) gene, can affect the Rb and p53 pathways, and the overexpression of the human epidermal growth factor have been discovered in osteosarcoma tumour cells (Bridge et al., 1997; Marina, 2004). Mutations in the *ATRX* gene have also been found in some osteosarcoma tumours. *ATRX* is part of a multiprotein complex involved in regulating chromatin remodeling, nucleosome assembly and telomere maintenance (Chen et al., 2014). A loss of chromosomes 6q and 10p, which may harbor tumour suppressor genes, has been suggested to be relevant to the initiation of osteosarcoma (Bridge et al., 1997; Marina, 2004). Radiation therapy and radiation exposure have been identified as risk factors for the disease, and exposure to alkylating agents may also contribute to it (Endicott et al., 2017; Ritter & Bielack, 2010). A dysregulation in Wnt and Notch signaling pathways has also been suggested to play a role in osteosarcoma formation (Gorlick & Khanna, 2010; Marina et al., 2004; McIntyre et al., 1994).

The canonical Wnt pathway (Wnt/β-catenin dependent) plays a role in normal osteogenesis, which involves the maturation and differentiation of osteoblasts (Cleton-

Jansen et al., 2009; Hartmann, 2006). Its activation has been suggested to decrease osteosarcoma proliferation, while the Notch pathway is suggested to play vital roles in mesenchymal stem cell differentiation and cell fate decision making and therefore to possibly be of relevance to osteosarcoma initiation and progression (Cleton-Jansen et al., 2009; Gorlick & Khanna, 2010). Notch signaling may have in both oncogenic and tumour suppressor roles depending on the levels of expression and the cell type, and it is context-dependent (Engin et al., 2009).

The World Health Organisation (WHO) classifies conventional osteosarcoma into three major subtypes according to the microscopic appearance of the predominant matrix: Osteoblastic, Chondroblastic, and Fibroblastic. Osteoblastic OS has osteoid or bone as the predominant matrix type, with the appearance of the matrix varying from dense sheets of osteoid and/or woven bone to interlacing trabeculae to delicate, arborizing wisps of osteoid. It is composed of a malignant plasmacytoid to epithelioid osteoblasts with variable numbers of smaller round to ovoid cells, spindle cells and anaplastic mono- or multinucleated giant cells. Chondroblastic OS has a predominance of a chondroid matrix with many malignant cells within the lacunae. Fibroblastic OS is composed of malignant spindle cells with scant osteoid/bone. Other OS variants include telangiectactic and small cells (Ritter & Bielack, 2010).

Treatment options for osteosarcoma include surgical resection and/or radiotherapy and chemotherapy. Despite an improvement in successful treatment after the introduction of multimodal chemotherapy, survival rates are still not satisfactory, with only 60% of 5-year event-free survival with localized OS and ~80% of patients undergoing surgical treatment experiencing recurrence or metastasis (Cleton-Jansen et al., 2009; Zhang et al.,

2019). Relapses and recurrences of the disease are more difficult to treat, and alternative targets and treatment options must be considered.

3.1. Long Non-Coding RNAs in Osteosarcoma

Many studies on osteosarcoma have revealed numerous lncRNAs that have been found to contribute to the disease. A report by Li et al. (2016) briefly discusses some lncRNAs that have been found to be involved in osteosarcoma including, *TUG1* (taurine upregulated gene 1) and *MALAT1*, which were found to be upregulated in OS cells and to contribute to cell proliferation and have been suggested to play oncogenic roles in the disease. Other lncRNAs, including *PVT1, MIAT* (myocardial infarction associated transcript), *LINC00511, CCAT2* (colon cancer associated transcript 2), *HOTTIP* (homeobox A transcript at the distal tip) and *LOC730101*, have been found to be upregulated in OS tissues and cell lines and have potential oncogenic roles in osteosarcoma, contributing to proliferation and at times the invasion and metastasis of the disease. Some of the lncRNAs mentioned have also been reported to play oncogenic roles in other cancers, including breast cancers, non-small cell lung cancers and colorectal cancers (Cheng et al., 2018a, 2018b; Yan et al., 2018; Yan et al., 2019; Zhang et al., 2019; Zhou et al., 2016). LncRNAs such as *LOC285194* and *BC04057, NBR2, SRA1* and *CASC2* have been found to play tumour suppressive roles in osteosarcoma with low expressions in cell lines and tissues and an overexpression leading to suppressed proliferation and a reduction in migration and invasion (Ba et al., 2018; Cai et al., 2019; Guo et al., 2019; Xie et al., 2017).

The action of the lncRNAs differ, with some acting as competing endogenous RNAs to sponge miRNAs, thereby regulating them and their target genes to affect the progression of the disease. Other lncRNAs target genes and proteins that affect EMT markers and other pathways, such as the canonical Wnt pathway, which is believed to have an influence on OS progression (Cleton-Jansen et al., 2009).

PVT1 is upregulated in OS tissues and cells and plays an oncogenic role in OS. Silencing its expression was found to inhibit cell proliferation, migration and invasion It negatively correlates with miR-195, binding to the miRNA and it subsequently affecting its other target genes including *BCL2* (B-cell lymphoma family 2), *CCND1* (cyclin D1) and *FASN* (Fatty acid synthase). Silencing *PVT1* has been found to suppress *BCL2*, *CCND1* and *FASN*. BCL2 is thought to release cell death factors, such as cytochrome-c, into the cytoplasm, CCND1 is relevant to cell cycle progression in the G1 to S phase and FASN is a key enzyme for endogenous lipogenesis that is suggested to be associated with cancer metastasis (Zhou et al., 2016).

Upregulated *MIAT* in OS promotes vascular endothelial growth factor C (VEGFC) expression, which is an important factor in regulating cell proliferation, apoptosis and metastasis, by competitively binding miR-128-3p. *MIAT* and VEGFC show the same binding sites for miR-128-3p, suggesting that both would competitively bind miRNA. Because miR-128-3p also binds and targets VEGFC, lncRNA *MIAT* binding of the miRNA prevents it from reaching its target, contributing to OS progression (Zhang et al., 2019).

LINC00511 promotes the growth, migration and colony formation of OS cells partly by regulating miR-765. MiR-765 has been reported to inhibit angiogenesis-related markers in osteosarcoma. *LINC00511* upregulation was also found to affect epithelial-mesenchymal-transition (EMT) related genes. E-cadherin (epithelial marker) expression decreased with *LINC00511* upregulation, and increased vimentin and N-cadherin (mesenchymal markers) expression, thus enhancing EMT progression and further contributing to the invasion of the disease (Yan et al., 2019).

A high expression of *CCAT2* (colon cancer associated transcript 2) is associated with poor disease-free survival and shorter overall survival time in OS. Its upregulation has been reported to increase EMT progression, leading to a reduction in the expression of E-cadherin, and increased expression of N-cadherin, vimentin and snail. The overexpression of *CCAT2* is also reported to increase the expression of large tumour suppressor 2 (LATS2) and c-Myc, which are gene regulators that are involved in tumourigenesis and are reported to promote cancer progression (Little et al., 1983; Wu et al., 2016; Yan et al., 2018).

LncRNA *HOTTIP* has been found to be highly expressed in OS tissues and cell lines, and its overexpression may facilitate migration, invasion and EMT via a positive feedback loop with c-Myc. *HOTTIP* has been suggested to activate the Wnt/β-catenin pathway, and c-Myc is reported to be an effector of the pathway. The lncRNA also affects EMT-related expression, suggesting that *HOTTIP* contributes to the progression of osteosarcoma by EMT, as well as through the Wnt/β-catenin pathway (Tang & Ji, 2019).

LncRNA *LOC730101*, which is also upregulated in OS cell lines and tissues shows increased levels when the OS cells undergo energy stress. This suggests that the condition of the cells could influence the expression of lncRNAs, which could further affect disease progression (Cheng et al., 2018a, 2018b).

LncRNAs *LOC285194* and *BC04057* were identified as tumour suppressors in a study by Xie et al. (2017), which analysed copy number variations in osteosarcoma samples. Both lncRNAs were found to express low levels in OS cell lines, and their ectopic expression inhibited proliferation (Xie et al., 2017).

CASC2 (cancer susceptibility candidate 2) downregulation in OS samples was found to be correlated with advanced TNM stage, and its overexpression suppressed proliferation in OS cell lines. *CASC2* was found to be negatively correlated with miR-181a, which has been found to target *RASSF6* (Ras association domain family member 6), a tumour suppressor gene found to inhibit tumour growth, invasion and metastasis. MiR-181a has been reported to enhance the proliferation of gastric cancer cell lines and it is increased in OS cell lines, suggesting it may play an oncogenic role (Ba et al., 2018).

LncRNAs *SRA1* (steroid receptor activator 1) and *NBR2* have also been reported to be downregulated in OS samples, with the former involved in sponging miR-208a when overexpressed and the latter upregulating E-cadherin and downregulating N-cadherin, thus preventing EMT in OS. MiR-208a is reported to be upregulated in OS samples and to promote the proliferation of cells, and increased levels of *SRA1* decrease levels of miR-208a, suppress proliferation, and promote apoptosis in OS cell lines (Cai et al., 2019; Guo et al., 2019).

Many other lncRNAs continue to be discovered that play potentially biologically relevant roles in osteosarcoma and that could serve as prognostic markers, as well as therapeutic targets.

3.2. Materials and Methods

3.2.1. Cell Lines and Cell Culture

The osteosarcoma cell lines KHOS-312H, HOS, G-292, 143B and JC and normal cell line WI38 were cultured with appropriate media supplemented with 10% fetal bovine serum (FBS) and maintained in a humidified incubator at 37°C with 5% carbon dioxide (CO_2). 143B, KHOS-312H, HOS and WI38 were cultured in EMEM (Multicell, Wisent Inc.) and G-292 was cultured in McCoy's 5A (Multicell, Wisent Inc.). JC was cultured in DMEM (Multicell, Wisent Inc.).

3.2.2. RNA Extraction and Sequencing

Total RNA from OS tumour samples and normal tissue were stored at -80°C until RNA isolation. A section from each of the samples provided was obtained to be used in the study. Total RNA was extracted using the Direct-zol™ RNA Miniprep Plus Kit (Zymo) adhering to the manufacturer's protocol. The RNA quality and quantity were determined by a NanoDrop 2000c spectrophotometer (v. 1.5). The RNA was sequenced using Illumina® NextSeq 500/550 kits following the manufacturer's instructions. Tissue samples were provided as a part of POETIC Genomics Consortium collaboration. The study was approved by the U of L Human Subject Committee, protocol number 2016-064.

3.2.3. In-Depth Transcriptome Profiling and Bioinformatics

Sequenced data were analysed through bioinformatics to determine the differential expression of long non-coding RNAs between normal and tumour samples to identify

significant differences in expression. The data were compared with a range of human RNA libraries to find significant changes in lncRNA levels using CASAVA software (v.1.8).

To confirm that the lncRNAs of interest did not randomly associate with other genes, the long non-coding RNAs identified were subjected to a correlation analysis with protein coding genes using Pearson's correlation method, and adjustments for multiple comparisons were made using the Benjamini–Hochberg (BH) method to determine significant interactions. Significant genes were determined using the adjusted p-value (padj) < 0.05.

Gene ontology (GO) enrichment and the KEGG pathway analysis were used to determine the enrichment of the biological themes of protein coding genes significantly correlated with the lncRNAs using geneSCF v.1.1-p2. Genes with an absolute Pearson's correlation coefficient >0.75 and a padj <0.05 were considered significant. Enrichments were calculated using Fisher's exact test, after which there were multiple comparison adjustments made using several alternative methods.

The prediction of the subcellular localisation of the lncRNAs of interest was performed using the lncATLAS database (Mas-Ponte et al., 2017). RIblast software was used to predict RNA-RNA interactions between the lncRNAs of interest and target genes. The top 25 target genes were compared to genes differentially expressed between OS tumour and normal samples provided.

3.2.4. Protein Extraction and Quantification

Cells were harvested by dislodging using trypsin/EDTA and rinsing with a cold 1× phosphate buffer sulfate (PBS). The mixture was centrifuged at 2,000rpm for five minutes. The supernatant was discarded, and the pellets were solubilised in a laboratory-prepared 1% sodium dodecyl sulfate (SDS) lysis buffer (BioUltraPure, Bioshop) by sonication using a Braunsonic model 1510 sonicator (B. Braun Germany) operating at 80% sonication intensity. The lysates were centrifuged at 15,000rpm for 10 min, and the supernatant was decanted for use. The protein concentrations were determined using the Bradford protein assay with bovine serum albumin as the standard protein using a Nanodrop 2000c spectrophotometer (v. 1.5)

3.2.5. SDS-PAGE and Western Blotting Analysis

Total proteins (50–100 μg) were separated by electrophoresis on 8-10% sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE). The proteins were then electro-transferred onto activated polyvinyl difluoride (PVDF) membranes (Amersham Hybond P 0.45, GE Healthcare). The membranes were incubated for one hour in a blocking solution (5% dry skimmed milk in PBS, 0.5% Tween 20) at room temperature and incubated with specific primary antibodies at 4°C overnight. Primary antibodies anti-Gsk3α/β monoclonal antibody (1:200, Santa Cruz Biotechnology- sc-7291), anti-EZH2 polyclonal antibody (1:1000, Cell Signaling Technologies, 4905), anti-ENX-1 (D-8) monoclonal antibody (1:200, Santa Cruz Biotechnology, sc-137255) and anti-Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) monoclonal antibody (1:1000, Santa Cruz Biotechnology, sc-47724) were used. Blots of the primary antibodies tested were developed with peroxidase-labelled secondary antibodies specific to the primaries.

The membranes were washed (5 times of 5 min washing with PBS-Tween) before and after adding secondary antibodies. The membranes were run in duplicates.

3.2.6. Immunofluorescence and Chemiluminescence Detection

The protein bands of interest were detected using an enhanced chemiluminescence (ECL) system by incubating for five minutes in ECL detection reagents (GE Healthcare, Amersham Biosciences) and were visualized using the FluorChem HD2 ALC detection system (software v.3.2.2.0805, Cell Biosciences).

3.2.7. Quantitative Real-Time PCR

The expressions of long non-coding RNAs *LINC00261* and *LINC01133* in OS cell lines were determined using the SsoFast EvaGreen Supermix (Bio-Rad, Foster City, CA, USA) and the Bio-Rad CFX96^{TM} Real-Time System (Bio-Rad, Foster City, CA, USA). The total reaction volume was 20 µL and contained 10 µL of Ssofast Evagreen Supermix, one µL cDNA template, five µM of each primer (forward and reverse), and nuclease-free water adjusted to the total volume. Primer sequences used are listed in Table 4. Conditions for amplification were 95 °C for 30 s, followed by 49 cycles of 95 °C for 5 s and of 60 °C for 10 s. The melting curves were obtained by slow heating (0.5 °C/s) at temperatures in the range of 65 to 95 °C. All samples were run in triplicate, and GAPDH was used as the internal control. Both *LINC00261* and *LINC01133* levels were standardised to the internal control using the ΔCT (relative cycle), and the fold changes were calculated using the $-2^{\Delta\Delta Ct}$ method.

3.2.8. Data Analyses

The results from the Western blots were analysed using the Image J software (1.4.3.67) to assess the levels of protein expressed in the cell lines and normalised using GAPDH.

Differences between the means of lincRNA expression and GAPDH were determined using the student's *t*-test (Microsoft excel software, version 1908). All p-values were two-sided, and $p<0.05$ was considered statistically significant. The results presented are means ± standard deviation (SD) of the experiments.

3.3. RESULTS

3.3.1. Identification of Differentially Expressed Long Non-Coding RNAs In OS Tissues

A bioinformatics analysis was performed to compare OS tumour samples with adjacent normal. The analysis performed on the tissues showed clustering of the osteosarcoma tissue and adjacent normal tissues at one end of the cluster dendrogram. One osteosarcoma tissue and adjacent normal pair was found within the neuroblastoma cluster (Figure 2a). The principal component analysis showed little variation in tissues and normal types for OS, as shown in Figure 2b.

A heat map was derived for differential gene expression in both osteosarcoma and neuroblastoma samples compared to their respective normal tissues. Figure 3 shows the top 500 differentially expressed genes within the samples. The results showed variations in expression between normal and cancer tissues as well as within normal tissue types observed in samples s37NC and s10NC. Figure 5 shows a heat map of a large number of

differentially expressed lincRNAs in OS tissues compared to normal tissues derived after the analysis. LincRNAs were sorted based on the magnitude of the differential fold expression of up- or down-regulation in cancer tissues compared to normal tissues using DE foldlog2 values and lncRNAs of interest were determined based on high differential expression value. The red colour on the heat maps indicates an upregulation in gene or lincRNA expression, and blue colour indicates a down-regulation.

3.3.2. *LINC01133, LINC00261* and *LINC01139* are Differentially Expressed in OS Tissues

In osteosarcoma tissue and adjacent normal samples, there were varying levels of the expression of *LINC00261*. One pair of the samples showed a higher expression in tumour compared to normal tissues (s12NC, s12TM), and another pair showed a similar expression of *LINC00261*, with tumour samples having a slightly higher expression than the normal tissues (s16NC, s16TM). Other pairs showed a differential expression, with one pair having more expression in the normal tissues, while the other had more expression in the tumour (Figure 8b).

LINC01133 expression in osteosarcoma samples was also differential. Some pairs similarly expressed *LINC01133* (s16NC, s16TM and s30N, s30T), with a slightly higher expression in tumour than in normal tissues, and another pair had a much higher expression in the normal tissues compared to the tumour, which had a low expression of *LINC01133* (s12NC, s12TM). There was also a much higher expression in tumour samples compared to normal tissues in the last pair (s29N, s29TM) shown in Figure 8a.

LINC01139 was consistently upregulated in all osteosarcoma tissues compared to the normal tissues (Figure 10).

3.3.3. *LINC00261* and *LINC01133* Are Variably Expressed in OS Cancer Cell Lines

Overall *LINC00261* expression in the OS cell lines showed upregulation compared to the WI-38 normal cell line. The JC cell line was the only cell line that showed significant upregulation ($p<0.05$), as shown in Figure 12a.

LINC01133 showed a differential expression in the OS cell lines. Cell lines HOS and KHOS showed downregulation in expression compared to the control, although not significant. The JC, 143B and G292 cell lines showed significant upregulation in the expression of *LINC01133* compared to the normal cell line (Figure 12b).

3.3.4. GSK3β and ENX-1 Showed Variable Expression in OS Cell Lines

The expression of GSK3β and ENX-1 were used to deduce a possible function of lncRNAs *LINC00261* and *LINC01133*. There was increased expression of GSK3β across most cell lines except for G292 which showed a low expression compared to WI38 normal cell lines. WI38 showed a high expression of ENX-1, and as such, the expressions of ENX-1 in other OS cell lines were lower compared to the normal cell line. The lowest expression was found in cell lines 143B and JC (Figure 14a-d).

3.3.5. Subcellular Localisation Prediction, GO and the KEGG Pathway Analyses

Based on the lncATLAS prediction, *LINC01133* was mostly expressed in the cytoplasm, and *LINC00261* showed expression in both nuclear and cytoplasm cell compartments, as shown in Figures 6a and 6b, respectively.

A GO analysis determined which differentially expressed lncRNAs of interest were enriched in terms of biological themes. The top 20 genes enriched in the GO analysis for *LINC00261* and *LINC01133* are listed in Tables 4 and 5, respectively.

The analysis from the top 25 predicted target genes of the lncRNAs of interest were compared to the list of genes differentially expressed in the osteosarcoma tissues. The list from the comparison can be seen in Table 8.

CHAPTER 4: DISCUSSION

Long non-coding RNAs are no longer considered ‘junk’ material because they have been implicated in regulating gene expression and other biological process and they play significant roles in the initiation and progression of tumours (Brosnan & Voinnet, 2009; Rinn & Chang, 2012). The dysregulation or aberrant expression of these >200bp non-coding transcripts contribute to many diseases, including cancers, and their altered expression could affect tumourigenesis, acting as either oncogenes or tumour suppressors.

LINC00261 expression has been found to be downregulated in many cancers, including gastric cancer (Yu et al., 2017), hepatocellular carcinoma (HCC) (Zhang et al., 2018), non-small cell lung cancer (NSCLC) (Liao & Dong, 2019; Liu et al., 2017; Shi et al., 2019), lung cancer (Dhamija et al., 2018) and colon cancer (Yan et al., 2019), suggesting it plays a tumour-suppressive role; however, *LINC00261* has also been found to be upregulated in cholangiocarcinoma with an upregulation of lncRNA, which is indicative of a poor prognosis in patients and is associated with a poor five-year overall survival, suggesting that the lncRNA could also have an oncogenic role (Gao et al., 2020).

LINC01133 has also been documented to play roles in different cancers, acting either in an oncogenic or a tumour suppressive role. *LINC01133* has been found to play oncogenic roles in lung cancers (Zang et al., 2016; Zhang et al., 2015), pancreatic ductal adenocarcinoma (PDAC) (Huang et al., 2018) and hepatocellular carcinomas (Zheng, Zhang, & Bu, 2019). It has also been found to play tumour suppressive roles in breast

cancers (Song et al., 2019), gastric cancer (Yang et al., 2018), oral squamous cell carcinoma (OSCC) (Kong et al., 2018) and colorectal cancer (Kong et al., 2016; Zhang et al., 2017).

A study workflow of the work done is shown in Figure 15. Based on the results, both *LINC00261* and *LINC01133* were found to have varying expressions in both NB and OS tissues as well as in cell lines. *LINC01268* showed an increased expression in the normal tissues compared to the NB tumours and *LINC01139* expression was increased in OS tumours compared to adjacent normal tissues. In NB tissues, *LINC00261* and *LINC01133* expressions were mostly expressed in the normal tissues suggesting that they may play tumour-suppressive roles in NB.

LINC00261 was reported to interact physically with GSK3β in a tumour-suppressive capacity in gastric cancer cells when it was highly expressed (Yu et al., 2017). *LINC00261* enhanced interaction between Slug (Snai2) and GSK3β, leading to slug degradation. Slug mediates EMT in cells, and hence its degradation would decrease EMT in cancer. The aim was to determine whether there could be similar expression of *LINC00261* and therefore similar potential interactions between *LINC00261* and GSK3β in neuroblastoma and osteosarcoma cell lines for future studies. It was found that *LINC00261* expression mostly correlated with the expression of GSK3β in NB cell lines. IMR5 showed the highest *LINC00261* and GSK3β expressions, and SK-N-AS and SK-N-MC showed lower *LINC00261* and GSK3β expressions. This suggests a likely correlation between *LINC00261* and GSK3β expression, and that *LINC00261* could possibly target GSK3β in neuroblastoma; however, additional research warranted to confirm the action of *LINC00261* in neuroblastoma.

LINC00261 expression in OS tissues was mostly upregulated in the tumour compared to the normal tissues except for one pair. The expression of *LINC00261* was also upregulated in all OS cell lines, suggesting that lncRNA could possibly be oncogenic in OS. GSK3β was slightly upregulated in most of the OS cell lines, also suggesting a possible interaction with *LINC00261*. *LINC00261* subcellular localisation data shows that it can be found in both nuclear and cytoplasmic regions, suggesting that lncRNA could regulate gene expression both transcriptionally and post-transcriptionally depending on the downstream targets.

LINC01133 showed varying expressions in OS tissues, all tumours having a slightly higher expression of lncRNA compared to the normal tissues with the exception of one pair. This is generally consistent with the findings of Zeng et al. (2018), who found that *LINC01133* was upregulated in osteosarcoma tumours compared to adjacent normal tissues. The expression of *LINC01133* in most cell lines was also upregulated compared to the baseline expression.

LINC01133 was reported to interact with EZH2 and LSD1 in non-small cell lung cancer cells to repress KLF2, p21 and E-cadherin. EZH2 has been found to be overexpressed in cancers, whereas KLF2 expression is diminished (Zang et al., 2016). *LINC01133* also interacts with EZH2 in breast cancer (Song et al., 2019) and NSCLC (Zang et al., 2016) cell lines to inhibit invasion and metastasis in breast cancer when overexpressed, and impair proliferation and induce apoptosis in the NSCLC lines when expression is knocked-down. The aim was to ascertain whether there could be similar

potential interactions of *LINC01133* and EZH2 in neuroblastoma and osteosarcoma by studying the expression of *LINC01133* and EZH2 in the cell lines for future studies.

There was a low expression of *LINC01133* in NB pediatric cell lines SK-N-BE (2), SK-N-MC and SK-N-AS compared to the baseline expression correlated with ENX-1 (EZH2) expression, where there was no expression. In contrast, IMR5 showed a high expression for *LINC01133* but not for ENX-1. A low expression of *LINC01133* suggests that the lncRNA might have a tumour-suppressive role, and no expression of ENX-1 might imply that *LINC01133* does not have a transcriptional regulatory role in neuroblastoma. This could also suggest that the target gene of *LINC01133* might be different for NB, given that *LINC01133* targets a variety of genes in different cancers.

The ENX-1 (also known as EZH2) protein expression observed in OS cell lines was mostly decreased. Due to the contrast in the expression of *LINC01133* and ENX-1, it might be possible that they are negatively correlated; however, additional research is warranted to determine whether they do interact with each other in osteosarcoma. In Zeng et al.'s study, *LINC01133* was found to bind miR-422a, and miR-422a expression was found to suppress OS cell proliferation and invasion. The lncRNA-miRNA interaction was suggested to contribute to an increase in proliferation in OS cells. A high *LINC01133* expression was therefore suggested to have an oncogenic function in the disease (Zeng et al., 2018).

Interestingly, two studies on the role of *LINC01133* in ovarian cancer (OC) reported opposing roles of lncRNA in the tissues. *LINC01133* was reported to be poorly expressed in ovarian cancer tissues, and its overexpression had a tumour-suppressive

function via sponging miR-205 and upregulating LRRK2 (leucine rich repeat kinase 2) (Liu et al., 2019). Hou et al. (2018) found *LINC01133* to be highly expressed in another OC data set as well as in OC tissues. It was reported that the inhibition of *LINC01133* increased the apoptosis of OC cells and suppressed tumour formation. *LINC01133* was found to sponge miR-126 to promote tumourigenesis (Hou et al., 2018). *LINC01133* was found to be mostly expressed in the cytoplasm in the lncATLAS, and it may regulate gene expression post-transcriptionally. Downstream targets outside the nucleus should be considered.

The findings of Hou et al. (2018) and Liu et al. (2019) suggest that the downstream targets of *LINC01133* play crucial roles in regulating the progression of cancers. *LINC01133* sponged different miRNAs in both studies, which resulted in different outcomes in OC. This also highlights the relevance of lncRNA-miRNA interactions. Differences in the localisation of lncRNA as well as the tissue-specific pattern of expression of lncRNAs have also been suggested to affect the roles lncRNA may have related to cancers (Derrien et al., 2012; Song et al., 2019). The differences in localisation and how it affects gene regulation can also be observed in NSCLC, and *LINC01133* was found to be expressed mostly in the nucleus, where it regulated gene expression transcriptionally by interacting with EZH2 (Zang et al., 2016).

The GO enrichment analysis of the genes associated with the expression of *LINC00261* showed that the oxidation-reduction process and the positive regulation of cell division were affected by *LINC00261* expression. For *LINC01133*, the GO enrichment of the genes associated with the expression of the lncRNA revealed that

processes such as the regulation of cell-cell adhesion mediated by integrin and tissue morphogenesis are affected by its expression. This indicates that the disruption in the expression of lncRNAs has potential effects on the regulation of the cell.

LINC00261 and *LINC01133* have also been implicated in affecting canonical Wnt/ β-catenin pathways as well as EMT-related genes in cancers. EMT occurs during embryonic development, and it is also implicated in tumourigenesis, contributing to invasion and metastasis (Yu et al., 2017). LncRNAs that affect EMT would therefore be of importance. *LINC00261* was found to promote Slug (Snai2) degradation by enhancing GSK3β-Slug interactions in gastric cancers. Slug is one of the transcription factors that helps to mediate EMT in cells (Yu et al., 2017). *LINC00261* overexpression also increases E-cadherin expression in HCC and NSCLC cells (Liao & Dong, 2019; Zhang et al., 2018).

Canonical Wnt signaling (Wnt/ β-catenin) has multiple roles in osteoblastogenesis including regulating osteoblast lineage differentiation in early development and regulating osteoblast proliferation and maturation post-natally. This suggests that the deregulation of the pathway could influence bone diseases such as osteosarcoma by affecting the differentiation of mature osteoblasts. GSK3β protein is part of a multiprotein complex involved in regulating the pathway (Hartmann, 2006). The Wnt/β-catenin pathway is also essential in regulating embryonic development, and its aberrant expression is known to play a role in the pathogenesis of tumours including neuroblastoma (Zhang et al., 2014).

LINC01268 is downregulated across all tissues in neuroblastoma compared to the normal tissues. From the list of predicted target genes for *LINC01268* that were differentially expressed in neuroblastoma; none have been experimentally confirmed to interact according to literature. However, *LINC01268* and one of the genes (*CCL22*) have both been reported to be downregulated in human glioblastoma-associated microglia/monocytes samples (Szulzewsky et al., 2016). *LINC01268* has been found to be significantly upregulated in acute myeloid leukemia (AML) samples and the lincRNA correlated with poor overall survival(Lei et al., 2018). *LINC01268* functions as an enhancer for histone deacetylase 2 (HDAC2) and regulates its expression in AML. HDAC2 expression was also increased in AML and its inhibition was found to increase cell apoptosis and reduce cell proliferation of AML cells (Lei et al., 2018). *LINC01268* was also reported to have higher expression in glioma tumour samples (Matjasic et al., 2017).

LINC01139 is upregulated across all osteosarcoma tissue samples compared to the normal tissues. There have been no confirmed experimental interactions between *LINC01139* and the list of predicted target genes differentially expressed in osteosarcoma tissue in literature. However, *LINC01139* has been found to be positively correlated with SNX8 in HCC cells (Li et al., 2020). *LINC01139* showed increased expression in HCC tumour samples and correlated with poor overall survival. Knockdown of *LINC01139* led to increased apoptosis in HCC cells and decreased invasion of the HCC cells. *LINC01139* modulates MYBL2 expression by sponging members of the miR-30 family, therefore promoting HCC progression (Li et al., 2020). *LINC01139* was also found to both

positively and inversely correlate with some genes in HCC and diabetes mellitus (Liu et al., 2019) that were also differentially expressed in the osteosarcoma tissues.

The potential gene targets for all the lncRNAs of interest listed in the study suggests that the lncRNAs may perform regulatory roles in the cell and may therefore contribute to disease progression. The higher summed interaction energies of the predicted gene targets and the larger number of interactions suggests a higher potential for regulation between the lncRNA of interest and the gene.

In the current study, one limitation involves the normal cells used to determine expression changes. These were not of the same tissue of origin as the diseases studied. It would be worthwhile to confirm significant expression changes of the cell lines relative to a normal cell of similar origin. A small sample size as well as limited tissue samples may also have affected the outcomes of the study.

Future studies that could improve on this work could be to include a larger database in conjunction with more tissue samples to confirm differentially expressed lncRNAs. The functional mechanism of the long non-coding RNAs should also be verified. The physical interactions of the lncRNAs with proteins and/or genes that affect the cell cycle pathways as well as affect the progression of the cancers should be analyzed. Gain-of-function and loss-of-function studies can also be used to determine the effect of the lncRNAs on the progression of osteosarcoma and neuroblastoma cell lines. In-vivo studies may also be used to ascertain the impact the expression of the lncRNAs in biological systems. It would also be interesting to determine whether the TNM stage, age and metastasis of the cancers affect lncRNA expression, allowing for a possible

prognostic diagnosis. LncRNAs that show significant differential expression in pediatric cancers could be used in future to provide evidences for diagnostic and prognostic purposes. This could also help in its application for therapeutics which would help advance treatment options for the disease.

Tables and Figures

Table 1. Main Classification of Childhood Cancers with subgroups according to the International Classification of Cancer, Third Edition (ICCC-3).

Main Diagnostic Group		Subgroups
I.	Leukemias, myeloproliferative diseases, and myelodysplastic diseases	a. Lymphoid leukemias b. Acute myeloid leukemias c. Chronic myeloproliferative diseases d. Myelodysplastic syndrome and other myeloproliferative diseases e. Unspecified and other leukemias
II.	Lymphomas and reticuloendothelial neoplasms	a. Hodgkin lymphomas b. Non-Hodgkin lymphomas (except Burkitt lymphoma) c. Burkitt lymphomas d. Miscellaneous lymphoreticular neoplasms e. Unspecified lymphomas
III.	CNS and miscellaneous intracranial and intraspinal neoplasms	a. Ependymomas and choroid plexus tumour b. Astrocytomas c. Intracranial and intraspinal embryonal tumours d. Other gliomas e. Other specified intracranial and intraspinal neoplasms f. Unspecified intracranial and intraspinal neoplasms
IV.	Neuroblastoma and other peripheral nervous cell tumours	a. Unspecified intracranial and intraspinal neoplasms b. Unspecified intracranial and intraspinal neoplasms
V.	Retinoblastoma	
VI.	Renal tumours	a. Nephroblastoma and other nonepithelial renal tumours b. Renal carcinomas c. Unspecified malignant renal tumours
VII.	Hepatic tumours	a. Hepatoblastoma b. Hepatic carcinomas c. Unspecified malignant hepatic tumours
VIII.	Malignant bone tumours	a. Osteosarcomas b. Chondrosarcomas

		c. Ewing tumour and related sarcomas of bone d. Other specified malignant bone tumours e. Unspecified malignant bone tumours
IX.	Soft tissue and other extraosseous sarcomas	a. Unspecified malignant hepatic tumours b. Fibrosarcomas, peripheral nerve sheath tumours, and other fibrous neoplasms c. Kaposi sarcoma d. Other specified soft tissue sarcomas e. Unspecified soft tissue sarcomas
X.	Germ cell tumours, trophoblastic tumours, and neoplasms of gonads	a. Intracranial and intraspinal germ cell tumours b. Malignant extracranial and extragonadal germ cell tumours c. Malignant gonadal germ cell tumours d. Gonadal carcinomas e. Other and unspecified malignant gonadal tumours
XI.	Other malignant epithelial neoplasms and malignant melanomas	a. Adrenocortical carcinomas b. Thyroid carcinomas c. Nasopharyngeal carcinomas d. Malignant melanomas e. Skin carcinomas f. Other and unspecified carcinomas
XII.	Other and unspecified malignant neoplasms	a. Other specified malignant tumours b. Other unspecified malignant tumours

Table 2a. International Neuroblastoma Risk Group Staging System (INRGSS) (Monclair et al., 2009; Mueller & Matthay, 2009).

Stage	Description
L1	Localized tumour not involving vital structures as defined by the list of image-defined risk factors and confined to one body compartment and may be ipsilaterally continuous within body compartments
L2	Locoregional tumour with presence of one or more image-defined risk factors
M	Distant metastatic disease (except stage MS) and includes distant lymph node involvement
MS	Metastatic disease in children younger than 18 months with metastases confined to skin, liver, and/or bone marrow

Table 2b. International Neuroblastoma Staging System (INSS) (Monclair et al., 2009; Mueller & Matthay, 2009).

Stage	Description
1	Localized tumour with complete gross excision; ± microscopic residual disease; representative ipsilateral lymph node negative for tumour microscopically
2A	Localized tumour with incomplete gross excision; representative ipsilateral lymph node negative for tumour microscopically
2B	Localized tumour with or without complete gross excision; ipsilateral lymph node positive for tumour microscopically; enlarged contralateral lymph nodes should be negative microscopically
3	Unresectable unilateral tumour infiltrating across the midline; ± regional lymph node involvement; or localized unilateral tumour with contralateral regional lymph node involvement or midline tumour with bilateral extension by infiltration (unresectable) or by lymph node involvement
4	Any primary tumour with dissemination to distant lymph nodes, bone, bone marrow, liver, skin, or other organs
4S	Localized primary tumour in infants younger than 1 year of age (localized as in stage 1, 2A, or 2B) with dissemination limited to skin, liver, or bone marrow (< 10% malignant cells)

Table 3. Neuroblastoma and Osteosarcoma samples (n=18).

Sample ID	Status	Disease	Gender	Age
s10NC	Normal	Neuroblastoma	NA	35
s10TM	Tumour	Neuroblastoma	NA	35
s17TM	Tumour	Neuroblastoma	NA	5
s18TM	Tumour	Neuroblastoma	NA	34
s19_11_Pof009	Tumour	Neuroblastoma	NA	3
s23T	Tumour	Neuroblastoma	NA	3.5
s26T	Tumour	Neuroblastoma	NA	2
s33T	Tumour	Neuroblastoma	M	7
s37NC	Normal	Neuroblastoma	F	4
S37T	Tumour	Neuroblastoma	F	4
s12NC	Normal	Osteosarcoma	NA	16
s12TM	Tumour	Osteosarcoma	NA	16
s16NC	Normal	Osteosarcoma	NA	19
s16TM	Tumour	Osteosarcoma	NA	19
s29NC	Normal	Osteosarcoma	M	14
s29TM	Tumour	Osteosarcoma	M	14
s30N	Normal	Osteosarcoma	F	19
s30T	Tumour	Osteosarcoma	F	19

Table 4. Primer sequences used for quantitative real-time PCR (qRT-PCR).

Primer	Sequence
LINC01133-F	5′-GGCAAGGTGAACCTCAAAAA-3′
LINC01133-R	5′-TTCCTGCAAGAGGAGAAAGC-3′
LINC00261-F	5′-ACATTTGGTAGCCCGTGGAG-3′
LINC00261-R	5′-TCTTCCCCGGAGAACTAGCA-3′

Table 5. Top 20 Gene Ontology Analysis for *LINC00261*

Process name	Number of Genes	Gene Group
Oxidation-reduction process	14	528
Xenobiotic metabolic process	6	94
Lung-associated mesenchyme development	3	9
Omega-hydroxylase P450 pathway	3	10
Fatty acid biosynthetic process	4	40

Organic anion transport	3	15
Lung development	5	80
Immunoglobulin transcytosis in epithelial cells mediated by polymeric immunoglobulin receptor	2	2
Activation of phospholipase A2 activity	2	2
Regulation of opsonization	2	2
Positive regulation of cell division	4	47
Negative regulation of complement activation, classical pathway	2	3
Prostate epithelial cord elongation	2	3
Mesenchymal cell proliferation involved in lung development	2	3
Aminophospholipid transport	2	4
Aromatic compound catabolic process	2	4
Leukotriene B4 catabolic process	2	4
Negative regulation of epithelial cell proliferation	4	61
Blood coagulation	6	171
Steroid catabolic process	2	5

Table 6. Top 20 Gene Ontology Analysis for *LINC01133*

Process name	**Number of Genes**	**Gene Group**
Response to osmotic stress	3	19
Germinal center B cell differentiation	2	3
Regulation of cell-cell adhesion mediated by integrin	2	3
Tissue morphogenesis	2	3
Glycoprotein biosynthetic process	2	9
Kidney morphogenesis	2	9
Calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules	3	45
Intestinal absorption	2	14
Heparan sulfate proteoglycan biosynthetic process	2	15
Execution phase of apoptosis	2	18
O-glycan processing	3	60
Carbohydrate metabolic process	4	121
Protein localisation to Golgi apparatus	2	21
Lipid metabolic process	4	128
Regulation of arginine metabolic process	1	1
Immunoglobulin production in mucosal tissue	1	1

Regulation of systemic arterial blood pressure mediated by a chemical signal	1	1
CMP-N-acetylneuraminate biosynthetic process	1	1
Deoxyadenosine catabolic process	1	1
Oligopeptide transport	1	1

Table 7. Predicted gene targets of LncRNAs of interest differentially expressed in Neuroblastoma tissues.

LINC00261	**Summed interactions energy**	**Number of interactions**	**Description**
ACSM2A	-541.4897	21	acyl-CoA synthetase medium chain family member 2A [Source: HGNC Symbol; Acc: HGNC:32017]
ACSM2B	-569.9564	21	acyl-CoA synthetase medium chain family member 2B [Source: HGNC Symbol; Acc: HGNC:30931]
C9orf152	-471.1126	21	chromosome 9 open reading frame 152 [Source: HGNC Symbol; Acc: HGNC:31455]
CENPB	-465.9442	17	centromere protein B [Source: HGNC Symbol; Acc: HGNC:1852]
FAM111A	-538.3492	18	family with sequence similarity 111 member A [Source: HGNC Symbol; Acc: HGNC:24725]
FAM98B	-1075.1604	51	family with sequence similarity 98 member B [Source: HGNC Symbol; Acc: HGNC:26773]
GAP43	-950.9897	43	growth associated protein 43 [Source: HGNC Symbol; Acc: HGNC:4140]
GPR26	-540.2754	20	G protein-coupled receptor 26 [Source: HGNC Symbol; Acc: HGNC:4481]

IGF2	-591.4193	32	insulin like growth factor 2 [Source: HGNC Symbol; Acc: HGNC:5466]
LSAMP	-570.5523	23	limbic system-associated membrane protein [Source: HGNC Symbol; Acc: HGNC:6705]
MECOM	-548.1299	20	MDS1 and EVI1 complex locus [Source: HGNC Symbol; Acc: HGNC:3498]
PARVG	-624.2935	34	parvin gamma [Source: HGNC Symbol; Acc: HGNC:14654]
PDX1	-574.2179	28	pancreatic and duodenal homeobox 1 [Source: HGNC Symbol; Acc: HGNC:6107]
PLEKHA4	-611.3922	30	pleckstrin homology domain containing A4 [Source: HGNC Symbol; Acc: HGNC:14339]
PLEKHS1	-831.6093	30	pleckstrin homology domain containing S1 [Source: HGNC Symbol; Acc: HGNC:26285]
SNX8	-1047.821	58	sorting nexin 8 [Source: HGNC Symbol; Acc: HGNC:14972]
TAF3	-544.9148	25	TATA-box binding protein associated factor 3 [Source: HGNC Symbol; Acc: HGNC:17303]
USB1	-573.7218	25	U6 snRNA biogenesis phosphodiesterase 1 [Source: HGNC Symbol; Acc: HGNC:25792]
LINC01268			
CCL22	-152.5022	7	C-C motif chemokine ligand 22 [Source: HGNC Symbol; Acc: HGNC:10621]
CYP20A1	-221.7973	9	cytochrome P450 family 20 subfamily A member 1 [Source: HGNC Symbol; Acc: HGNC:20576]
FAM98B	-634.8042	31	family with sequence similarity 98 member B [Source: HGNC Symbol; Acc: HGNC:26773]

GNL3L	-230.2848	10	G protein nucleolar 3 like [Source: HGNC Symbol; Acc: HGNC:25553]
GPR68	-207.5766	11	G protein-coupled receptor 68 [Source: HGNC Symbol; Acc: HGNC:4519]
IRGQ	-225.856	9	immunity related GTPase Q [Source: HGNC Symbol; Acc: HGNC:24868]
NOL9	-173.5881	7	nucleolar protein 9 [Source: HGNC Symbol; Acc: HGNC:26265]
ORAI2	-220.3061	8	ORAI calcium release-activated calcium modulator 2 [Source: HGNC Symbol; Acc: HGNC:21667]
PARVG	-400.3241	21	parvin gamma [Source: HGNC Symbol; Acc: HGNC:14654]
PCBD2	-234.193	9	pterin-4 alpha-carbinolamine dehydratase 2 [Source: HGNC Symbol; Acc: HGNC:24474]
PDZD4	-202.7942	11	PDZ domain containing 4 [Source: HGNC Symbol; Acc: HGNC:21167]
RBMS2	-184.5863	7	RNA binding motif single stranded interacting protein 2 [Source: HGNC Symbol; Acc: HGNC:9909]
SCAI	-212.0198	9	suppressor of cancer cell invasion [Source: HGNC Symbol; Acc: HGNC:26709]
SLC25A51	-183.7751	7	solute carrier family 25 member 51 [Source: HGNC Symbol; Acc: HGNC:23323]
SLC6A17	-195.1953	10	solute carrier family 6 member 17 [Source: HGNC Symbol; Acc: HGNC:31399]
SNX8	-662.2965	34	sorting nexin 8 [Source: HGNC Symbol; Acc: HGNC:14972]
SYNCRIP	-160.1443	9	synaptotagmin binding cytoplasmic RNA interacting protein [Source: HGNC Symbol;Acc:HGNC:16918]

TMEM120B	-190.0096	8	transmembrane protein 120B [Source: HGNC Symbol; Acc: HGNC:32008]
TOR1AIP2	-178.8972	7	torsin 1A interacting protein 2 [Source: HGNC Symbol; Acc: HGNC:24055]
ZNF264	-181.5751	7	zinc finger protein 264 [Source: HGNC Symbol; Acc: HGNC:13057]
ZNF490	-211.2918	8	zinc finger protein 490 [Source: HGNC Symbol; Acc: HGNC:23705]
ZNF738	-182.3049	7	zinc finger protein 738 [Source: HGNC Symbol; Acc: HGNC:32469]
ZNF850	-167.3959	7	zinc finger protein 850 [Source: HGNC Symbol; Acc: HGNC:27994]
LINC01133			
A1BG	-691.832	21	alpha-1-B glycoprotein [Source: HGNC Symbol;Acc:HGNC:5]
ARSA	-933.6488	46	arylsulfatase A [Source: HGNC Symbol;Acc:HGNC:713]
CNOT6L	-570.7348	34	CCR4-NOT transcription complex subunit 6 like [Source: HGNC Symbol;Acc:HGNC:18042]
EFNA5	-1388.479	78	ephrin A5 [Source: HGNC Symbol;Acc:HGNC:3225]
EIF4EBP2	-667.6579	28	eukaryotic translation initiation factor 4E binding protein 2 [Source: HGNC Symbol;Acc:HGNC:3289]
FEM1C	-647.6936	25	fem-1 homolog C [Source: HGNC Symbol;Acc:HGNC:16933]
FOXP1	-615.0819	34	forkhead box P1 [Source: HGNC Symbol;Acc:HGNC:3823]
HAUS5	-609.4399	33	HAUS augmin like complex subunit 5 [Source: HGNC Symbol;Acc:HGNC:29130]
HIF3A	-560.4812	27	hypoxia inducible factor 3 alpha subunit [Source: HGNC Symbol;Acc:HGNC:15825]

KCNJ2	-1101.8074	66	potassium voltage-gated channel subfamily J member 2 [Source: HGNC Symbol;Acc:HGNC:6263]
KIN	-1181.1896	61	Kin17 DNA and RNA binding protein [Source: HGNC Symbol;Acc:HGNC:6327]
LPP	-786.5945	37	LIM domain containing preferred translocation partner in lipoma [Source: HGNC Symbol;Acc:HGNC:6679]
MTPAP	-595.3981	29	mitochondrial poly(A) polymerase [Source: HGNC Symbol;Acc:HGNC:25532]
NSL1	-580.0256	33	NSL1, MIS12 kinetochore complex component [Source: HGNC Symbol;Acc:HGNC:24548]
PARP11	-713.2726	12	poly(ADP-ribose) polymerase family member 11 [Source: HGNC Symbol;Acc:HGNC:1186]
PCYT1B	-564.0685	14	phosphate cytidylyltransferase 1, choline, beta [Source: HGNC Symbol;Acc:HGNC:8755]
PGPEP1	-589.4255	14	pyroglutamyl-peptidase I [Source: HGNC Symbol;Acc:HGNC:13568]
PHC3	-594.6789	26	polyhomeotic homolog 3 [Source: HGNC Symbol;Acc:HGNC:15682]
RSPO4	-664.0217	20	R-spondin 4 [Source: HGNC Symbol; Acc: HGNC:16175]
S1PR2	-1361.0499	75	sphingosine-1-phosphate receptor 2 [Source: HGNC Symbol; Acc: HGNC:3169]
SLC9A7	-567.2491	25	solute carrier family 9 member A7 [Source: HGNC Symbol; Acc: HGNC:17123]
SNX27	-1162.3881	37	sorting nexin family member 27 [Source: HGNC Symbol; Acc: HGNC:20073]
TLCD2	-765.1659	37	TLC domain containing 2 [Source: HGNC Symbol; Acc: HGNC:33522]

ZNF738	-555.6835	24	zinc finger protein 738 [Source: HGNC Symbol;Acc:HGNC:32469]

Table 8. Predicted gene targets of LncRNAs of interest differentially expressed in Osteosarcoma tissues.

LINC00261	**Summed interaction energy**	**Number of interactions**	**Description**
ACSM2A	-541.4897	21	acyl-CoA synthetase medium chain family member 2A [Source: HGNC Symbol; Acc: HGNC:32017]
ACSM2B	-569.9564	21	acyl-CoA synthetase medium chain family member 2B [Source: HGNC Symbol; Acc: HGNC:30931]
CENPB	-465.9442	17	centromere protein B [Source: HGNC Symbol; Acc: HGNC:1852]
FAM111A	-538.3492	18	family with sequence similarity 111 member A [Source: HGNC Symbol; Acc: HGNC:24725]
FAM98B	-1075.1604	51	family with sequence similarity 98 member B [Source: HGNC Symbol; Acc: HGNC:26773]
GPR26	-540.2754	20	G protein-coupled receptor 26 [Source: HGNC Symbol; Acc: HGNC:4481]
LSAMP	-570.5523	23	limbic system-associated membrane protein [Source: HGNC Symbol; Acc: HGNC:6705]

MECOM	-548.1299	20	MDS1 and EVI1 complex locus [Source: HGNC Symbol; Acc: HGNC:3498]
MYOCD	-547.9394	19	myocardin [Source: HGNC Symbol; Acc: HGNC:16067]
NBPF20	-4160.1166	203	NBPF member 20 [Source: HGNC Symbol; Acc: HGNC:32000]
PARVG	-624.2935	34	parvin gamma [Source: HGNC Symbol; Acc: HGNC:14654]
PDX1	-574.2179	28	pancreatic and duodenal homeobox 1 [Source: HGNC Symbol; Acc: HGNC:6107]
PLEKHS1	-831.6093	30	pleckstrin homology domain containing S1 [Source: HGNC Symbol; Acc: HGNC:26285]
SNX8	-1047.821	58	sorting nexin 8 [Source: HGNC Symbol; Acc: HGNC:14972]
TAF3	-544.9148	25	TATA-box binding protein associated factor 3 [Source: HGNC Symbol; Acc: HGNC:17303]
USB1	-573.7218	25	U6 snRNA biogenesis phosphodiesterase 1 [Source: HGNC Symbol; Acc: HGNC:25792]
LINC01133			
A1BG	-691.832	21	alpha-1-B glycoprotein [Source: HGNC Symbol; Acc: HGNC:5]
ARSA	-933.6488	46	arylsulfatase A [Source: HGNC Symbol; Acc: HGNC:713]
CNOT6L	-570.7348	34	CCR4-NOT transcription complex subunit 6 like [Source: HGNC Symbol; Acc: HGNC:18042]
EFNA5	-1388.479	78	ephrin A5 [Source: HGNC Symbol; Acc: HGNC:3225]
EIF4EBP2	-667.6579	28	eukaryotic translation initiation factor 4E binding protein 2 [Source: HGNC Symbol; Acc: HGNC:3289]

ESRRG	-981.9133	25	estrogen related receptor gamma [Source: HGNC Symbol; Acc: HGNC:3474]
FEM1C	-674.6936	25	fem-1 homolog C [Source: HGNC Symbol; Acc: HGNC:16933]
FOXP1	-615.0819	34	forkhead box P1 [Source: HGNC Symbol; Acc: HGNC:3823]
HAUS5	-609.4399	33	HAUS augmin like complex subunit 5 [Source: HGNC Symbol; Acc: HGNC:29130]
HIF3A	-560.4812	27	hypoxia inducible factor 3 alpha subunit [Source: HGNC Symbol; Acc: HGNC:15825]
KCNJ2	-1101.8074	66	potassium voltage-gated channel subfamily J member 2 [Source: HGNC Symbol; Acc: HGNC:6263]
KIN	-1181.1896	61	Kin17 DNA and RNA binding protein [Source: HGNC Symbol; Acc: HGNC:6327]
LPP	-786.5945	37	LIM domain containing preferred translocation partner in lipoma [Source: HGNC Symbol; Acc: HGNC:6679]
MTPAP	-595.3981	29	mitochondrial poly(A) polymerase [Source: HGNC Symbol; Acc: HGNC:25532]
NSL1	-580.0256	33	NSL1, MIS12 kinetochore complex component [Source: HGNC Symbol; Acc: HGNC:24548]
PARP11	-713.2726	12	poly(ADP-ribose) polymerase family member 11 [Source: HGNC Symbol; Acc: HGNC:1186]
PCYT1B	-564.0685	14	phosphate cytidylyltransferase 1, choline, beta [Source: HGNC Symbol; Acc: HGNC:8755]
PGPEP1	-589.4255	14	pyroglutamyl-peptidase I [Source: HGNC Symbol; Acc: HGNC:13568]
PHC3	-594.6789	26	polyhomeotic homolog 3 [Source: HGNC Symbol; Acc: HGNC:15682]

RSPO4	-664.0217	20	R-spondin 4 [Source: HGNC Symbol; Acc: HGNC:16175]
S1PR2	-1361.0499	75	sphingosine-1-phosphate receptor 2 [Source: HGNC Symbol; Acc: HGNC:3169]
SLC9A7	-567.2491	25	solute carrier family 9 member A7 [Source: HGNC Symbol; Acc: HGNC:17123]
SNX27	-1162.3881	37	sorting nexin family member 27 [Source: HGNC Symbol; Acc: HGNC:20073]
TLCD2	-765.1659	37	TLC domain containing 2 [Source: HGNC Symbol; Acc: HGNC:33522]
ZNF738	-555.6835	24	zinc finger protein 738 [Source: HGNC Symbol; Acc: HGNC:32469]
LINC01139			
C6orf89	-212.437	9	chromosome 6 open reading frame 89 [Source: HGNC Symbol; Acc: HGNC:21114]
CENPB	-389.653	15	centromere protein B [Source: HGNC Symbol; Acc: HGNC:1852]
FAM111A	-232.372	10	family with sequence similarity 111 member A [Source: HGNC Symbol; Acc: HGNC:24725]
FAM98B	-766.137	34	family with sequence similarity 98 member B [Source: HGNC Symbol; Acc: HGNC:26773]
FIZ1	-284.46	14	FLT3 interacting zinc finger 1 [Source: HGNC Symbol; Acc: HGNC:25917]
GAB2	-218.745	11	GRB2 associated binding protein 2 [Source: HGNC Symbol; Acc: HGNC:14458]
GATAD2B	-249.553	23	GATA zinc finger domain containing 2B [Source: HGNC Symbol;Acc:HGNC:30778]
GPR68	-385.128	15	G protein-coupled receptor 68 [Source: HGNC Symbol; Acc: HGNC:4519]

HDDC2	-242.955	10	HD domain containing 2 [Source: HGNC Symbol;Acc:HGNC:21078]
HEXIM1	-224.651	11	HEXIM P-TEFb complex subunit 1 [Source: HGNC Symbol; Acc: HGNC:24953]
KRT9	-277.154	15	keratin 9 [Source: HGNC Symbol; Acc: HGNC:6447]
LHFPL4	-218.23	11	LHFPL tetraspan subfamily member 4 [Source: HGNC Symbol; Acc: HGNC:29568]
LMNTD1	-258.8708	6	lamin tail domain containing 1 [Source: HGNC Symbol; Acc: HGNC:26683]
MYOCD	-268.8245	12	myocardin [Source: HGNC Symbol; Acc: HGNC:16067]
NBPF20	-880.6611	53	NBPF member 20 [Source: HGNC Symbol; Acc: HGNC:32000]
NFAM1	-245.569	13	NFAT activating protein with ITAM motif 1 [Source: HGNC Symbol; Acc: HGNC:29872]
NUTM2A	-254.3284	11	NUT family member 2A [Source: HGNC Symbol; Acc: HGNC:23438]
PARVG	-723.323	32	parvin gamma [Source: HGNC Symbol; Acc: HGNC:14654]
PDZD4	-296.885	12	PDZ domain containing 4 [Source: HGNC Symbol; Acc: HGNC:21167]
PNMA8B	-238.035	11	PNMA family member 8B [Source: HGNC Symbol; Acc: HGNC:29206]
SNX8	-1008.96	41	sorting nexin 8 [Source: HGNC Symbol; Acc: HGNC:14972]
SYNCRIP	-236.92	10	synaptotagmin binding cytoplasmic RNA interacting protein [Source: HGNC Symbol; Acc: HGNC:16918]
ZNF326	-220.393	10	zinc finger protein 326 [Source: HGNC Symbol; Acc: HGNC:14104]

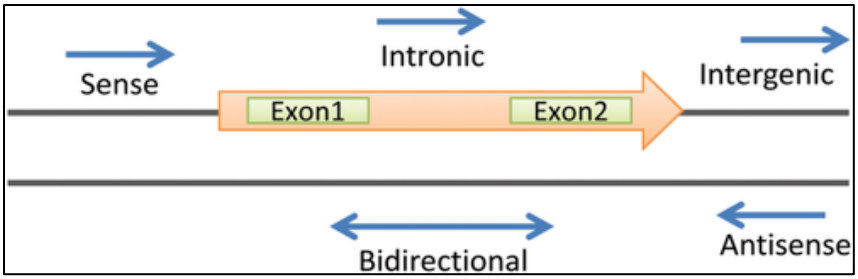


Figure 1a. Origin of lncRNA transcripts. LncRNAs can be distinct from or can overlap regions that can encode for mRNA and proteins. The relationship of lncRNAs relative to that of nearby protein-coding genes can be used to describe their location. Orange arrow represents a protein-coding region. LncRNAs are shown using blue arrows (Takahashi et al., 2014)

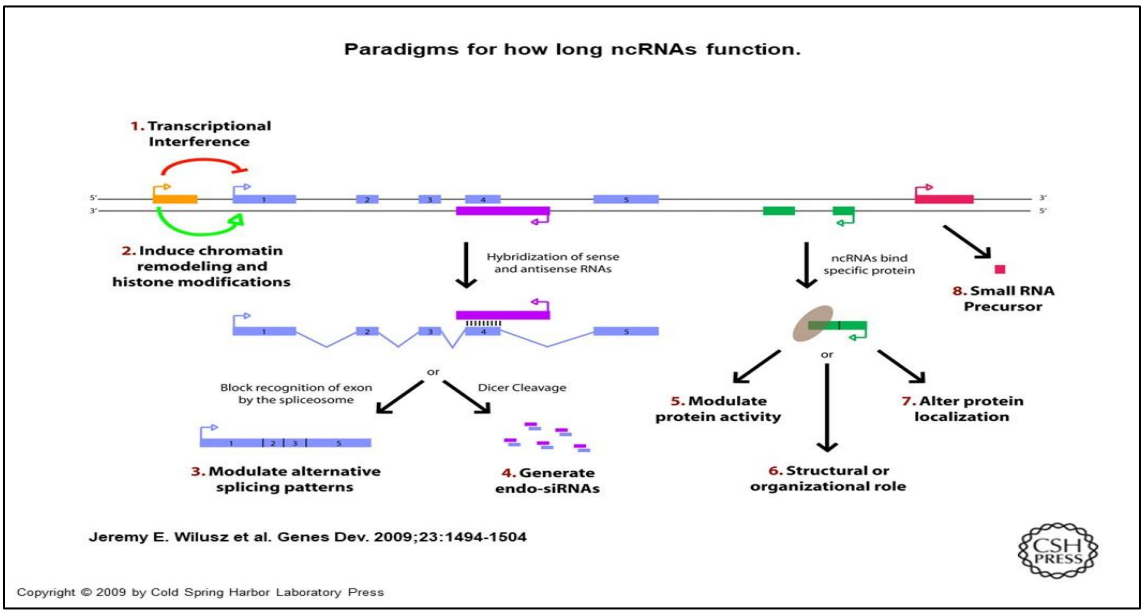


Figure 1b. Functional roles of long non-coding RNAs in transcriptional and post-transcriptional roles as depicted by Wilusz et al., 2009.

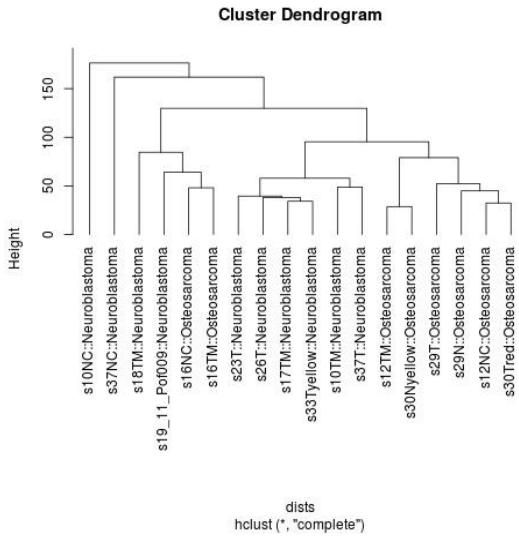


Figure 2a. Cluster dendrogram of Neuroblastoma (NB) and Osteosarcoma (OS) tissues. Pairs of OS tissue and adjacent normal n=4. NB adjacent normal tissue n=2; NB tumour tissue n=8.

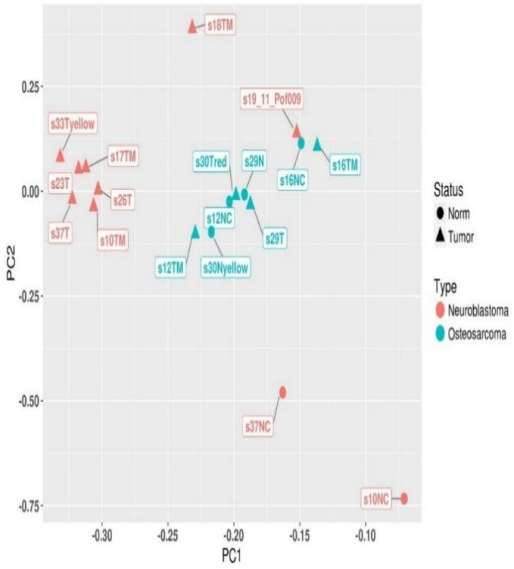


Figure 2b. Principal Component Analysis plot of NB and OS tissues. Pairs of OS tissue and adjacent normal n=4. NB adjacent normal tissue n=2; NB tumour tissue n=8.

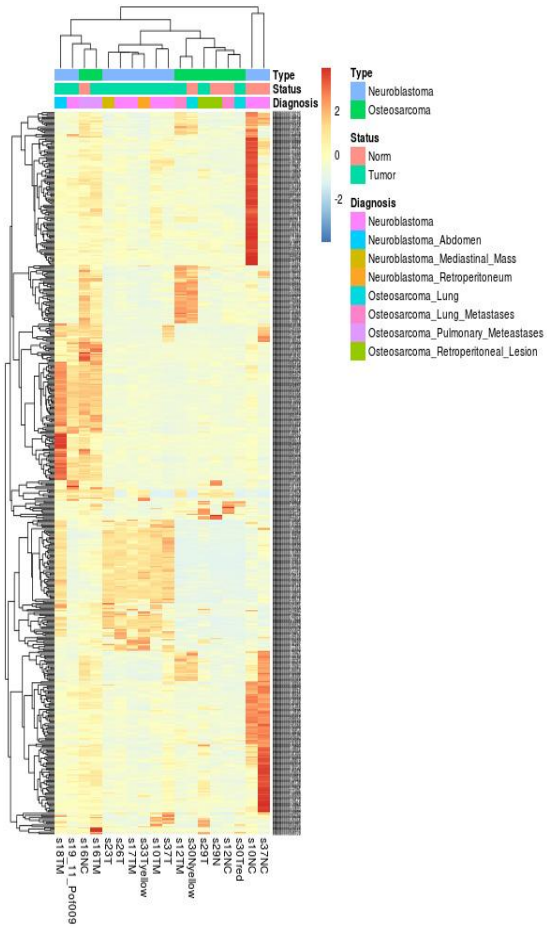


Figure 3. Heat map of Neuroblastoma and Osteosarcoma tissue samples showing differential expression of genes. Red denotes up-regulation, blue denotes down-regulation. In NB tissues, gene expression is compared to adjacent normal tissue (or closest normal by age). Pairs of OS tissue and adjacent normal n=4. NB adjacent normal tissue n=2; NB tumour tissue n=8.

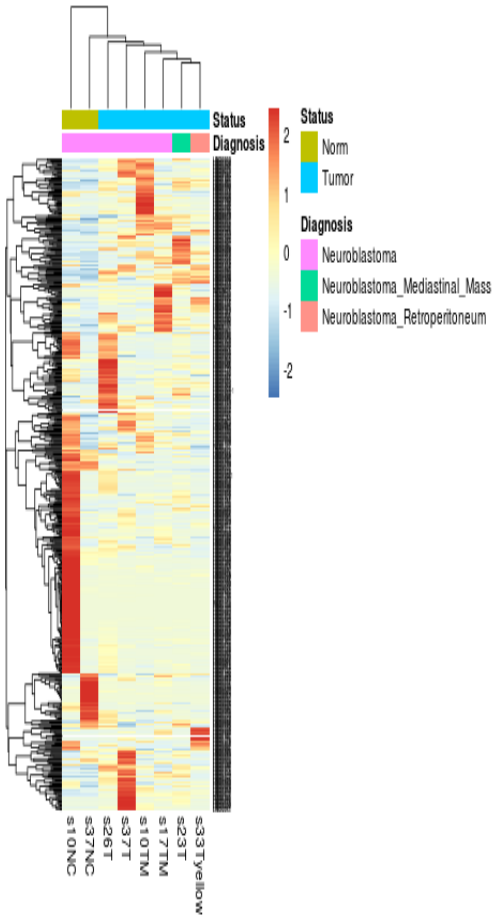


Figure 4. Heat map of Neuroblastoma tissue samples showing differential expression of long non-coding RNAs. Red denotes up-regulation, blue denotes down-regulation. NB adjacent normal tissue n=2; NB tumour tissue n=6.

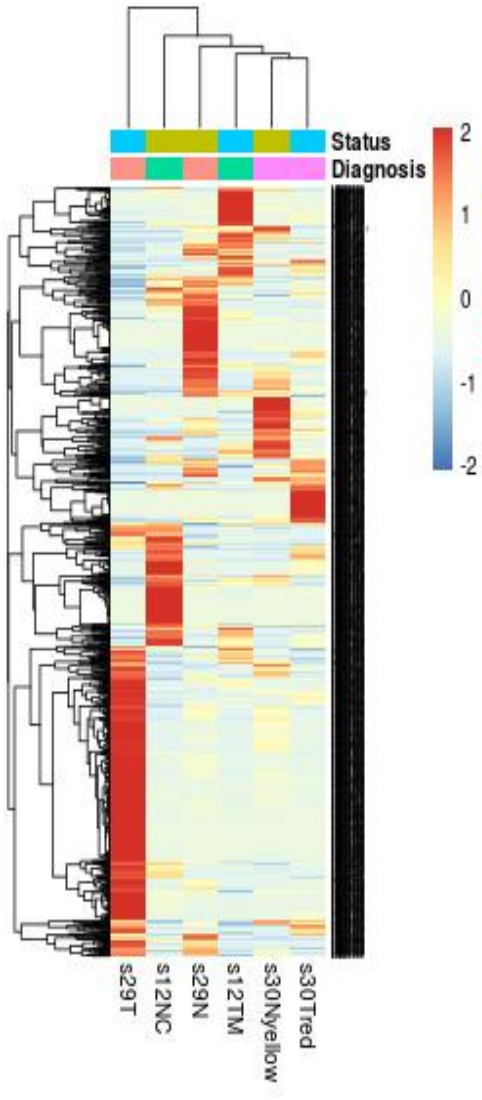
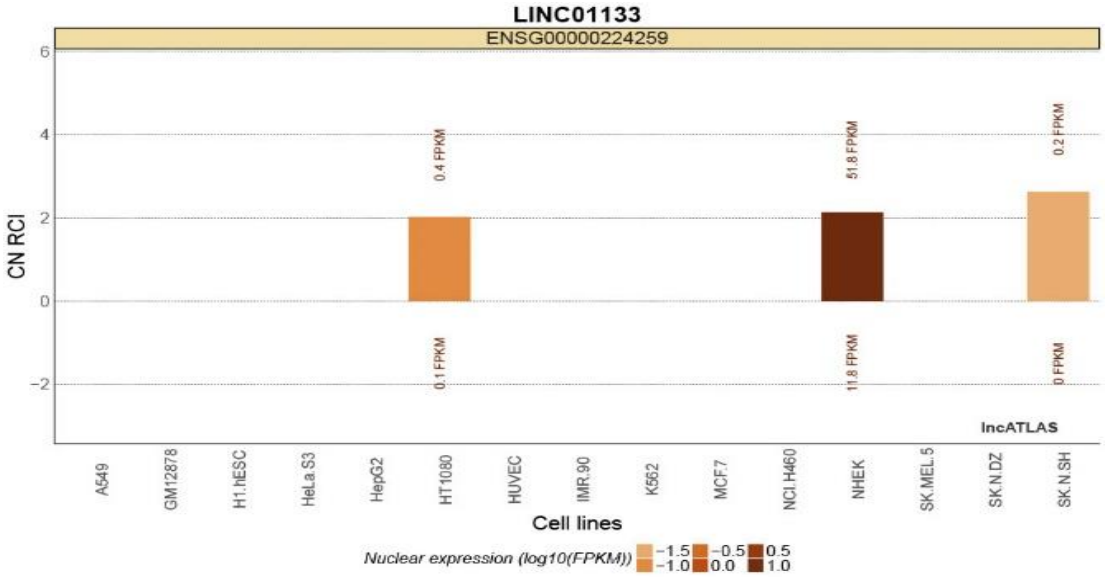


Figure 5. Heatmap of Osteosarcoma tissue samples showing differential expression of long non-coding RNAs. Red denotes up-regulation, blue denotes down-regulation. Pairs of OS tissue and adjacent normal n=3.

A)



B)

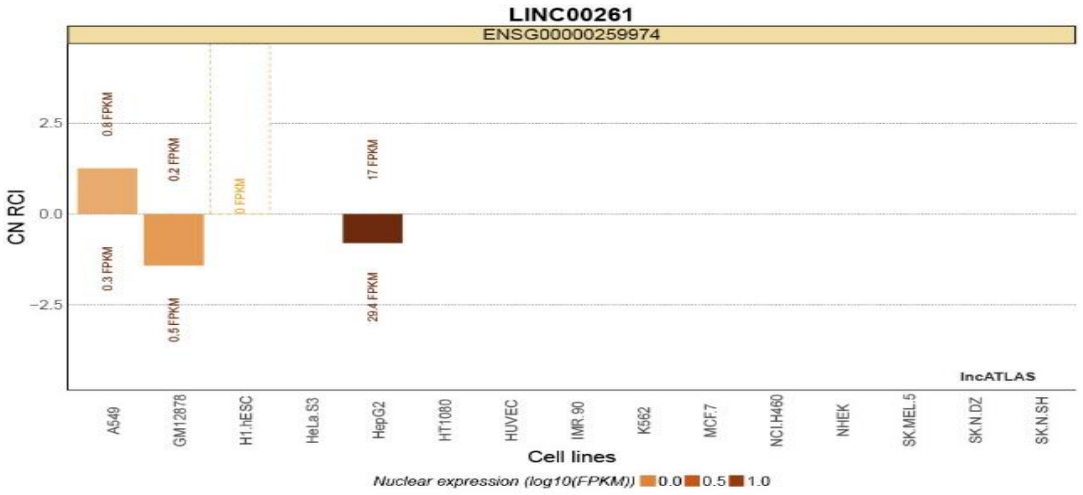


Figure 6A). The prediction of subcellular localisation of LINC0133 using long non-coding RNA ATLAS. **B)** The prediction of subcellular localisation of *LINC00261* using long non-coding RNA ATLAS.

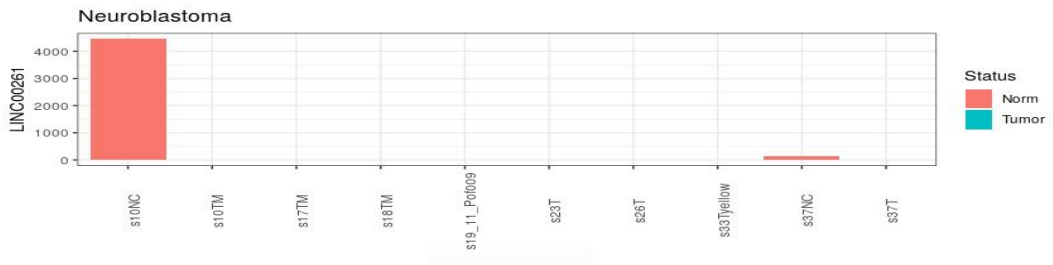


Figure 7a. Differential expression of long non-coding RNA *LINC00261* in neuroblastoma tissue samples. NB adjacent normal tissue n=2 ; NB tumour tissue n=8.

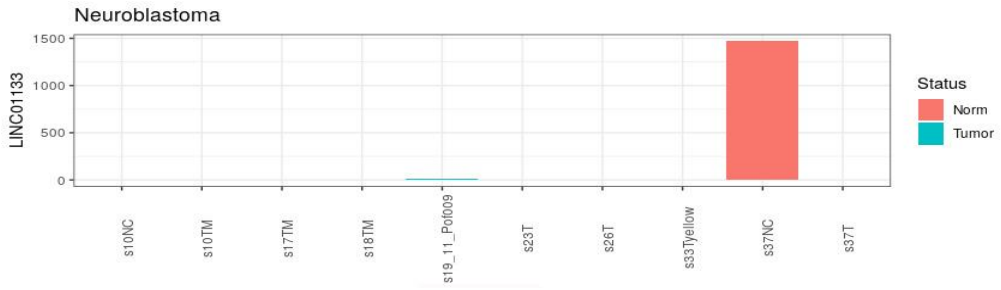


Figure 7b. Differential expression of long non -coding RNA *LINC01133* in neuroblastoma tissue samples. NB adjacent normal tissue n=2; NB tumour tissue n=8.

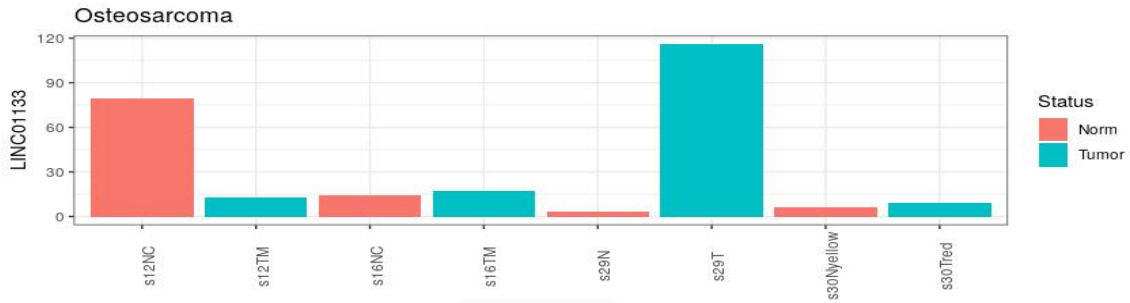


Figure 8a. Differential expression of long non-coding RNA *LINC01133* in osteosarcoma tissue samples. Pairs of OS tissue and adjacent normal n=4.

Figure 8b. Bar graph showing differential expression of long non-coding RNA *LINC00261* in osteosarcoma tissue samples. Pairs of OS tissue and adjacent normal n=4

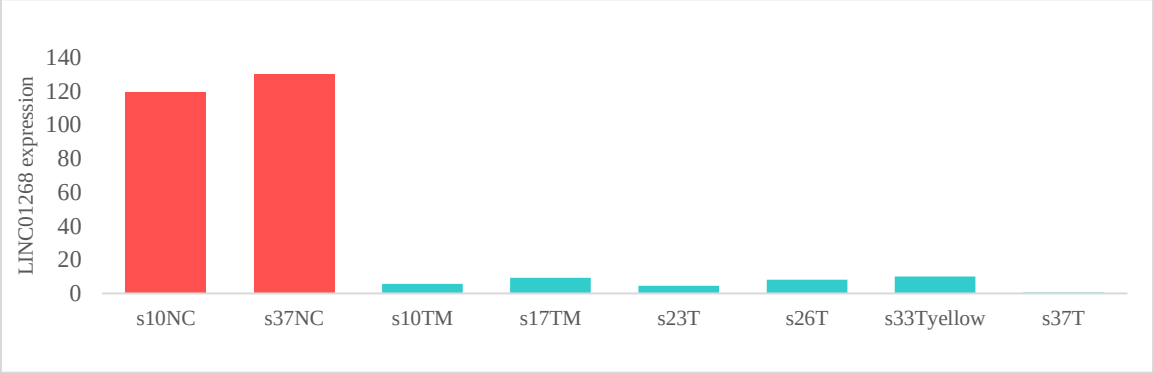


Figure 9. Differential expression of long non-coding RNA *LINC01268* in neuroblastoma tissue samples. NB adjacent normal tissue n=2; NB tumour tissue n=6.

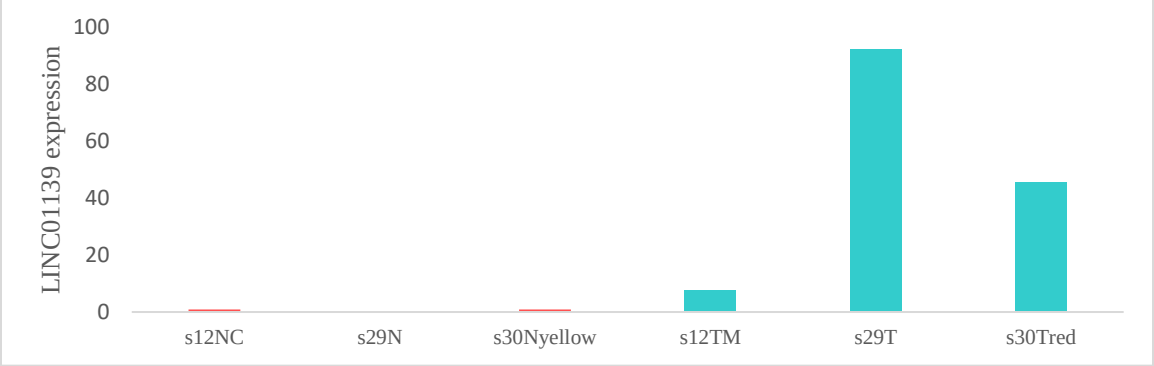


Figure 10. Differential expression of long non-coding RNA *LINC01139* in osteosarcoma tissue samples. Pairs of OS tissue and adjacent normal n=3.

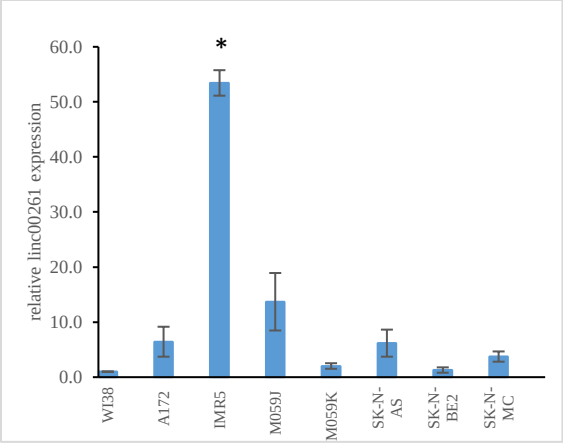


Figure 11a. Relative expression of *LINC00261* in brain cell lines from qRT-PCR. NB cell lines n =4; n=3 GBM. Bars are mean ± SD *denotes p<0.05

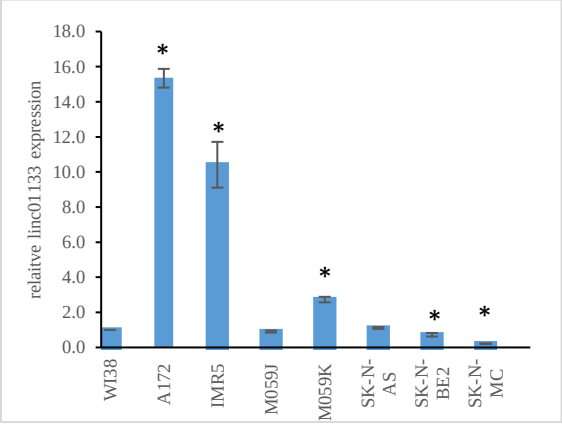


Figure 11b. Relative expression of *LINC01133* in brain cell lines from qRT-PCR. NB cell lines n =4; n=3 GBM. Bars are mean ± SD *denotes p<0.05

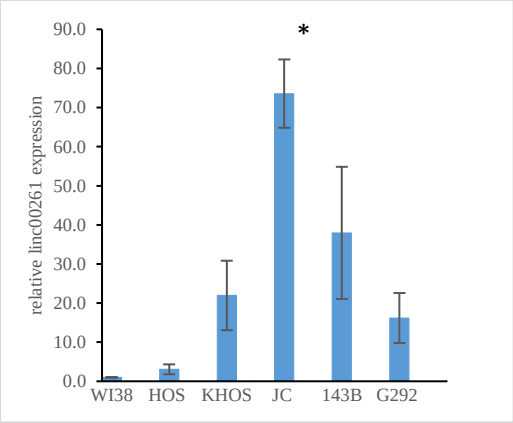


Figure 12a. Relative expression of *LINC00261* in OS cell lines; n=5. Bars are mean ± SD *denotes p<0.05.

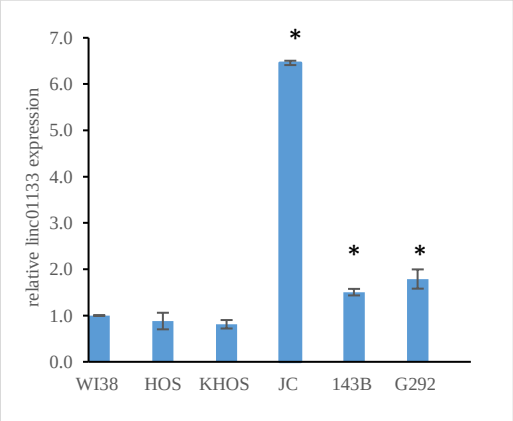


Figure 12b. Relative expression of *LINC01133* in OS cell lines; n=5. Bars are mean ± SD *denotes p<0.05

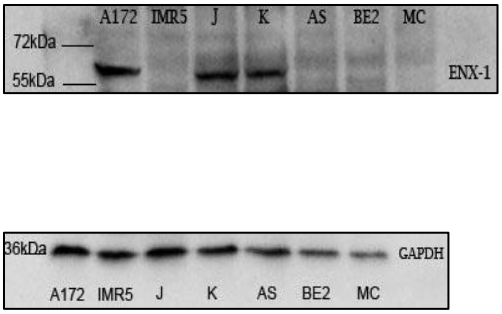


Figure 13a. ENX-1 (EZH2) protein expression in brain cell lines compared to GAPDH control. NB cell lines n =4; GBM, n=3.

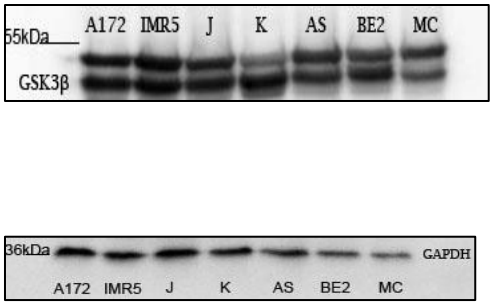


Figure 13b. GSK3β protein expression in brain cell lines compared to GAPDH control. NB cell lines n =4; GBM, n=3.

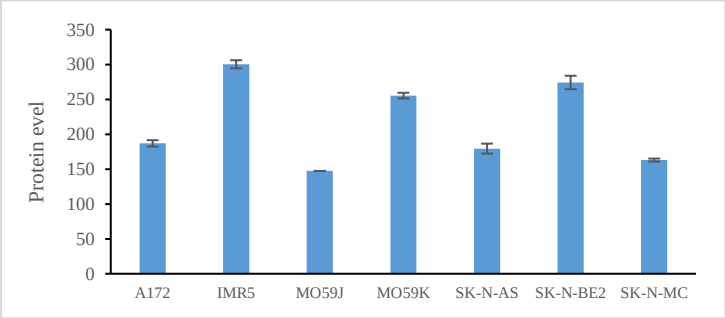


Figure 13c. Bar plot of protein expression of GSK3β in brain cell lines from Image J analysis. NB cell lines n =4; GBM, n=3

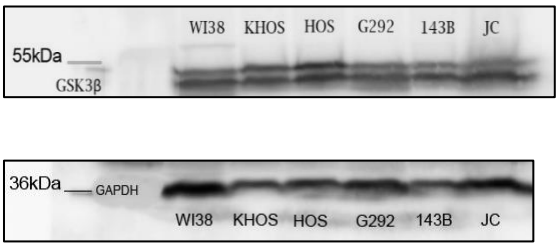


Figure 14a. GSK3β protein expression in OS cell lines; n=5 and WI-38 normal.

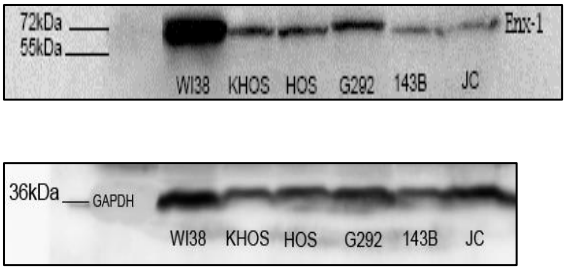


Figure 14b. ENX-1(EZH2) protein expression in osteosarcoma cell lines; n=5 and WI-38 normal compared to GAPDH control

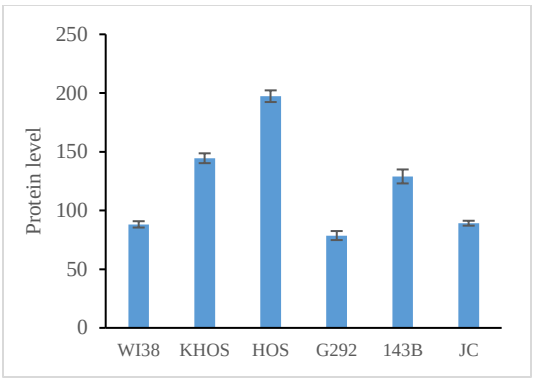


Figure 14c. Bar plot of GSK3β protein expression in OS cell lines; n=5 and WI-38 from Image J analysis.

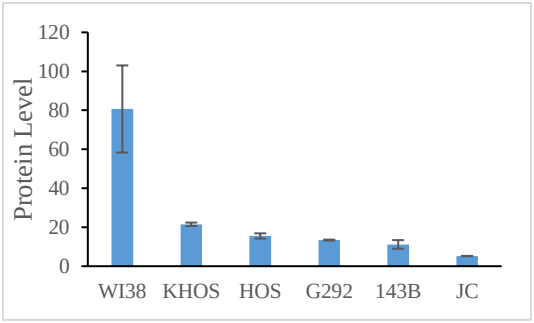


Figure 14d. Bar plot of ENX-1 (EZH2) expression in osteosarcoma cell lines; n=5 and WI-38 from Image J analysis.

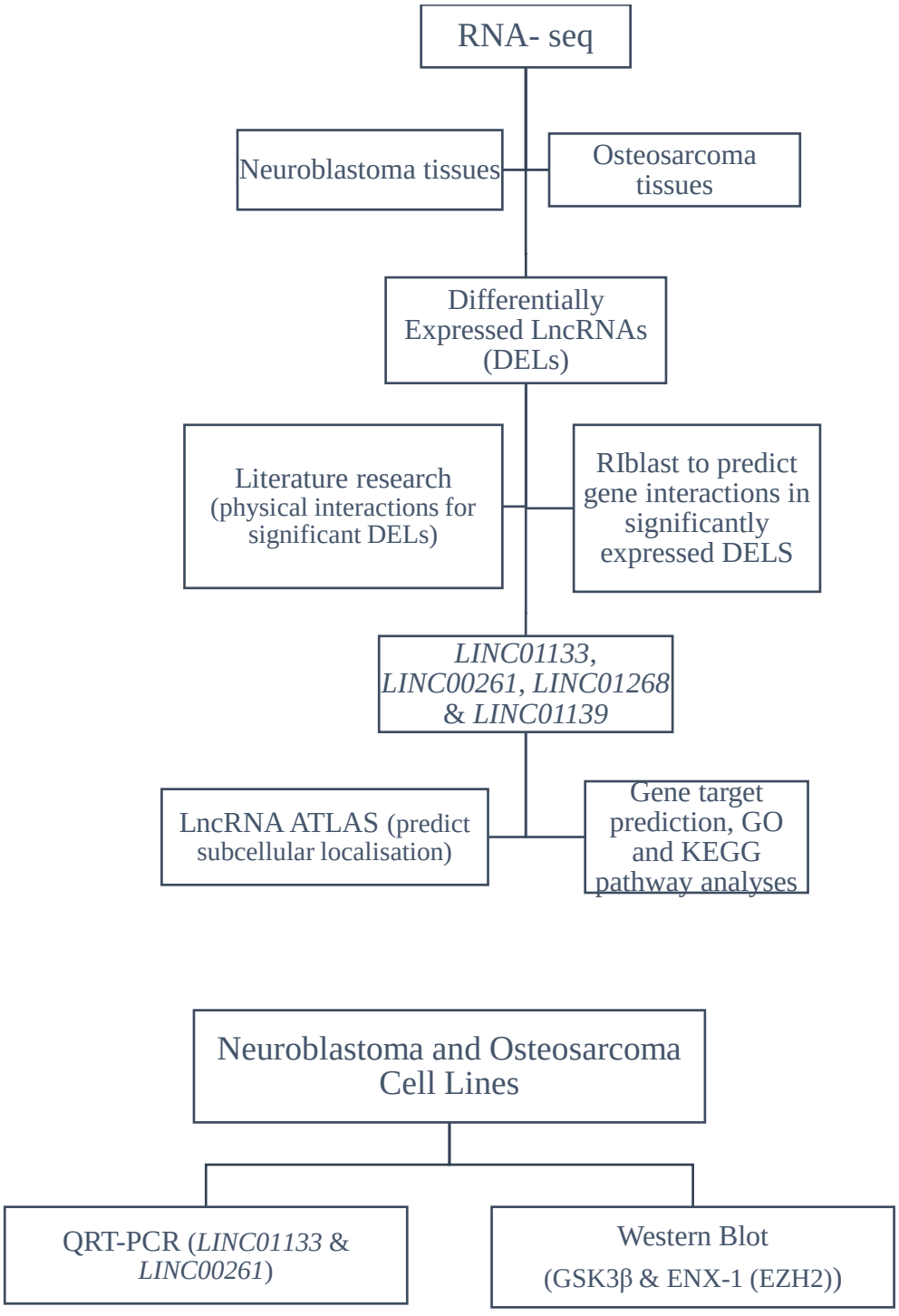


Figure 15. Study workflow to identify potentially relevant lncRNAs in neuroblastoma and osteosarcoma.

REFERENCES

Alaiyan, B., Ilyayev, N., Stojadinovic, A., Izadjoo, M., Roistacher, M., Pavlov, V., … Nissan, A. (2013). Differential expression of colon cancer associated transcript1 (CCAT1) along the colonic adenoma-carcinoma sequence. *BMC Cancer, 13*(1), 196. https://doi.org/10.1186/1471-2407-13-196.

Attiyeh, E. F., London, W. B., Mossé, Y. P., Wang, Q., Winter, C., Khazi, D., … Maris, J. M. (2005). Chromosome 1p and 11q deletions and outcome in neuroblastoma. *New England Journal of Medicine, 353*(21), 2243–2253. https://doi.org/10.1056/NEJMoa052399

Ba, Z., Gu, L., Hao, S., Wang, X., Cheng, Z., & Nie, G. (2018). Downregulation of lncRNA CASC2 facilitates osteosarcoma growth and invasion through miR-181a. *Cell Proliferation, 51*(1), e12409. https://doi.org/10.1111/cpr.12409

Bevilacqua, V., Gioia, U., Di Carlo, V., Tortorelli, A. F., Colombo, T., Bozzoni, I., … Caffarelli, E. (2015). Identification of linc-NeD125, a novel long non coding RNA that hosts miR-125b-1 and negatively controls proliferation of human neuroblastoma cells. *RNA Biology, 12*(12), 1323–1337. https://doi.org/10.1080/15476286.2015.1096488

Bi, S., Wang, C., Li, Y., Zhang, W., Zhang, J., Lv, Z., & Wang, J. (2017). LncRNA-MALAT1-mediated Axl promotes cell invasion and migration in human neuroblastoma. *Tumor Biology, 39*(5), 101042831769979. https://doi.org/10.1177/1010428317699796

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development, 16*(1), 6–21. https://doi.org/10.1101/gad.947102

Bountali, A., Tonge, D. P., & Mourtada-Maarabouni, M. (2019). RNA sequencing reveals a key role for the long non-coding RNA MIAT in regulating neuroblastoma and glioblastoma cell fate. *International Journal of Biological Macromolecules, 130*(1), 878–891. https://doi.org/10.1016/j.ijbiomac.2019.03.005

Bridge, J. A., Nelson, M., McComb, E., McGuire, M. H., Rosenthal, H., Vergara, G., … Neff, J. R. (1997). Cytogenetic findings in 73 osteosarcoma specimens and a review of the literature. *Cancer Genetics and Cytogenetics, 95*(1), 74–87. https://doi.org/10.1016/S0165-4608(96)00306-8

Brosnan, C. A., & Voinnet, O. (2009). The long and the short of noncoding RNAs. *Current Opinion in Cell Biology*, *21*(3), 416–425. https://doi.org/10.1016/j.ceb.2009.04.001

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., & Willard, H. F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, *349*(6304), 38–44. https://doi.org/10.1038/349038a0

Cai, W., Wu, B., Li, Z., He, P., Wang, B., Cai, A., & Zhang, X. (2019). LncRNA NBR2 inhibits epithelial-mesenchymal transition by regulating Notch1 signaling in osteosarcoma cells. *Journal of Cellular Biochemistry*, *120*(2), 2015–2027. https://doi.org/10.1002/jcb.27508

Canadian Cancer Society's Advisory Committee on Cancer Statistics. (2019). Canadian Cancer Statistics 2019. Toronto, ON, ON: Canadian Cancer Society.

Canadian Cancer Society's Advisory Committee on Cancer Statistics, & Canadian Cancer Society. (2017). *Canadian Cancer Statistics 2017*. *Canadian Cancer Society* (Vol. 2017). https://doi.org/0835-2976

Canadian Cancer Society. (2017). Canadian Cancer Statistics 2017. *Canadian Cancer Society*, *2017*, 1–132. https://doi.org/0835-2976

Chen, L., Wang, W., Cao, L., Li, Z., & Wang, X. (2016). Long non-coding RNA CCAT1 acts as a competing endogenous rna to regulate cell growth and differentiation in acute myeloid leukemia. *Molecules and Cells*, *39*(4), 330–336. https://doi.org/10.14348/molcells.2016.2308

Chen, X., Bahrami, A., Pappo, A., Easton, J., Dalton, J., Hedlund, E., … Dyer, M. A. (2014). Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. *Cell Reports*, *7*(1), 104–112. https://doi.org/10.1016/j.celrep.2014.03.003

Chen, Y., Takita, J., Choi, Y. L., Kato, M., Ohira, M., Sanada, M., … Ogawa, S. (2008). Oncogenic mutations of ALK kinase in neuroblastoma. *Nature*, *455*(7215), 971–974. https://doi.org/10.1038/nature07399

Cheng, G., Li, B., Sun, X., Cao, Z., Zhang, G., Zhao, Z., … Liu, W. (2018a). Corrigendum to "LncRNA LOC730101 promotes osteosarcoma cell survival under energy stress" [Biochem. Biophys. Res. Commun. 496 (1) (2018) 1–6]. *Biochemical*

and Biophysical Research Communications, *499*(2), 396. https://doi.org/10.1016/j.bbrc.2018.03.115

Cheng, G., Li, B., Sun, X., Cao, Z., Zhang, G., Zhao, Z., … Liu, W. (2018b). LncRNA LOC730101 promotes osteosarcoma cell survival under energy stress. *Biochemical and Biophysical Research Communications*, *496*(1), 1–6. https://doi.org/10.1016/j.bbrc.2017.12.074

Cheng, J., Su, H., Zhu, R., Wang, X., Peng, M., Song, J., & Fan, D. (2014). Maternal coffee consumption during pregnancy and risk of childhood acute leukemia: A metaanalysis. *American Journal of Obstetrics and Gynecology*, *210*(2), 151-e1. https://doi.org/10.1016/j.ajog.2013.09.026

Cheung, N.-K. V. (2012). Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *JAMA*, *307*(10), 1062. https://doi.org/10.1001/jama.2012.228

Chiaroni-Clarke, R. C., Munro, J. E., & Ellis, J. A. (2016). Sex bias in paediatric autoimmune disease – Not just about sex hormones? *Journal of Autoimmunity*, *69*, 12–23. https://doi.org/10.1016/j.jaut.2016.02.011

Cleton-Jansen, A.-M., Anninga, J. K., Briaire-de Bruijn, I. H., Romeo, S., Oosting, J., Egeler, R. M., … Hogendoorn, P. C. W. (2009). Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, *101*(11), 1909–1918. https://doi.org/10.1038/sj.bjc.6605405

Colnaghi, R., Carpenter, G., Volker, M., & O'Driscoll, M. (2011). The consequences of structural genomic alterations in humans: Genomic Disorders, genomic instability and cancer. *Seminars in Cell and Developmental Biology*. https://doi.org/10.1016/j.semcdb.2011.07.010

Costa, F. F. (2010). Non-coding RNAs: Meet thy masters. *BioEssays*. https://doi.org/10.1002/bies.200900112

D'Angio, G., Evans, A., & Koop, C. E. (1971). Special pattern of widespread neuroblastoma with a favourable prognosis. *The Lancet*, *297*(7708), 1046–1049. https://doi.org/10.1016/S0140-6736(71)91606-0

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., … Guigo, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of

their gene structure, evolution, and expression. *Genome Research*, *22*(9), 1775–1789. https://doi.org/10.1101/gr.132159.111

Dhamija, S., Becker, A., Sharma, Y., Myacheva, K., Seiler, J., & Diederichs, S. (2018). *LINC00261* and the adjacent gene FOXA2 are epithelial markers and are suppressed during lung cancer tumorigenesis and progression. *Non-Coding RNA*, *5*(1), 2. https://doi.org/10.3390/ncrna5010002

Downing, JR, Wilson, RK, Zhang, J, Mardis, ER, Pui, C-H, Ding, L, Ley, TJ & Evans, W. (2012). Erratum: The pediatric cancer genome project. *Nature Genetics*, *44*(9), 1072.

Downing, J. R., Wilson, R. K., Zhang, J., Mardis, E. R., Pui, C.-H., Ding, L., … Evans, W. E. (2012). The pediatric cancer genome project. *Nature Genetics*, *44*(6), 619–622. https://doi.org/10.1038/ng.2287

Endicott, A. A., Morimoto, L. M., Kline, C. N., Wiemels, J. L., Metayer, C., & Walsh, K. M. (2017). Perinatal factors associated with clinical presentation of osteosarcoma in children and adolescents. *Pediatric Blood and Cancer*, *64*(6), e26349.https://doi.org/10.1002/pbc.26349

Engin, F., Bertin, T., Ma, O., Jiang, M. M., Wang, L., Sutton, R. E., … Lee, B. (2009). Notch signaling contributes to the pathogenesis of human osteosarcomas. *Human Molecular Genetics*, *18*(8), 1464–1470. https://doi.org/10.1093/hmg/ddp057

Erwin, J. A., & Lee, J. T. (2008). New twists in X-chromosome inactivation. *Current Opinion in Cell Biology*, *20*(3), 349–355. https://doi.org/10.1016/j.ceb.2008.04.007

Feng, C., Zhao, Y., Li, Y., Zhang, T., Ma, Y., & Liu, Y. (2019). LncRNA MALAT1 promotes lung cancer proliferation and gefitinib resistance by acting as a mir-200a sponge. *Archivos de Bronconeumología*, *55*(12), 627–633. https://doi.org/10.1016/j.arbres.2019.03.026

Galderisi, U., Di Bernardo, G., Cipollaro, M., Peluso, G., Cascino, A., Cotrufo, R., & Melone, M. A. B. (1999). Differentiation and apoptosis of neuroblastoma cells: Role of N-myc gene product. *Journal of Cellular Biochemistry*, *73*(1), 97–105. https://doi.org/10.1002/(SICI)1097-4644(19990401)73:1<97::AID-JCB11>3.0.CO;2-M

Gao, J., Qin, W., Kang, P., Xu, Y., Leng, K., Li, Z., … Zhong, X. (2020). Up-regulated *LINC00261* predicts a poor prognosis and promotes a metastasis by EMT process in

cholangiocarcinoma. *Pathology - Research and Practice*, *216*(1), 152733. https://doi.org/10.1016/j.prp.2019.152733

George, R. E., Sanda, T., Hanna, M., Fröhling, S., II, W. L., Zhang, J., … Look, A. T. (2008). Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature*, *455*(7215), 975–978. https://doi.org/10.1038/nature07397

Gestblom, C., Hoehner, J. C., Hedborg, F., Sandstedt, B., & Påhlman, S. (1997). In vivo spontaneous neuronal to neuroendocrine lineage conversion in a subset of neuroblastomas. *American Journal of Pathology*, *150*(1), 107.

Gorlick, R., & Khanna, C. (2010). Osteosarcoma. *Journal of Bone and Mineral Research*, *25*(4), 683–691. https://doi.org/10.1002/jbmr.77

Guo, W., Jiang, H., Li, H., Li, F., Yu, Q., Liu, Y., … Zhang, M. (2019). LncRNA-SRA1 suppresses osteosarcoma cell proliferation while promoting cell apoptosis. *Technology in Cancer Research & Treatment*, *18*, 153303381984143. https://doi.org/10.1177/1533033819841438

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., … Chang, H. Y. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, *464*(7291), 1071–1076. https://doi.org/10.1038/nature08975

Gutschner, T., Hammerle, M., Eissmann, M., Hsu, J., Kim, Y., Hung, G., … Diederichs, S. (2013). The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Research*, *73*(3), 1180–1189. https://doi.org/10.1158/0008-5472.CAN-12-2850

Harries, L. W. (2012). Long non-coding RNAs and human disease. *Biochemical Society Transactions*, *40*(4), 902–906. https://doi.org/10.1042/BST20120020

Hartmann, C. (2006). A Wnt canon orchestrating osteoblastogenesis. *Trends in Cell Biology*, *16*(3), 151–158. https://doi.org/10.1016/j.tcb.2006.01.001

Hayter, S. M., & Cook, M. C. (2012). Updated assessment of the prevalence, spectrum and case definition of autoimmune disease. *Autoimmunity Reviews*, *11*(10), 754–765. https://doi.org/10.1016/j.autrev.2012.02.001

He, J.-H., Han, Z.-P., Liu, J.-M., Zhou, J.-B., Zou, M.-X., Lv, Y.-B., … Cao, M.-R. (2017). Overexpression of long non-coding rna MEG3 inhibits proliferation of

hepatocellular carcinoma huh7 cells via negative modulation of miRNA-664. *Journal of Cellular Biochemistry*, *118*(11), 3713–3721. https://doi.org/10.1002/jcb.26018

He, X., Tan, X., Wang, X., Jin, H., Liu, L., Ma, L., … Fan, Z. (2014). C-Myc-activated long noncoding RNA CCAT1 promotes colon cancer cell proliferation and invasion. *Tumor Biology*, *35*(12), 12181–12188. https://doi.org/10.1007/s13277-014-2526-4

Hou, B., Hou, X., & Ni, H. (2018). Long non-coding RNA LNC01133 promotes the tumorigenesis of ovarian cancer by sponging miR-126. *International Journal of Clinical and Experimental Pathology*, *11*(12), 5809–5819.

Huang, C.-S., Chu, J., Zhu, X.-X., Li, J.-H., Huang, X.-T., Cai, J.-P., … Yin, X.-Y. (2018). The C/EBPβ-*LINC01133* axis promotes cell proliferation in pancreatic ductal adenocarcinoma through upregulation of CCNG1. *Cancer Letters*, *421*, 63–72. https://doi.org/10.1016/j.canlet.2018.02.020

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., … Rinn, J. L. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*. https://doi.org/10.1016/j.cell.2010.06.040

International Human Genome Sequencing Consortium. (2004). International human genome sequencing consortium. Finishing the euchromatic sequence of the human genome. *Nature*, *431*, 931–945. https://doi.org/nature03001 [pii]\r10.1038/nature03001

Janoueix-Lerosey, I., Penther, D., Thioux, M., De Crémoux, P., Derré, J., Ambros, P., … Delattre, O. (2000). Molecular analysis of chromosome arm 17q gain in neuroblastoma. *Genes Chromosomes and Cancer*, *28*(3), 276–284. https://doi.org/10.1002/1098-2264(200007)28:3<276::AID-GCC5>3.0.CO;2-P

Jones, P. A., & Martienssen, R. (2005). A blueprint for a human epigenome project: The AACR human epigenome workshop. *Cancer Research*, *65*(24), 11241–11246. https://doi.org/10.1158/0008-5472.CAN-05-3865

Kamijo, T., & Nakagawara, A. (2012). Molecular and genetic bases of neuroblastoma. *International Journal of Clinical Oncology*, *17*(3), 190–195. https://doi.org/10.1007/s10147-012-0415-7

Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., …

Gingeras, T. R. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*, *316*(5830), 1484–1488. https://doi.org/10.1126/science.1138341

Kim, H. J., Lee, D. W., Yim, G. W., Nam, E. J., Kim, S., Kim, S. W., & Kim, Y. T. (2015). Long non-coding RNA HOTAIR is associated with human cervical cancer progression. *International Journal of Oncology*, *46*(2), 521–530. https://doi.org/10.3892/ijo.2014.2758

Knudson, A. G. (1971). Mutation and cancer: Statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences*, *68*(4), 820–823. https://doi.org/10.1073/pnas.68.4.820

Kong, J., Sun, W., Li, C., Wan, L., Wang, S., Wu, Y., … Lai, M. (2016). Long non-coding RNA *LINC01133* inhibits epithelial–mesenchymal transition and metastasis in colorectal cancer by interacting with SRSF6. *Cancer Letters*, *380*(2), 476–484. https://doi.org/10.1016/j.canlet.2016.07.015

Kong, J., Sun, W., Zhu, W., Liu, C., Zhang, H., & Wang, H. (2018). Long noncoding RNA *LINC01133* inhibits oral squamous cell carcinoma metastasis through a feedback regulation loop with GDF15. *Journal of Surgical Oncology*, *118*(8), 1326–1334. https://doi.org/10.1002/jso.25278

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, *128*(4), 693–705. https://doi.org/10.1016/j.cell.2007.02.005

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*, 860–921. https://doi.org/10.1038/35057062

Lee, J. T. (2012). Epigenetic regulation by long noncoding RNAs. *Science*, *338*(6113), 1435–1439. https://doi.org/10.1126/science.1231776

Lei, L., Xia, S., Liu, D., Li, X., Feng, J., Zhu, Y., … He, C. (2018). Genome-wide characterization of lncRNAs in acute myeloid leukemia. *Briefings in Bioinformatics*. https://doi.org/10.1093/bib/bbx007

Li, S., Mei, Z., Hu, H.-B., & Zhang, X. (2018). The lncRNA MALAT1 contributes to non-small cell lung cancer development via modulating miR-124/STAT3 axis. *Journal of Cellular Physiology*, *233*(9), 6679–6688. https://doi.org/10.1002/jcp.26325

Li, Z.-B., Chu, H.-T., Jia, M., & Li, L. (2020). Long noncoding RNA *LINC01139* promotes the progression of hepatocellular carcinoma by upregulating MYBL2 via competitively binding to miR-30 family. *Biochemical and Biophysical Research Communications*. https://doi.org/10.1016/j.bbrc.2020.02.116

Li, Z., Yu, X., & Shen, J. (2016). Long non-coding RNAs: emerging players in osteosarcoma. *Tumor Biology*, *37*(3), 2811–2816. https://doi.org/10.1007/s13277-015-4749-4

Liao, J., & Dong, L. P. (2019). *LINC00261* suppresses growth and metastasis of non-small cell lung cancer via repressing epithelial-mesenchymal transition. *European Review for Medical and Pharmacological Sciences*, *23*(9), 3829–3837. https://doi.org/10.26355/eurrev_201905_17810

Little, C. D., Nau, M. M., Carney, D. N., Gazdar, A. F., & Minna, J. D. (1983). Amplification and expression of the c-myc oncogene in human lung cancer cell lines. *Nature*, *306*(5939), 194–196. https://doi.org/10.1038/306194a0

Liu, G.-M., Zeng, H.-D., Zhang, C.-Y., & Xu, J.-W. (2019). Key genes associated with diabetes mellitus and hepatocellular carcinoma. *Pathology - Research and Practice*, *215*(11), 152510. https://doi.org/10.1016/j.prp.2019.152510

Liu, M., Shen, C., & Wang, C. (2019). Long noncoding RNA *LINC01133* confers tumor-suppressive functions in ovarian cancer by regulating leucine-rich repeat kinase 2 as an miR-205 sponge. *The American Journal of Pathology*, *189*(11), 2323–2339. https://doi.org/10.1016/j.ajpath.2019.07.020

Liu, P. Y., Erriquez, D., Marshall, G. M., Tee, A. E., Polly, P., Wong, M., … Liu, T. (2014a). Effects of a novel long noncoding RNA, lncUSMycN, on N-Myc expression and neuroblastoma progression. *JNCI: Journal of the National Cancer Institute*, *106*(7), dju113. https://doi.org/10.1093/jnci/dju113

Liu, P. Y., Erriquez, D., Marshall, G. M., Tee, A. E., Polly, P., Wong, M., … Liu, T. (2014b). Erratum. *JNCI Journal of the National Cancer Institute*, *107*(1), dju359–dju359. https://doi.org/10.1093/jnci/dju359

Liu, Y., Xiao, N., & Xu, S. F. (2017). Decreased expression of long non-coding RNA *LINC00261* is a prognostic marker for patients with non-small cell lung cancer: A preliminary study. *European Review for Medical and Pharmacological Sciences*, *21*(24), 5691–5695. https://doi.org/10.26355/eurrev_201712_14014

Lund, A. H. (2004). Epigenetics and cancer. *Genes & Development*, *18*(19), 2315–2335. https://doi.org/10.1101/gad.1232504

Luo, M., Li, Z., Wang, W., Zeng, Y., Liu, Z., & Qiu, J. (2013). Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression. *Cancer Letters*, *333*(2), 213–221. https://doi.org/10.1016/j.canlet.2013.01.033

Lynch, H. T., Fusaro, R. M., & Lynch, J. (1995). Hereditary cancer in adults. *Cancer Detection and Prevention*, *19*(3), 219–233.

Macklin, M. T. (1932). Human tumours and their inheritance. *Canadian Medical Association Journal*, *27*(2), 182.

Marina, N. (2004). Biology and therapeutic advances for pediatric osteosarcoma. *The Oncologist*, *9*(4), 422–441. https://doi.org/10.1634/theoncologist.9-4-422

Maris, J. M., Kyemba, S. M., Rebbeck, T. R., White, P. S., Sulman, E. P., Jensen, S. J., … Brodeur, G. M. (1997). Molecular genetic analysis of familial neuroblastoma. *European Journal of Cancer*, *33*(12), 1923–1928. https://doi.org/10.1016/S0959-8049(97)00265-7

Maris, John M., Guo, C., White, P. S., Hogarty, M. D., Thompson, P. M., Stram, D. O., … Brodeur, G. M. (2001). Allelic deletion at chromosome bands 11q14-23 is common in neuroblastoma. *Medical and Pediatric Oncology: The Official Journal of SIOP—International Society of Pediatric Oncology (Societé Internationale d'Oncologie Pédiatrique)*, *36*(1), 24–27. https://doi.org/10.1002/1096-911x(20010101)36:1<24::aid-mpo1007>3.0.co;2-7

Maris, John M., Hogarty, M. D., Bagatell, R., & Cohn, S. L. (2007). Neuroblastoma. *The Lancet*, *369*(9579), 2106–2120. https://doi.org/10.1016/S0140-6736(07)60983-0

Mas-Ponte, D., Carlevaro-Fita, J., Palumbo, E., Hermoso Pulido, T., Guigo, R., & Johnson, R. (2017). LncATLAS database for subcellular localization of long noncoding RNAs. *RNA*, *23*(7), 1080–1087. https://doi.org/10.1261/rna.060814.117

Matjasic, A., Popovic, M., Matos, B., & Glavac, D. (2017). Expression of LOC285758, a potential long non-coding biomarker, is methylation-dependent and correlates with glioma malignancy grade. *Radiology and Oncology*. https://doi.org/10.1515/raon-2017-0004

Mattick, J. S. (2001). Non-coding RNAs: The architects of eukaryotic complexity. *EMBO Reports, 2*(11), 986–991. https://doi.org/10.1093/embo-reports/kve230

McIntyre, J. F., Smith-Sorensen, B., Friend, S. H., Kassell, J., Borresen, A. L., Yan, Y. X., … Miser, J. (1994). Germline mutations of the p53 tumor suppressor gene in children with osteosarcoma. *Journal of Clinical Oncology, 12*(5), 925–930. https://doi.org/10.1200/JCO.1994.12.5.925

Monclair, T., Brodeur, G. M., Ambros, P. F., Brisse, H. J., Cecchetto, G., Holmes, K., … Pearson, A. D. J. (2009). The international neuroblastoma risk group (INRG) staging system: An INRG task force report. *Journal of Clinical Oncology, 27*(2), 298–303. https://doi.org/10.1200/JCO.2008.16.6876

Monnier, P., Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S., & Dandolo, L. (2013). H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proceedings of the National Academy of Sciences, 110*(51), 20693–20698. https://doi.org/10.1073/pnas.1310201110

Monsalve, J., Kapur, J., Malkin, D., & Babyn, P. S. (2011). Imaging of cancer predisposition syndromes in children. *RadioGraphics, 31*(1), 263–280. https://doi.org/10.1148/rg.311105099

Mossé, Y. P., Laudenslager, M., Longo, L., Cole, K. A., Wood, A., Attiyeh, E. F., … Maris, J. M. (2008). Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature, 455*(7215), 930–935. https://doi.org/10.1038/nature07261

Mueller, S., & Matthay, K. K. (2009). Neuroblastoma: Biology and staging. *Current Oncology Reports, 11*(6), 431–438. https://doi.org/10.1007/s11912-009-0059-6

Nagano, T., Mitchell, J. A., Sanz, L. A., Pauler, F. M., Ferguson-Smith, A. C., Feil, R., & Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science, 322*(5908), 1717–1720. https://doi.org/10.1126/science.1163802

Narod, S. A., Lenoir, G. M., & Stiller, C. (1991). An estimate of the heritable fraction of childhood cancer. *British Journal of Cancer, 63*(6), 993. https://doi.org/10.1038/bjc.1991.216

O'Brien, E. M., Selfe, J. L., Martins, A. S., Walters, Z. S., & Shipley, J. M. (2018). The

long non-coding RNA MYCNOS-01 regulates MYCN protein levels and affects growth of MYCN-amplified rhabdomyosarcoma and neuroblastoma cells. *BMC Cancer*, *18*(1), 217. https://doi.org/10.1186/s12885-018-4129-8

Paci, P., Colombo, T., & Farina, L. (2014). Computational analysis identifies a sponge interaction network between long non-coding RNAs and messenger RNAs in human breast cancer. *BMC Systems Biology*, *8*(1), 83. https://doi.org/10.1186/1752-0509-8-83

Pandey, G. K., Mitra, S., Subhash, S., Hertwig, F., Kanduri, M., Mishra, K., … Kanduri, C. (2014). The risk-associated long noncoding RNA NBAT-1 Controls neuroblastoma progression by regulating cell proliferation and neuronal differentiation. *Cancer Cell*, *26*(5), 722–737. https://doi.org/10.1016/j.ccell.2014.09.014

Pandey, R. R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., … Kanduri, C. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Molecular Cell*, *32*(2), 232–246. https://doi.org/10.1016/j.molcel.2008.08.022

Ponting, C. P., Oliver, P. L., & Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell*, *136*(4), 629–641. https://doi.org/10.1016/j.cell.2009.02.006

Prajapati, B., Fatma, M., Fatima, M., Khan, M. T., Sinha, S., & Seth, P. K. (2019). Identification of lncRNAs associated with neuroblastoma in cross-sectional databases: Potential biomarkers. *Frontiers in Molecular Neuroscience*, *12*. https://doi.org/10.3389/fnmol.2019.00293

Qiu, J.-J., Lin, Y.-Y., Ding, J.-X., Feng, W.-W., Jin, H.-Y., & Hua, K.-Q. (2015). Long non-coding RNA ANRIL predicts poor prognosis and promotes invasion/metastasis in serous ovarian cancer. *International Journal of Oncology*, *46*(6), 2497–2505. https://doi.org/10.3892/ijo.2015.2943

Rinn, J. L., & Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry*, *81*(1), 145–166. https://doi.org/10.1146/annurev-biochem-051410-092902

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., … Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, *129*(7), 1311–1323. https://doi.org/10.1016/j.cell.2007.05.022

Ritter, J., & Bielack, S. S. (2010). Osteosarcoma. *Annals of Oncology*, *21*(Supplement 7), vii320–vii325. https://doi.org/10.1093/annonc/mdq276

Romeo, S., Duim, R. A. J., Bridge, J. A., Mertens, F., de Jong, D., Dal Cin, P., … Hogendoorn, P. C. W. (2010). Heterogeneous and complex rearrangements of chromosome arm 6q in chondromyxoid fibroma. *The American Journal of Pathology*, *177*(3), 1365–1376. https://doi.org/10.2353/ajpath.2010.091277

Ron, E. (2002). Ionizing radiation and cancer risk: Evidence from epidemiology. *Pediatric Radiology*, *32*(4), 232–237. https://doi.org/10.1007/s00247-002-0672-0

Ron, E., Modan, B., Boice, J. D., Alfandary, E., Stovall, M., Chetrit, A., & Katz, L. (1988). Tumors of the brain and nervous system after radiotherapy in childhood. *New England Journal of Medicine*, *319*(16), 1033–1039. https://doi.org/10.1056/NEJM198810203191601

Saxena, A., & Carninci, P. (2011). Long non-coding RNA modifies chromatin. *BioEssays*, *33*(11), 830–839. https://doi.org/10.1002/bies.201100084

Sharma, S., Kelly, T. K., & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis*, *31*(1), 27–36. https://doi.org/10.1093/carcin/bgp220

Shi, J., Ma, H., Wang, H., Zhu, W., Jiang, S., Dou, R., & Yan, B. (2019). Overexpression of *LINC00261* inhibits non–small cell lung cancer cells progression by interacting with miR-522-3p and suppressing Wnt signaling. *Journal of Cellular Biochemistry*, *120*(10), 18378–18387. https://doi.org/10.1002/jcb.29149

Signal, B., Gloss, B. S., & Dinger, M. E. (2016). Computational approaches for functional prediction and characterisation of long noncoding RNAs. *Trends in Genetics*, *32*(10), 620–637. https://doi.org/10.1016/j.tig.2016.08.004

Song, Z., Zhang, X., Lin, Y., Wei, Y., Liang, S., & Dong, C. (2019). *LINC01133* inhibits breast cancer invasion and metastasis by negatively regulating SOX4 expression through EZH2. *Journal of Cellular and Molecular Medicine*, *23*(11), 7554–7565. https://doi.org/10.1111/jcmm.14625

Spector, L. G., Pankratz, N., & Marcotte, E. L. (2015). Genetic and nongenetic risk factors for childhood cancer. *Pediatric Clinics of North America*, *62*(1), 11–25. https://doi.org/10.1016/j.pcl.2014.09.013

Spitz, R., Hero, B., Ernestus, K., & Berthold, F. (2003). Deletions in chromosome arms 3p and 11q are new prognostic markers in localized and 4s neuroblastoma. *Clinical Cancer Research*, *9*(1), 52–58.

St. Laurent, G., Wahlestedt, C., & Kapranov, P. (2015). The Landscape of long noncoding RNA classification. *Trends in Genetics*, *31*(5), 239–251. https://doi.org/10.1016/j.tig.2015.03.007

Steliarova-Foucher, E., Stiller, C., Lacour, B., & Kaatsch, P. (2005). International classification of childhood cancer, third edition. *Cancer*, *103*(7), 1457–1467. https://doi.org/10.1002/cncr.20910

Strahm, B., & Malkin, D. (2006). Hereditary cancer predisposition in children: Genetic basis and clinical implications. *International Journal of Cancer*, *119*(9), 2001–2006. https://doi.org/10.1002/ijc.21962

Suenaga, Y., Islam, S. M. R., Alagu, J., Kaneko, Y., Kato, M., Tanaka, Y., … Nakagawara, A. (2014). NCYM, a cis-antisense gene of MYCN, encodes a de novo evolved protein that inhibits GSK3β resulting in the stabilization of MYCN in human neuroblastomas. *PLoS Genetics*, *10*(1), e1003996. https://doi.org/10.1371/journal.pgen.1003996

Sun, H., Wang, G., Peng, Y., Zeng, Y., Zhu, Q.-N., Li, T.-L., … Zhu, Y.-S. (2015). H19 lncRNA mediates 17β-estradiol-induced cell proliferation in MCF-7 breast cancer cells. *Oncology Reports*, *33*(6), 3045–3052. https://doi.org/10.3892/or.2015.3899

Sun, T., Warrington, N. M., Luo, J., Brooks, M. D., Dahiya, S., Snyder, S. C., … Rubin, J. B. (2014). Sexually dimorphic RB inactivation underlies mesenchymal glioblastoma prevalence in males. *Journal of Clinical Investigation*. https://doi.org/10.1172/JCI71048

Sun, Y., & Zhang, H. (2005). A unified mode of epigenetic gene silencing: RNA meets polycomb group proteins. *RNA Biology*, *2*(1), 8–10. https://doi.org/10.4161/rna.2.1.1465

Szulzewsky, F., Arora, S., de Witte, L., Ulas, T., Markovic, D., Schultze, J. L., … Kettenmann, H. (2016). Human glioblastoma-associated microglia/monocytes express a distinct RNA profile compared to human control and murine samples. *Glia*, *64*(8), 1416–1436. https://doi.org/10.1002/glia.23014

Szymanski, M., Barciszewska, M. Z., Erdmann, V. A., & Barciszewski, J. (2005). A new frontier for molecular medicine: Noncoding RNAs. *Biochimica et Biophysica Acta - Reviews on Cancer*, *1756*(1), 65–75. https://doi.org/10.1016/j.bbcan.2005.07.005

Taft, R. J., Pang, K. C., Mercer, T. R., Dinger, M., & Mattick, J. S. (2010). Non-coding RNAs: Regulators of disease. *The Journal of Pathology*, *220*(2), 126–139. https://doi.org/10.1002/path.2638

Takahashi, K., Yan, I., Haga, H., & Patel, T. (2014). Long noncoding RNA in liver diseases. *Hepatology*, *60*(2), 744–753. https://doi.org/10.1002/hep.27043

Tang, N., Song, W.-X., Luo, J., Haydon, R. C., & He, T.-C. (2008). Osteosarcoma development and stem cell differentiation. *Clinical Orthopaedics and Related Research*, *466*(9), 2114–2130. https://doi.org/10.1007/s11999-008-0335-z

Tang, Y., & Ji, F. (2019). lncRNA HOTTIP facilitates osteosarcoma cell migration, invasion and epithelial-mesenchymal transition by forming a positive feedback loop with c-Myc. *Oncology Letters*, *18*(2), 1649–1656. https://doi.org/10.3892/ol.2019.10463

Tsai, M.-C. M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., … Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, *329*(5992), 689–693. https://doi.org/10.1126/science.1192002

Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine*, *10*(8), 789. https://doi.org/10.1038/nm1087

Waddington, C. H. (2011). The epigenotype. *Endeavour*, *1*(1), 10–13. https://doi.org/10.1093/ije/dyr184

Wang, Q. (2016). Cancer predisposition genes: Molecular mechanisms and clinical impact on personalized cancer care: examples of Lynch and HBOC syndromes. *Acta Pharmacologica Sinica*, *37*(2), 143–149. https://doi.org/10.1038/aps.2015.89

Wang, S.-H., Zhang, W.-J., Wu, X.-C., Zhang, M.-D., Weng, M.-Z., Zhou, D., … Quan, Z.-W. (2016). Long non-coding RNA MALAT1 promotes gallbladder cancer development by acting as a molecular sponge to regulate miR-206. *Oncotarget*, *7*(25), 37857–37867. https://doi.org/10.18632/oncotarget.9347

Warthin, A. S. (1913). Heredity with reference to carcinoma. *Archives of Internal Medicine*, *XII*(5), 546. https://doi.org/10.1001/archinte.1913.00070050063006

Weiss, W. A. (1997). Targeted expression of MYCN causes neuroblastoma in transgenic mice. *The EMBO Journal*, *16*(11), 2985–2995. https://doi.org/10.1093/emboj/16.11.2985

Wen, Y., Gong, X., Dong, Y., & Tang, C. (2020). Long non coding RNA SNHG16 facilitates proliferation, migration, invasion and autophagy of neuroblastoma cells via sponging miR-542-3p and upregulating ATG5 expression. *OncoTargets and Therapy*, *13*, 263–275. https://doi.org/10.2147/OTT.S226915

White, P. S., Maris, J. M., Beltinger, C., Sulman, E., Marshall, H. N., Fujimori, M., … Hilliard, C. (1995). A region of consistent deletion in neuroblastoma maps within human chromosome 1p36.2-36.3. *Proceedings of the National Academy of Sciences*, *92*(12), 5520–5524. https://doi.org/10.1073/pnas.92.12.5520

Wilusz, J. E., Sunwoo, H., & Spector, D. L. (2009). Long noncoding RNAs: Functional surprises from the RNA world. *Genes & Development*, *23*(13), 1494–1504. https://doi.org/10.1101/gad.1800909

Wongtrakoongate, P. (2015). Epigenetic therapy of cancer stem and progenitor cells by targeting DNA methylation machineries. *World Journal of Stem Cells*, *7*(1), 137. https://doi.org/10.4252/wjsc.v7.i1.137

Wu, A., Li, J., Wu, K., Mo, Y., Luo, Y., Ye, H., … Yang, Z. (2016). LATS2 as a poor prognostic marker regulates non-small cell lung cancer invasion by modulating MMPs expression. *Biomedicine & Pharmacotherapy*, *82*, 290–297. https://doi.org/10.1016/j.biopha.2016.04.008

Xie, J., Lin, D., Lee, D. H. T., Akunowicz, J., Hansen, M., Miller, C., … Koeffler, H. P. (2017). Copy number analysis identifies tumor suppressive lncRNAs in human osteosarcoma. *International Journal of Oncology*, *50*(3), 863–872. https://doi.org/10.3892/ijo.2017.3864

Xiong, Y., Wang, L., Li, Y., Chen, M., He, W., & Qi, L. (2017). The long non-coding RNA XIST interacted with miR-124 to modulate bladder cancer growth, invasion and migration by targeting androgen receptor (AR). *Cellular Physiology and Biochemistry*, *43*(1), 405–418. https://doi.org/10.1159/000480419

Yan, L., Wu, X., Yin, X., Du, F., Liu, Y., & Ding, X. (2018). LncRNA CCAT2 promoted

osteosarcoma cell proliferation and invasion. *Journal of Cellular and Molecular Medicine*, *22*(5), 2592–2599. https://doi.org/https://doi.org/10.1111/jcmm.13518

Yan, D., Liu, W., Liu, Y., & Luo, M. (2019). *LINC00261* suppresses human colon cancer progression via sponging miR-324-3p and inactivating the Wnt/β-catenin pathway. *Journal of Cellular Physiology*, *234*(12), 22648–22656. https://doi.org/10.1002/jcp.28831

Yan, L., Wu, X., Liu, Y., & Xian, W. (2019). LncRNA LINC00511 promotes osteosarcoma cell proliferation and migration through sponging miR-765. *Journal of Cellular Biochemistry*, *120*(5), 7248–7256. https://doi.org/10.1002/jcb.27999

Yan, R., Wang, K., Peng, R., Wang, S., Cao, J., Wang, P., & Song, C. (2016). Genetic variants in lncRNA SRA and risk of breast cancer. *Oncotarget*, *7*(16), 22486. https://doi.org/10.18632/oncotarget.7995

Yang, F., Xue, X., Bi, J., Zheng, L., Zhi, K., Gu, Y., & Fang, G. (2013). Long noncoding RNA CCAT1, which could be activated by c-Myc, promotes the progression of gastric carcinoma. *Journal of Cancer Research and Clinical Oncology*, *139*(3), 437–445. https://doi.org/10.1007/s00432-012-1324-x

Yang, X.-Z., Cheng, T.-T., He, Q.-J., Lei, Z.-Y., Chi, J., Tang, Z., … Cui, S.-Z. (2018). *LINC01133* as ceRNA inhibits gastric cancer progression by sponging miR-106a-3p to regulate APC expression and the Wnt/β-catenin pathway. *Molecular Cancer*, *17*(1), 126. https://doi.org/10.1186/s12943-018-0874-1

Yao, Y., Ma, J., Xue, Y., Wang, P., Li, Z., Liu, J., … Liu, Y. (2015). Knockdown of long non-coding RNA XIST exerts tumor-suppressive functions in human glioblastoma stem cells by up-regulating miR-152. *Cancer Letters*, *359*(1), 75–86. https://doi.org/10.1016/j.canlet.2014.12.051

Yelin, R., Dahary, D., Sorek, R., Levanon, E. Y., Goldstein, O., Shoshan, A., … Rotman, G. (2003). Widespread occurrence of antisense transcription in the human genome. *Nature Biotechnology*, *21*(4), 379. https://doi.org/10.1038/nbt808

Yu, Y., Li, L., Zheng, Z., Chen, S., Chen, E., & Hu, Y. (2017). Long non-coding RNA *LINC00261* suppresses gastric cancer progression via promoting Slug degradation. *Journal of Cellular and Molecular Medicine*, *21*(5), 955–967. https://doi.org/10.1111/jcmm.13035

Zang, C., Nie, F., Wang, Q., Sun, M., Li, W., He, J., … Lu, K. (2016). Long non-coding

RNA *LINC01133* represses KLF2, P21 and E-cadherin transcription through binding with EZH2, LSD1 in non small cell lung cancer. *Oncotarget*, *7*(10), 11696. https://doi.org/10.18632/oncotarget.7077

Zeng, H.-F., Qiu, H.-Y., & Feng, F.-B. (2018). Long noncoding RNA *LINC01133* functions as an miR-422a sponge to aggravate the tumorigenesis of human osteosarcoma. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*, *26*(3), 335–343. https://doi.org/10.3727/096504017X14907375885605

Zhang, C., Xie, L., Liang, H., & Cui, Y. (2019). LncRNA MIAT facilitates osteosarcoma progression by regulating miR-128-3p/VEGFC axis. *IUBMB Life*, *71*(7), 845–853. https://doi.org/10.1002/iub.2001

Zhang, H.-F., Li, W., & Han, Y.-D. (2018). *LINC00261* suppresses cell proliferation, invasion and Notch signaling pathway in hepatocellular carcinoma. *Cancer Biomarkers*, *21*(3), 575–582. https://doi.org/10.3233/CBM-170471

Zhang, J. H., Li, A. Y., & Wei, N. (2017). Downregulation of long non-coding RNA *LINC01133* is predictive of poor prognosis in colorectal cancer patients. *European Review for Medical and Pharmacological Sciences*, *21*(9), 2103–2107.

Zhang, Jiao, Li, W.-Y., Yang, Y., Yan, L.-Z., Zhang, S.-Y., He, J., & Wang, J.-X. (2019). LncRNA XIST facilitates cell growth, migration and invasion via modulating H3 histone methylation of DKK1 in neuroblastoma. *Cell Cycle*, *18*(16), 1882–1892. https://doi.org/10.1080/15384101.2019.1632134

Zhang, Jiao, Zhou, B., Liu, Y., Chen, K., Bao, P., Wang, Y., … Li, Y. (2014). Wnt inhibitory factor-1 functions as a tumor suppressor through modulating Wnt/β-catenin signaling in neuroblastoma. *Cancer Letters*, *348*(1–2), 12–19. https://doi.org/10.1016/j.canlet.2014.02.011

Zhang, Jing, Zhu, N., & Chen, X. (2015). A novel long noncoding RNA *LINC01133* is upregulated in lung squamous cell cancer and predicts survival. *Tumor Biology*, *36*(10), 7465–7471. https://doi.org/10.1007/s13277-015-3460-9

Zhang, Jinghui, Walsh, M. F., Wu, G., Edmonson, M. N., Gruber, T. A., Easton, J., … Downing, J. R. (2015). Germline mutations in predisposition genes in pediatric cancer. *New England Journal of Medicine*, *373*(24), 2336–2346. https://doi.org/10.1056/NEJMoa1508054

Zhang, R., Xia, L. Q., Lu, W. W., Zhang, J., & Zhu, J.-S. (2016). LncRNAs and cancer. *Oncology Letters*, *12*(2), 1233–1239. https://doi.org/10.3892/ol.2016.4770

Zheng, Y., Zhang, X., & Bu, Y. (2019). *LINC01133* aggravates the progression of hepatocellular carcinoma by activating the PI3K/AKT pathway. *Journal of Cellular Biochemistry*, *120*(3), 4172–4179. https://doi.org/10.1002/jcb.27704

Zhou, Q., Chen, F., Zhao, J., Li, B., Liang, Y., Pan, W., … Zheng, D. (2016). Long non-coding RNA PVT1 promotes osteosarcoma development by acting as a molecular sponge to regulate miR-195. *Oncotarget*, *7*(50), 82620–82633. https://doi.org/10.18632/oncotarget.13012