



LUND UNIVERSITY

Genomic and Transcriptomic Analyses of Osteogenic Tumours of Bone

Saba, Karim

2020

[Link to publication](#)

Citation for published version (APA):

Saba, K. (2020). *Genomic and Transcriptomic Analyses of Osteogenic Tumours of Bone*. [Doctoral Thesis (compilation), Department of Laboratory Medicine]. Lund University, Faculty of Medicine.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

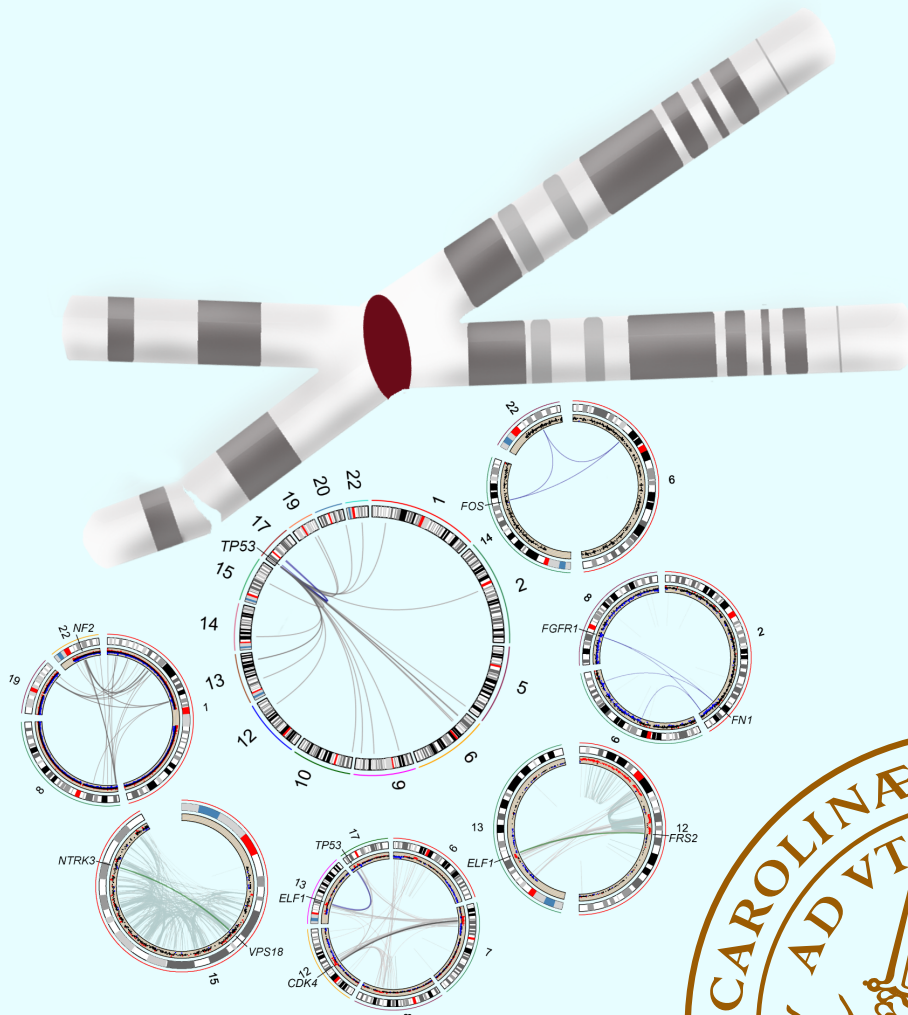
LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Genomic and Transcriptomic Analyses of Osteogenic Tumours of Bone

KARIM H. SABA

FACULTY OF MEDICINE | LUND UNIVERSITY



It all began a few years ago, in a nation not so far away. Looking back onto our first impressions, it is amazing to see how we have developed. After our initial misconceptions, we became neighbours, flatmates, colleagues, and foremost friends. We have been through good times, bad times, times we wish we could forget and times we will never forget. Karim's dedication to his work and ability to simultaneously stay committed to family and friends is impressive and a testament to his humble and loyal character. We are very proud of you and cannot wait to watch you strive for even bigger goals. Whichever passion you choose to pursue next, we wish you the best of luck and look forward to accompanying you along the way. Whatever the future may bring, you are destined for greatness!

- You lost the game, love Jimmie
- Best wishes from your dedicated friend and fellow beer aficionado, Oskar



**FACULTY OF
MEDICINE**

Department of Laboratory Medicine, Lund
Division of Clinical Genetics

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2020:119
ISBN 978-91-7619-982-4
ISSN 1652-8220



Genomic and Transcriptomic Analyses of Osteogenic Tumours of Bone

Karim H. Saba



LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Belfragesalen, BMC D15, Lund

Friday 20 November 2020, at 09:00.

Faculty opponent

Professor Adrienne M. Flanagan

University College London Cancer Institute, London, United Kingdom
Royal National Orthopaedic Hospital NHS Trust, Stanmore, United Kingdom

Organization LUND UNIVERSITY Faculty of Medicine Department of Laboratory Medicine Division of Clinical Genetics Author: Karim H. Saba		Document name Doctoral Dissertation	
		Date of issue 20 November 2020	
		Sponsoring organization	
Genomic and Transcriptomic Analyses of Osteogenic Tumours of Bone			
Abstract			
<p>Primary tumours of bone are heterogenous and infrequent neoplasms. Distinguishing between benign, intermediate and malignant entities can in some instances pose a clinical challenge. For some tumour types, there is still much to be learned about the genetic mechanisms that give rise to and drive these tumours forward. With the hope of improving diagnostic accuracy and treatment outcomes on the longer term, this thesis will deal with the genetic mutational mechanisms that characterise the primary osteogenic neoplasms osteoblastoma and osteosarcoma.</p> <p>In Article I, we show that a subgroup of non-<i>FOS</i>-rearranged, preferentially epithelioid osteoblastomas harbour homozygous loss of the <i>NF2</i> gene. Additionally, we find a lower proportion of <i>FOS</i>-rearranged cases than previously reported and a high normal cell content.</p> <p>In Article II, we genetically characterise a rare chondroblastoma-like osteosarcoma/malignant phosphaturic mesenchymal tumour of bone. We detect a potentially targetable <i>FN1-FGFR1</i> gene fusion and homozygous loss of the <i>CDKN2A</i> and <i>DMD</i> genes.</p> <p>In Article III, an RNA-sequencing screen of conventional osteosarcomas reveals that <i>NTRK</i> fusions are rare and most likely non-functional events.</p> <p>In Article IV, we demonstrate, for the first time, the existence of a recurrent gain-of-function mechanism involving the promoter region of the <i>TP53</i> tumour suppressor gene in a subset of conventional osteosarcoma. We show that structural variants abrogate <i>TP53</i> expression but also relocate its promoter region. By responding to ongoing DNA damage, it in turn leads to upregulation of known or putative oncogenes erroneously translocated into its vicinity.</p> <p>In Article V, we subdivide 12q-amplified osteosarcomas into four distinct groups and show that recurrent promoter swapping events involving the <i>FRS2</i> and <i>PLEKHA5</i> regulatory regions occur in many high-grade and dedifferentiated osteosarcomas with <i>CDK4</i> and <i>MDM2</i> amplification.</p> <p>In conclusion, this thesis will highlight the role chromosome remodelling plays in the development of primary osteogenic tumours of bone.</p>			
Key words: bone sarcoma, promoter swapping, gene fusions, osteoblastoma, osteosarcoma, <i>TP53</i> , <i>FRS2</i> , <i>PLEKHA5</i> , <i>NF2</i> , <i>FN1-FGFR1</i> , <i>NTRK</i>			
Classification system and/or index terms (if any)			
Supplementary bibliographical information		Language: English	
ISSN and key title: 1652-8220		ISBN: 978-91-7619-982-4	
Recipient's notes	Number of pages 67		Price
	Security classification		

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature 

Date 2020-10-15

Genomic and Transcriptomic Analyses of Osteogenic Tumours of Bone

Karim H. Saba



LUND
UNIVERSITY

Cover photo by Karim H. Saba

Copyright Karim H. Saba

Lund University
Faculty of Medicine
Department of Laboratory Medicine
Division of Clinical Genetics

ISBN 978-91-7619-982-4

ISSN 1652-8220

Lund University, Faculty of Medicine Doctoral Dissertation Series 2020:119

Printed in Sweden by Media-Tryck, Lund University
Lund 2020



Media-Tryck is a Nordic Swan Ecolabel
certified provider of printed material.
Read more about our environmental
work at www.mediatryck.lu.se

MADE IN SWEDEN 

Table of Contents

Original Articles.....	7
Abbreviations	9
Introduction.....	11
Neoplasia and the cell cycle	11
Bone tumours.....	14
Genetic mutational mechanisms in bone tumours	16
Single nucleotide variants.....	16
Copy number variants	17
Structural variants.....	19
Gene fusions and promoter swapping	22
Current clinical problems	25
Osteoblastoma	25
Conventional osteosarcoma	26
Low-grade and dedifferentiated osteosarcoma	27
Novel investigations into primary osteogenic bone tumours	29
Aim	29
Materials and Methods	30
Patients and tumour samples	30
DNA and RNA extraction	30
Polymerase chain reaction	31
Sanger sequencing.....	31
Fluorescence <i>in situ</i> hybridisation.....	31
SNP array	32
Next generation sequencing	32
Bioinformatics	35
<i>In vitro</i> cell models.....	36
Results.....	38
Article I	38

Article II	38
Article III.....	39
Article IV.....	39
Article V	40
General Discussion.....	42
<i>FOS</i> stabilisation versus <i>NF2</i> loss in osteoblastoma: two genetic subgroups of the same disease	42
Bone sarcomas are rare, their subtypes even more so	43
Targeting genomic alterations as part of future personalised medicine practices: any considerations?	44
Hijacked regulatory elements are a central theme in osteosarcoma	44
Conclusions.....	47
Popular science summary.....	49
Populärvetenskaplig sammanfattning	51
Acknowledgements	53
References	57

Original Articles

- I. Saba KH, Cornmark L, Hofvander J, Magnusson L, Nilsson J, van den Bos, H, Spierings DCJ, Foijer F, Staaf J, Brosjö O, Sumathi VP, Lam SW, Szuhai K, Bovée JVMG, Kovac M, Baumhoer D, Styring E, Nord KH. **Loss of *NF2* defines a genetic subgroup of non-*FOS*-rearranged osteoblastoma.** *J Pathol Clin Res.* 2020; doi: 10.1002/cjp2.172. Online ahead of print.
- II. Saba KH, Cornmark L, Rissler M, Fioretos T, Åström K, Haglund F, Rosenberg AE, Brosjö O, Nord KH. **Genetic profiling of a chondroblastoma-like osteosarcoma/malignant phosphaturic mesenchymal tumor of bone reveals a homozygous deletion of *CDKN2A*, intragenic deletion of *DMD*, and a targetable *FNI-FGFR1* gene fusion.** *Genes Chromosomes Cancer.* 2019; 58: 731-736.
- III. Ameline B*, Saba KH*, Kovac M, Magnusson L, Witt O, Bielack S, Nathrath M, Nord KH, Baumhoer D. ***NTRK* fusions in osteosarcoma are rare and non-functional events.** *J Pathol Clin Res.* 2020; 6: 107-112. (*Shared first authors)
- IV. Saba KH, Cornmark L, Kovac M, Magnusson L, Nilsson J, van den Bos H, Spierings DCJ, Bidgoli M, Jonson T, Sumathi VP, Brosjö O, Staaf J, Foijer F, Styring E, Nathrath M, Baumhoer D†, Nord KH†. **Oncogenes hijack a constitutively active *TP53* promoter in osteosarcoma.** Manuscript. Preprint available at *bioRxiv* doi.org/10.1101/2020.04.20.050252 (†Jointly directed this work)
- V. Saba KH, Nilsson J, Magnusson L, Nathrath M, Brosjö O, Styring E, Haglund F, Baumhoer D, Nord KH. **Genetic rearrangements in osteosarcoma result in ectopic gene expression through transposition of regulatory elements and selective amplification.** Manuscript.

Article not discussed in this thesis:

Magnusson L, Hansen N, Saba KH, Nilsson J, Fioretos T, Rissler P, Nord KH. **Loss of the tumour suppressor gene *AIP* mediates the browning of human brown fat tumours.** *J Pathol.* 2017; 243: 160-164.

Abbreviations

CNV	Copy number variant
FISH	Fluorescence <i>in situ</i> hybridisation
NGS	Next generation sequencing
NTRK	Neurotrophic tyrosine receptor kinase
PCR	Polymerase chain reaction
RT-PCR	Reverse transcription polymerase chain reaction
RT-qPCR	Real-time quantitative polymerase chain reaction
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SV	Structural variant
TKI	Tyrosine kinase inhibitor
WES	Whole-exome sequencing
WGS	Whole-genome paired end sequencing

Introduction

Neoplasia and the cell cycle

A neoplasm is an aberrant and excessive growth of tissue. The word is derived from the ancient Greek words for new (néos) and formation (plasma). As early as 1914, Theodor Boveri postulated that neoplasia is caused by chromosomal rearrangements resulting from multipolar mitoses¹, during a time when the structure of the DNA double-helix was not yet known, let alone that the composition of chromosomes had yet to be determined to even be DNA^{2,3}. Throughout the 20th century, advancements in medical and molecular research have expanded our understanding of cancer to the point where we know which genes tumours tend to over-activate – so called oncogenes – and which genes tumours often deactivate or knockout – so called tumour suppressor genes. Advances in genome sequencing technologies in the early 21st century have again expanded our understanding of how tumours form, develop and spread leading to death of the patient in many instances.

The growth and replication rates of the cells that make up our tissues and organs are normally intricately controlled by a process called the cell cycle (Figure 1). In the first of its main steps, termed interphase, the cells grow and accumulate the necessary requirements for cell division during the first gap (G1) phase, then replicate their DNA during the synthesis (S) phase, which is then followed by a second gap (G2) phase. During G2, the cells continue to grow before entering the second major step, the mitotic (M) phase. During mitosis, replicated DNA is separated to the respective poles of the cells in a multi-step process, which culminates in the cell splitting in half and the creation of two identical daughter cells. During the transition from one phase to the other, there are elaborate control mechanisms in place to ensure that no errors have been made along the way before the cell commits to the next phase of its cell cycle. These control mechanisms are termed ‘cell cycle checkpoints’. If an error, or mutation, in the replicated DNA is discovered, the cell cycle is stopped and the damage is assessed. These errors can occur in the S phase, or even be influenced by external factors, such as mutagens, during S or other phases. During the M phase, chromosome missegregation may even occur. If the damage is deemed repairable by the cell, a variety of DNA repair mechanisms can be employed to right the error. If the damage is too

great, the cell may enter into apoptosis, or programmed cell death, to avoid carrying catastrophic mistakes into the next generation of daughter cells. Cells may even decide not to replicate their DNA in the first place and enter a quiescent mode where they cease to grow and divide, termed G0. The cell cycle is however not completely fool proof and mutations or breaks in the DNA may occur that go undetected by the cell cycle checkpoints. The general dogma states that if the cell acquires enough mutations over time in key genes, it may evade the different cell cycle checkpoints and ignore cues from its surrounding environment, thereby becoming able to grow and proliferate uncontrollably. It becomes neoplastic, capable of growing into a cancer, manipulating its surrounding microenvironment and maybe even metastasising^{4, 5}. Recent advances brought upon by modern, large-scale human and cancer genome sequencing studies have, however, shown that numerous mutations and rearrangements can occur at once due to a single catastrophic chromosomal event or an orchestrated series of rearrangements, as opposed to occurring sequentially over time^{6, 7}.

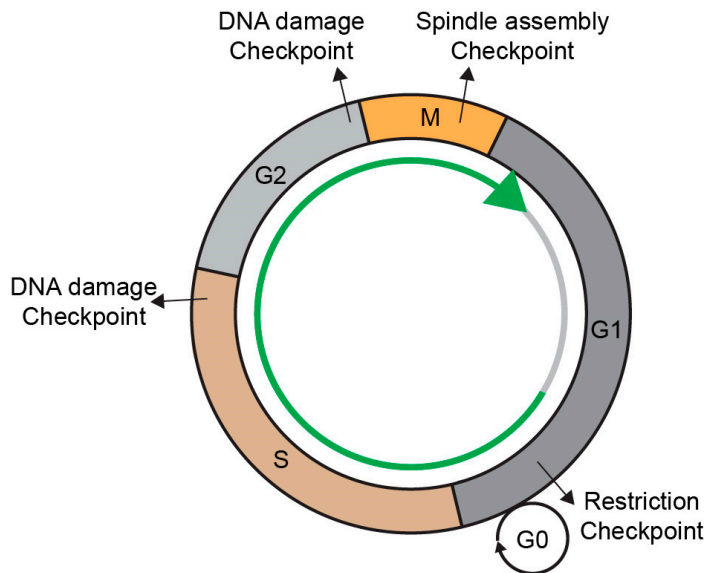


Figure 1. *The mammalian cell cycle.*

Cells grow and prepare for DNA replication during the G1 phase. At the restriction checkpoint, the cell decides whether to commit to the cell cycle or to pause its growth, become quiescent and enter G0. During the S phase, the DNA is replicated and checked for damage before entering G2. In G2, the cell continues growing and produces proteins required for mitosis. Once a last DNA damage check has been conducted, the cell enters mitosis. In the middle of mitosis, the spindle assembly checkpoint makes sure everything is in place before the sister chromosomes are pulled to their respective poles, after which the cell splits into two identical daughter cells.

The exponential growth of knowledge during the late 20th and early 21st centuries has led to the classification of various ‘hallmarks of cancer’, summarised and described by Douglas Hanahan and Robert Weinberg in two field-encompassing reviews^{8, 9}. This started with the identification of an initial six hallmarks, or characteristics, that cancer cells have acquired that allow them to continue growing and replicating endlessly. By being able to (i) sustain proliferative signalling, (ii) evade growth suppressors, (iii) invade and metastasise, (iv) achieve replicative immortality, (v) induce angiogenesis, or the formation of blood vessels, and (vi) resist programmed cell death, cancerous cells can thrive and may lead to death if left untreated. In the follow up review a decade later, new emerging hallmarks were identified, namely (vii) the deregulation of cellular energetics – to ensure a continued supply of nutrients – and (viii) the avoidance of immune destruction – to give cancer cells an extra advantage in evading the body’s own defence mechanisms. Taking into consideration a background of genetic instability that allows for new mutations to arise and be selected for and a growing understanding of the role the tumour microenvironment, which is made up of normal cells, plays in cancer progression, one can appreciate how ingenious tumours really are and how much of an uphill battle the fight against them is.

There are over 100 different types of cancers identified to date, and each can be considered individual diseases on their own. In general, cancers can be divided into three main categories: (i) carcinomas that arise in the epithelial tissues of the body, (ii) leukaemia that are cancers of the blood and (iii) sarcomas that are cancers of mesenchymal origin. While there are a plethora of benign or locally aggressive neoplasms, cancers are instead malignant entities that have metastatic potential. Just because a tumour is classified as benign or locally aggressive does not necessarily mean that it lacks potential to turn into a much more aggressive disease, however. This thesis will deal with the genomic (DNA-level) and transcriptomic (RNA-level) changes underlying the development of a variety of osteogenic, or bone-producing, primary skeletal tumours including a spectrum from locally aggressive osteblastoma through malignant low-grade osteosarcoma to high-grade conventional osteosarcoma. Primary skeletal or bone tumours are tumours that arise initially within a certain bone in the body and are not metastases of other types of cancers that settle and grow in a bone.

Bone tumours

Primary bone tumours are heterogeneous and very infrequent tumour types, and bone sarcomas, *i.e.*, malignant bone tumours, represent less than 0.2% of all diagnosed malignancies¹⁰. As of the latest World Health Organisation update (2020), there are more than 60 subtypes of recognised bone tumours which are classified on a 4-level malignancy scale¹¹. Tumours are classified as **benign** if they tend not to recur and are generally not destructive in the rare event that they do. They are almost always treatable by complete local excision (surgical removal) or curettage (tissue removal by scraping). **Intermediate** tumours can be either **locally aggressive** or **rarely metastasising**. Intermediate locally aggressive tumours have the potential to recur locally and can infiltrate surrounding tissue in a destructive manner. They generally do not have metastatic potential but require removal with wide surgical margins (removal of surrounding normal tissue). Local chemotherapy may also be used. Intermediate rarely metastasising tumours of bone are similar to their locally aggressive counterparts, with the exception that they have been documented to metastasise, albeit at low rates. The most common metastatic site is the lungs. Finally, bone tumours are classified as **malignant**, and are thereby bone sarcomas, if they exhibit locally destructive growth patterns, can recur and come with a profound risk of metastasising to distant parts of the body.

Bone tumours are most commonly either chondrogenic (cartilage-producing) or osteogenic (bone-forming) in nature, although other types of lesions exist such as fibrogenic, vascular and osteoclastic giant cell-rich, to name a few. Ewing sarcoma, while technically classified as an undifferentiated small round cell sarcoma¹¹, may arise in either bones or soft tissues, but most commonly does so in bones and will be used later on in this introduction for explanatory purposes. Primary bone tumours generally show unique age distributions, with incidence rates varying greatly based on patient age. For example, Ewing sarcoma and osteosarcoma are much more common in children, adolescents and younger adults, whereas chondrosarcoma is much more frequent in middle-aged and older patients¹². Interestingly, osteosarcoma shows a bimodal age distribution with a second, albeit smaller peak in older individuals. Most bone lesions are discovered due to persistent non-mechanical pain in a certain bone or due to the presence of swelling. The tumours investigated in this thesis, osteoblastoma and osteosarcoma, are osteogenic in nature.

The first, osteoblastoma, is an intermediate locally aggressive tumour of bone that mostly affects patients between 10-30 years old, although it can affect individuals of almost any age¹³. It is much more common in males, with a 2:1 male to female ratio. Around half of osteoblastoma cases arise in posterior elements of the spine and sacrum,

but the long bones of the leg (femur and tibia) are also common sites. It can be divided into two morphological subgroups, namely conventional osteblastoma and epithelioid osteblastoma. Osteblastoma is morphologically similar to but, per definition, larger than osteoid osteoma, a benign tumour of bone. Not much was known about the genetics of osteblastoma until recently^{14, 15}, described in more detail in the next sections and in **Article I**. Osteblastomas are usually surgically removed and have excellent prognosis, although cases with difficult surgical access may recur over time and disrupt nearby normal tissue.

The second, osteosarcoma, is a collection of malignant tumour subtypes of varying prognosis. Conventional osteosarcoma is a high-grade malignancy and has a rather dismal prognosis, with 5-year survival rates at 60-70% for localised disease treated with both chemotherapy and surgery, but much lower for metastatic or recurrent disease¹⁶⁻²⁰. Conventional osteosarcoma most commonly affects growing children and adolescents with the most frequent site being the distal femur and proximal tibia, near the active growth plates around the knee²¹. The proximal humerus is also a common site, but osteosarcoma can arise in any bone. Osteosarcoma in older patients can commonly present in the hip, pelvis or jaw bones, for example. Males are more likely to be affected than females. Although the introduction of multi-drug chemotherapy in the 1980s dramatically improved survival rates, very little has happened since. Surviving patients have lifelong scars from the radical surgical techniques and intensive chemotherapy regimens that need to be employed, with a negative impact on quality of life^{22, 23}. Conventional osteosarcoma had long been known to harbour extreme genetic complexity^{24, 25}, something highlighted again but in much further detail in recent genome sequencing efforts²⁶⁻³². Conventional osteosarcoma is discussed in **Articles III-V** and an exceedingly rare subtype of it, chondroblastoma-like osteosarcoma, is discussed in **Article II**.

Low-grade osteosarcoma is a malignant tumour of bone but has a good prognosis with a 90% 5-year survival rate^{33, 34}. It can be divided into parosteal osteosarcoma and low-grade central osteosarcoma depending on the tumour's relative location to the bone cortex. Parosteal osteosarcomas arise on the cortical surface of the bone, while low-grade central osteosarcomas arise in the medullary cavity of the bone. Low-grade osteosarcomas peak in the third decade of life and affect females slightly more than males. They make up about 5-6% of all diagnosed osteosarcomas and most commonly occur in the tibia and femur, although they may affect other parts of the skeleton. Low-grade osteosarcoma may dedifferentiate and thereby transition into a high-grade tumour with a worsened prognosis^{35, 36}. Low-grade and dedifferentiated osteosarcoma will be one of the subjects of **Article V**.

Genetic mutational mechanisms in bone tumours

Single nucleotide variants

The DNA molecules in our cells are made up of four nucleotide bases named adenine, thymine, cytosine and guanine (A, T, C and G). These bases are arranged in a way that provides cues to certain enzymes on where to start ‘transcribing’ DNA into RNA which can then be ‘translated’ by cellular machinery to proteins that carry out various functions necessary for survival. These nucleotides can also constitute regulatory regions that do not code for any proteins, but rather control the expression, or transcription level, of genes. Single nucleotide variants (SNVs) are thereby mutations that represent a substitution of one nucleotide to another, such as an A to T substitution (A>T), causing an error in the sequence once compared to the reference human genome. There are millions of naturally occurring variants that are known to be harmless and are part of the natural variation in the human genome. However, SNVs within **exons** – the coding parts of genes – or within key regulatory regions can lead to errors in protein production and gene dysregulation, respectively. A SNV affecting a coding sequence can be either **synonymous**, *i.e.*, it does not lead to a change in protein sequence due to the redundancy of the genetic code, or **non-synonymous**, leading to a change in the amino acid sequence of a protein (**missense mutation**). In some instances, an interruption of the reading-frame, via an erroneous ‘stop’ signal, can be caused by a SNV which results in a truncated protein (**nonsense mutation**). Some SNVs can occur in specific locations within **introns** – the sequences that separate a gene’s exons – resulting in incorrect intron splicing, or removal, which can have deleterious consequences on protein structure and function³⁷.

Some SNVs confer a gain-of-function to genes, thereby increasing or prolonging their intended activity. Such mutations normally affect a specific nucleotide in a gene as it is rather unlikely that a random mutation anywhere in a gene leads to overactivation and are therefore known as hotspot mutations. Characteristic hotspot mutations occur in various tumours of bone such as the G>T mutation leading to the G35W amino acid substitution in the protein encoded by the *H3-3A* gene (previously *H3F3A*) in giant cell tumour of bone³⁸. In chondroblastoma, an A>T substitution causes the K37M amino acid change in the protein encoded by the *H3-3B* gene (previously *H3F3B*), a phenomenon observed, although not as frequently, in the *H3-3A* gene as well in chondroblastomas³⁸. Both of these genes encode identical H3.3 histones that function in the spatial organisation of DNA. Any alteration in their activity can lead to dysregulation of the expression of numerous other genes with the consequence being the development of a tumour. More examples of such defining hotspot SNVs are given

in **Article II**. In about 4% of osteosarcomas, 23% of chondrosarcomas and in almost all cases of the soft tissue tumour myxoid liposarcoma, SNVs in the promoter, or one of the regulatory regions, of the *TERT* gene have been observed^{39, 40}. These lead to increased activity and thereby increased expression of the gene resulting in cellular immortality, one of the key hallmarks of cancer, and can also be found in bladder and thyroid cancers, melanoma and gliomas^{41, 42}. The majority of SNVs, however, represent loss-of-function mutations. These can occur in numerous places within a gene's coding sequences and often lead to improper protein structure or impaired polymerisation activity normally required for the protein to function correctly. Such acquired inactivating point mutations are known to affect the *TP53* and *RB1* tumour suppressor genes in many conventional osteosarcomas²⁶⁻²⁸.

Copy number variants

The normal human genome is diploid in nature, meaning that we have two copies of every sequence of DNA, a maternal one and a paternal one. The DNA is organised into structures called chromosomes and, normally, 22 pairs of autosomes and one pair of sex chromosomes exist in the nuclei of our cells. Chromosomes are divided into two arms, the short arm and the long arm, or the **p arm** and **q arm**, respectively. The two arms are separated by a **centromere**. Proximal sequences are those closer to the centromere and distal sequences are those farther out towards the **telomeres**, or chromosomal caps. One can therefore describe the relative position of a gene as being either distal or proximal to another one or to a breakpoint in the DNA, for example. Different chromosomal regions are numbered starting from 1 closest to the centromere and then outward along each arm towards the telomeres, such as p1, p2, q1, q2 and so on. This varies depending on how the regions on each chromosome appear as seen under the microscope using a specific stain. Regions are further divided into bands which are in turn also numbered starting from 1 outward towards the telomeres, such as p12 or q34. 4p15 is therefore region 1, band 5 of the p arm of chromosome 4 and 4q28 is region 2, band 8 of the q arm of chromosome 4 (Figure 2). Some bands can once again be divided into sub-bands. Five pairs of human chromosomes have very short p arms and are known as acrocentric chromosomes. The nucleotides that make up a chromosome are numbered in an increasing order from 1 starting from the telomere of the p arm until the end of the telomere of the q arm. A unit of length derived from the number of nucleotides, the **base-pair**, can be used to describe the length of genomic elements or distances across a chromosome (often used with the kilo- and mega- prefixes such as kb for 1 000 base-pairs or Mb for 1 000 000 base-pairs). For example, chromosome 4 is 191 154 276 base-pairs in length, or approximately 191 Mb long (according to the hg19/GRCh37 reference human genome build).

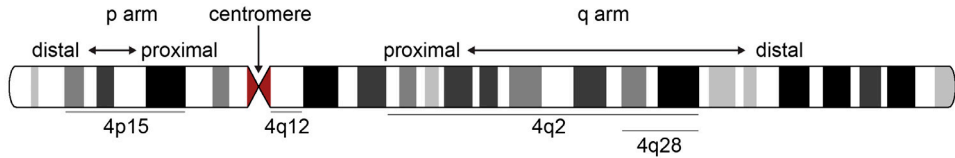


Figure 2. General chromosomal structure.

Human chromosomes, such as chromosome 4 displayed above, are divided into two arms separated by a centromere. The shorter arm is denoted as the p arm and the longer one as the q arm. The proximal parts of arms are those closer to the centromere, while the distal ones are those farther away towards the ends of each arm, or the telomeres. Chromosome arms are divided into regions (*e.g.* 4q2) which are, in turn, further divided into bands (*e.g.* 4q28) and sub-bands.

Copy number variants (CNVs) involve either the loss or gain of DNA sequences, be it whole chromosomes or pieces of them. Loss of DNA can be *hemizygous* if one copy is lost or *homozygous* if both copies are lost. Tumour suppressor genes such as *CDKN2A* in osteosarcoma and chondrosarcoma are often hemizygously and sometimes homozygously lost and *RBI* is deleted in about half of osteosarcomas^{30, 43-45}. It is, however, quite common that these deletions affect large parts of chromosomes and do not target these genes specifically. Conversely, gain of DNA involves the duplication of a sequence to end up with three or more copies. Amplification of certain oncogenes so that cancer cells harbour tens of copies of the gene is common in sarcomas, especially the *CDK4* and *MDM2* genes in almost all parosteal osteosarcoma, well-differentiated and dedifferentiated liposarcoma and about 10% of high-grade conventional osteosarcoma⁴⁶⁻⁵². Other commonly gained or amplified chromosomal regions in conventional osteosarcoma include parts of chromosome arms 8q and 17p, including the *MYC* and *COPS3* genes for example, as well as numerous genes in chromosome arm 6p that are involved in cell cycle regulation or bone development^{26, 53-56}. For an illustration of CNV detection, see Figure 3. Copy number gains and losses are generally accompanied by structural rearrangements due to the introduction of breakpoints in the DNA, leading to a spatial reorganisation of the DNA on top of the acquired copy number changes.

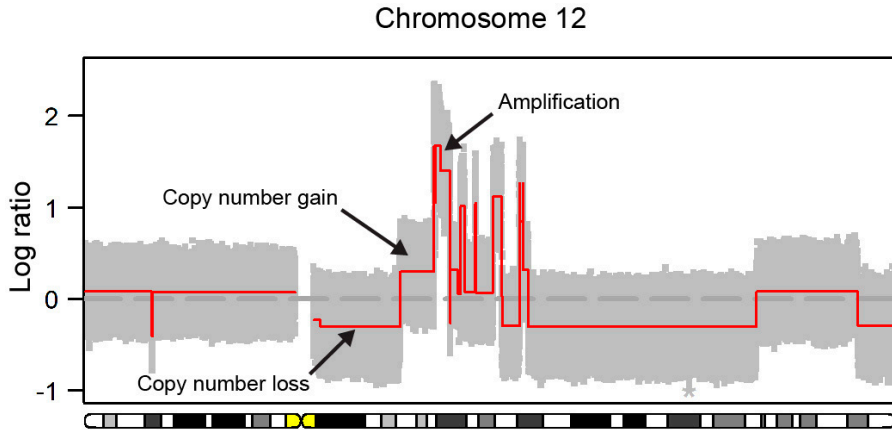


Figure 3. Copy number variants affecting chromosome 12 in a high-grade osteosarcoma.

Copy number losses (red segments with a log ratio below zero) affect multiple regions of chromosome 12. Copy number gains (red segments above zero) were also detected. In some instances, numerous extra copies of some regions were observed, a phenomenon known as copy number amplification.

Structural variants

In tumour cells, the DNA strands that make up our chromosomes can be rearranged. This creates “modified” chromosomes that contain a mixture of DNA sequences that would have normally resided on two or more separate chromosomes, or much further away on the same chromosome. In some instances, such as in low-grade osteosarcoma, gained and/or amplified material can reside in ring or giant rod marker chromosomes^{47, 52}, structures that generally should not exist in normal cells. A structural variant (SV) is thereby a physical rearrangement of one or more chromosomes or pieces of them. SVs can be either balanced or unbalanced in nature, where balanced rearrangements do not lead to a change in copy number, while unbalanced rearrangements involve material being either lost and/or gained. Structural variants change the spatial arrangement of DNA and lead to abnormal karyotypes as observed using standard cytogenetic techniques. *Interchromosomal* rearrangements are structural variants that bring together two or more separate chromosomes, whereas *intrachromosomal* rearrangements occur within the same chromosome. Examples of these could be **tandem duplications** where a duplicated sequence is placed right beside the original copy, **inversions** where part of the chromosome is inverted in place, or simple **deletions** where a part of the chromosome is lost and the remaining pieces are simply attached to each other (Figure 4A-C). Simple **duplications** can also occur, but the extra copies are then spread out on other chromosomes or further away on the same chromosome.

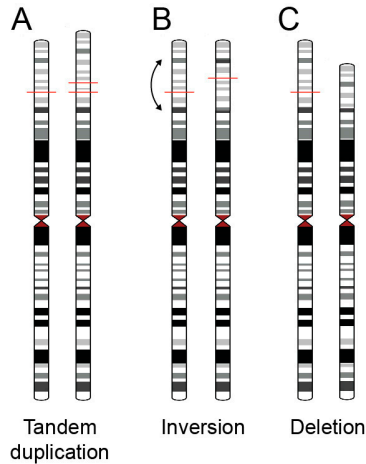


Figure 4. Examples of intrachromosomal rearrangements.

The red line represents the locus of the *EVA1B* gene on chromosome 1 and is used as an illustrative example. (A) Tandem duplications involve the duplication of a certain DNA sequence one or multiple times and the subsequent insertion of the duplicated sequence directly next to the original one. (B) Inversions involve two double stranded DNA breaks and the subsequent reattachment of the DNA sequences in-between the breaks in the opposite sense. (C) Deletions involve the loss of a DNA sequence and the subsequent reattachment of the two remaining pieces of the chromosome.

More complex types of chromosomal rearrangements have been observed throughout the years that involve a plethora of events occurring, either at once or over multiple cell cycles, that cannot be described as a simple rearrangement. For example, **breakage-fusion-bridge** cycles were first described by Barbara McClintock over 80 years ago^{57, 58}. Breakage-fusion-bridge cycles typically occur when an *interchromosomal* rearrangement creates a dicentric derivative chromosome. As this structure contains two centromeres, it may be pulled to opposite poles of the cell simultaneously during mitosis, forming a DNA bridge. This bridge will at one point break and when the broken chromosome is replicated, it may fuse to itself or another broken chromosome. This may once again create a dicentric derivative chromosome that may break during mitosis. This cycle repeats until the structure manages to cap itself and only contain a single centromere. As this normally takes several cell cycles to achieve, it leads to copy number amplifications due to the recurrent cycles of replication and the duplicated sequences ending up in the same daughter cell (Figures 5 and 6A). Breakage-fusion-bridge cycles are known to occur in sarcomas and are thought to cause massive *intratumour* heterogeneity⁵⁹.

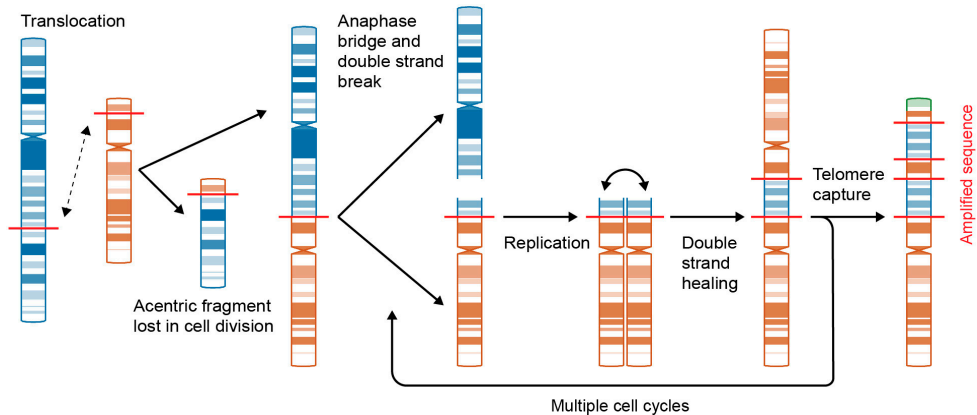


Figure 5. An example of how breakage-fusion-bridge cycles can occur.

An interchromosomal rearrangement creates an unstable dicentric chromosome and an acentric fragment which will be lost in cell division. During anaphase, or when the sister chromosomes are pulled to the opposite poles of the cell, the dicentric chromosome may be simultaneously pulled to both poles, forming a DNA bridge that will then break. During the next cell cycle, the broken chromosome will replicate and may then ligate to its sister chromosome leading to the creation of a new dicentric chromosome. This may repeat for multiple cell cycles until the broken end manages to capture a telomere and become stable. This results in multiple rounds of duplications of certain sequences, leading to copy number amplification.

In the past decade, high-throughput sequencing has identified two phenomena that cause significant chromosomal rearrangement. The first, **chromothripsis**, or chromosome shattering, involves an extreme number of *de novo* rearrangements resulting from a single catastrophic event: the shattering of a chromosome into numerous pieces^{7, 60}. The cell then attempts to salvage these pieces by stitching them together randomly, often leading to copy number losses as some pieces are lost completely during the cell cycle and to copy number gains as some pieces may be duplicated. Chromothripsis thereby leads to interspersed copy number losses and gains in addition to massive genomic rearrangement. This may affect a single chromosome, a part of a chromosome or parts of multiple chromosomes (Figure 6B-C). In a recent massive whole-genome sequencing study, chromothripsis was demonstrated to occur in a wide range of cancer types, albeit to varying degrees⁶¹. The second, **chromoplexy**, or chromosome weaving/braiding, involves multiple DNA breaks that are ligated to each other leading to new chromosomal configurations and loss of some genetic material. These rearrangements look like chains that bridge the gap between deleted material leading to the observation of **deletion bridges**. Chromoplexy was first described in prostate cancer⁶ but has also been observed in various soft tissue and bone tumours such as Ewing sarcoma, chondromyxoid fibroma, synovial sarcoma and phosphaturic mesenchymal tumour⁶². Chromoplexy is thought to be responsible for

the generation of characteristic gene fusions or the deactivation of certain genes in the aforementioned tumour types.

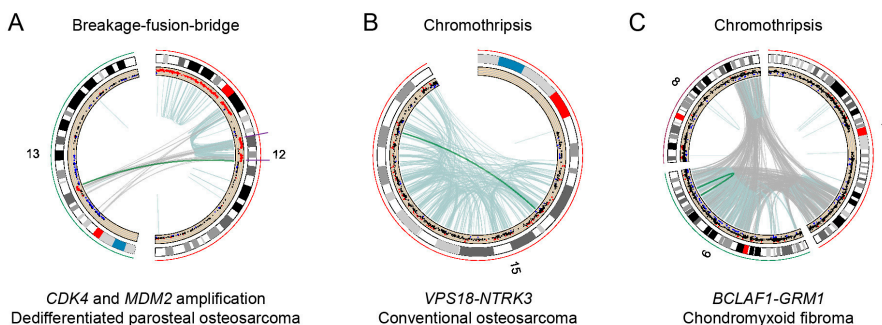


Figure 6. Examples of breakage-fusion-bridge cycles and chromothripsis in tumour samples generating driver and passenger mutations.

(A) In a dedifferentiated parosteal osteosarcoma, a dicentric chromosome resulting from the rearrangement of chromosomes 12 and 13 instigates a series of breakage-fusion-bridge cycles leading to amplification of numerous DNA sequences, including the *CDK4* and *MDM2* genes (purple lines). Only amplified material (elevated red dots) is rearranged, unlike in classical chromothripsis where amplification is generally not observed. (B) In a high-grade conventional osteosarcoma, chromothripsis of chromosome 15 was observed. Copy number states oscillate between $n=1$ (blue dots), $n=2$ (black dots) and $n=3$ (red dots) and there is extreme intrachromosomal rearrangement (light blue arcs). A non-functional *VPS18-NTRK3* fusion is created as a result (dark green arc). (C) In a chondromyxoid fibroma, chromothripsis of chromosome 6 and parts of chromosomes 1 and 8 occurred. Copy number states of the affected regions oscillate between $n=1$, $n=2$ and $n=3$ and there is an extreme number of inter- (grey arcs) and intrachromosomal rearrangements (light blue arcs). One of the rearrangements creates the *BCLAF1-GRM1* promoter swapping event (dark green arc).

Gene fusions and promoter swapping

The first ever recurrent neoplastic rearrangement was discovered in 1960 and was termed the ‘Philadelphia chromosome’ in chronic myeloid leukaemia⁶³. In the early 1970s, the Philadelphia chromosome was determined to be a result of the rearrangement of chromosomes 9 and 22, denoted as $t(9;22)$, leading to the creation of the first ever identified fusion gene, *BCR-ABL1*, discovered in the mid-1980s^{64, 65}. Advances that followed led to the understanding that chromosomal rearrangements often create gene fusions and by the year 2007, around 350 gene fusions had been discovered in various cancer types. With the gradual introduction of novel deep sequencing techniques in the following years, this number began to grow exponentially and has by now reached to more than 32 550 gene fusions involving around 14 000 genes (as of July 2020) reported in the Mitelman Database of Chromosome Aberrations

and Gene Fusions in Cancer⁶⁶. This staggering number of fusions was reached with the use of algorithms, and not all identified gene fusions are driver events that contribute to tumour progression. Some may simply be passenger events due to chromosomal rearrangement, while others may be technical noise due to algorithmic inaccuracies. Importantly, the creation of a gene fusion is not a given when a structural rearrangement occurs. Some tumour suppressor genes are often partially deleted, and as a breakpoint was created and must be fixed, a structural variant occurs that does not necessarily create a functional gene fusion but is rather a way to heal a DNA break.

Fusion genes are usually comprised of two partner genes denoted as 5' (five prime) and 3' (three prime) partners. Due to the chemical structure of the nucleotides that make up DNA and RNA molecules, nucleic acids are synthesized in the 5' to 3' direction. The end result is a 5' end where the 5th carbon atom will be bound to a phosphate group, and a 3' end where the 3rd carbon atom is bound to a hydroxyl group. The first part of a fusion gene, the 5' partner, provides regulatory sequences such as promoter and enhancer regions and often, but not always, protein coding regions. The second part of a fusion gene, the 3' partner, provides protein coding regions. In a classical gene fusion, both partners provide protein coding sequences leading to the production of a chimeric protein, driven by the active promoter of the 5' partner gene. For example, in Ewing sarcoma, the *EWSR1* gene in chromosome 22 is recurrently rearranged with a member of the *ETS* family of transcription factors, the most common being *EWSR1-FLI1* in 85% of cases and *EWSR1-ERG* in 10% of cases, although other combinations not involving *EWSR1* exist⁶⁷. The observed fusions in Ewing sarcoma are normally generated via balanced structural rearrangements and encode chimeric oncoproteins that are considered a driving event in Ewing sarcoma (Figure 7A). In 'promoter swapping' or 'enhancer hijacking' scenarios, the 5' partner only provides non-coding regulatory sequences that lead to upregulation of the entire coding sequences of the 3' partner gene. For instance, in chondromyxoid fibroma, various 5' partners can be rearranged with the entire coding sequence of the *GRM1* gene. These rearrangements provide a strong active promoter that leads to upregulation of *GRM1*, such as the *BCLAF1* promoter⁶⁸ (Figures 6C and 7B). Similarly, in aneurysmal bone cysts, the *USP6* gene is placed under the control of a variety of active promoters leading to upregulation⁶⁹. No chimeric proteins are created in these cases. Upregulation of a 3' partner gene can over-activate a specific signalling pathway or cellular function that most likely fits in to one of the hallmarks of cancer. Interestingly, the strong promoter or enhancer of the 5' partner gene need not be placed in the direct vicinity of the 3' partner gene. There are numerous examples of rearrangements of the immunoglobulin loci that place their active enhancer elements elsewhere in the genome leading to overactivation of oncogenes in B-cell malignancies⁷⁰. The *MYC* gene is often

overexpressed via such a rearrangement, and the breakpoint can be up to 1 Mb away from *MYC*. Over the past decade, a number of gene fusions involving three partner genes have been described in leukaemia and in lung and prostate carcinomas⁷¹⁻⁷⁵. In the papers included in this thesis, we describe a recurrent promoter swapping event involving the promoter of a tumour suppressor gene that reacts to ongoing DNA damage and a three-partner fusion in an osteosarcoma, amongst other findings.

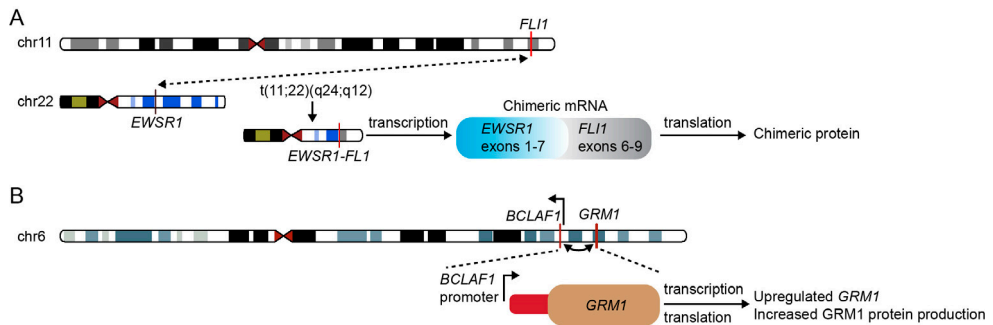


Figure 7. The difference between classical gene fusions and promoter swapping.

(A) In 85% of Ewing sarcoma cases, a balanced rearrangement between chromosomes 11 and 22 leads to the creation of the *EWSR1-FLI1* fusion gene. The fusion gene is then transcribed into RNA and subsequently translated into a chimeric oncoprotein. (B) In chondromyxoid fibroma, the complete *GRM1* gene is placed under the control of an active promoter, in this case the *BCLAF1* promoter. *BCLAF1* does not contribute any coding sequences and the result is an upregulation of the *GRM1* gene and likely an increased *GRM1* protein production.

Current clinical problems

Despite the recent advances in cancer research, there is still much to be learned about the molecular mechanisms underlying the development of bone tumours. In routine diagnostics, it can sometimes be difficult to distinguish between benign, intermediate and malignant entities based on morphology and other biological parameters. This has necessitated the search for new characteristic genomic alterations that can be used to improve diagnostic accuracy. The hope on the longer term is to improve patient outcome via novel treatment strategies that increase survival rates while simultaneously reducing long-term side effects. Are there recurrent targetable alterations that can have an almost immediate effect on patient outcomes waiting to be discovered? We set out to elucidate genetic subgroups of the primary osteogenic tumours osteoblastoma and osteosarcoma as described next.

Osteoblastoma

Very little was known about the genetics underlying osteoblastoma development until the recent description of the ubiquity of structural rearrangements affecting the *FOS* gene in chromosome band 14q24 or, more rarely, its paralogue *FOSB* in chromosome band 19q13 in osteoblastoma and osteoid osteoma¹⁴. This finding was based on genome sequencing studies on six osteoblastomas that demonstrated that the terminal exon of the *FOS* gene is rearranged in a mostly non-recurrent manner. It is thought that such rearrangements result in an increased activity of the FOS protein, achieved by placing extra nucleotides within the terminal exon of *FOS* that lead to stabilisation of the *FOS* mRNA, thereby producing more protein and bestowing an increased resistance to protein degradation^{14, 76}. The six sequenced osteoblastomas did not show any copy number aberrations. The *FOS/FOSB* finding was then corroborated with lower resolution techniques like fluorescence *in-situ* hybridisation, or FISH for short – described in the Materials and Methods section – and immunohistochemistry on a larger series that demonstrated that up to 89% of osteoblastomas have *FOS*, or more rarely, *FOSB* rearrangements. Additional studies using FISH and immunohistochemistry were also carried out yielding similar results^{77, 78}. There nonetheless remained a portion of cases that did not display a rearrangement or immunohistochemical reactivity of *FOS* or *FOSB*. Our group had previously conducted copy number investigations on a series of osteoblastoma, finding recurrent deletions affecting a part of chromosome band 22q12 in a subset of cases and identifying a few genes, such as *NF2* and *ZNRF3*, as potential targets of these deletions¹⁵. As osteoblastoma is locally aggressive and can in some instances be

morphologically confused with osteoblastic osteosarcoma^{13, 79}, a subtype of conventional osteosarcoma, it is imperative to diagnose properly. This can be achieved with the help of defining genetic alterations. Patients wrongly diagnosed with a malignant osteosarcoma can suffer greatly during its aggressive treatment when, in reality, surgical excision of what was an osteoblastoma would be sufficient in almost all cases. In **Article I**, we set out to pinpoint the target gene(s) affected by homozygous deletions in a subset of osteoblastoma in an expanded series of tumours, as this subgroup with deletions may be the “type” of osteoblastoma that is negative for *FOS* or *FOSB* rearrangements.

Conventional osteosarcoma

Conventional osteosarcoma has been shown to display an extreme number of copy number and structural alterations²⁴⁻³¹, yet there are still no reports on genetic alterations specific to this disease. This genetic complexity has long acted as a barrier towards understanding how this cancer arises. The fact that it presents with such complexity in children and adolescents is also baffling, as it is thought that such young individuals have not lived long enough to acquire a staggering amount of genetic alterations, often seen in carcinomas of the elderly. Patients with predisposing genetic disorders, such as Li-Fraumeni syndrome or hereditary retinoblastoma have higher risks of developing osteosarcoma, likely due to them already having germline loss-of-function mutations in the *TP53* and *RBI* tumour suppressor genes, respectively⁸⁰⁻⁸². It was not until the last two decades that some recurrent aberrations started to be delineated in high-grade conventional osteosarcoma, such as copy number gains and amplifications of chromosome arms 6p, 8q, 12p and 17p^{53, 56, 59, 83}. In the last decade, the *TP53* and *RBI* genes have been shown to be mutated in about 90% and 50% of sporadic osteosarcomas, respectively^{27, 28, 45}. Homozygous loss of *CDKN2A*, another tumour suppressor gene, occurs in around 10% of conventional osteosarcomas⁴⁴. Even though the osteosarcoma genome has been thoroughly investigated, we still do not understand how such genetically complex tumours can thrive and whether there are genetically unique subgroups of conventional osteosarcoma. Conventional osteosarcoma can be divided into multiple histological subtypes including, amongst others, osteoblastic, chondroblastic, fibroblastic and chondroblastoma-like. Chondroblastoma-like osteosarcoma is an extremely rare subtype of conventional osteosarcoma representing less than 1% of cases^{84, 85}. As no reports have been published about the underlying genetics of chondroblastoma-like osteosarcoma, we investigated a rare case using genomic and transcriptomic methods in **Article II**.

Since there are currently no known targetable alterations in osteosarcoma, we searched for the potential presence of targetable neurotropic tyrosine receptor kinase (*NTRK*) gene fusions. Such fusions were first identified in 1982 and have since been described in a multitude of malignant tumour types^{86, 87}. The most usual form of such fusions is composed of a 5' partner that, via dimerization domains, enables constitutive activation of the 3' tyrosine kinase domain of either of the *NTRK1*, *NTRK2* or *NTRK3* genes⁸⁸. In some malignancies, *NTRK* fusions may be found in the majority (>80%) of cases, whereas in others they are much rarer (<5%)⁸⁹. Nonetheless, the pan-NTRK tyrosine kinase inhibitors (TKIs) larotrectinib and entrectinib have shown response rates above 75% in clinical trials on various tumour types^{90, 91}. As no new treatments for osteosarcoma have been introduced in the last three to four decades¹⁶, a survey of RNA-sequencing data on 113 osteosarcoma samples was conducted in **Article III** in search of potentially targetable *NTRK* gene fusions.

It is unknown whether the genetic complexity seen in conventional osteosarcoma can actually drive the tumour forward. For example, the *TP53* gene is often affected by structural variants in conventional osteosarcoma, but this was explained as the tumour's way of silencing the gene^{26, 27, 29, 32}. While this is true, some osteosarcomas actually preserve and relocate the *TP53* promoter region, and there have been no studies exploring the effects of such a relocated regulatory region of a tumour suppressor gene that responds to DNA damage. Rearrangements of *TP53* in osteosarcoma were already observed in the late 1980s using what would now be considered outdated molecular techniques, leading researchers to speculate that relocation of the *TP53* promoter to control another unidentified gene was a possible result of such rearrangements^{92, 93}. Given this advancement of knowledge in osteosarcoma, there are still no defining genetic alterations specific to the tumour, and it may be that there are subgroups of osteosarcoma that have differing genetic characteristics. As studying the effects of a relocated promoter requires access to matched genomic and transcriptomic information, we integrated copy number, structural variant and transcriptomic information from osteosarcoma biopsies and *in vitro* cell models to study the effects of such a relocated promoter region in **Article IV**.

Low-grade and dedifferentiated osteosarcoma

Low-grade parosteal osteosarcomas have long been known to harbour supernumerary ring and/or giant rod marker chromosomes containing amplified material from chromosome bands 12q13-15, including the *CDK4* and *MDM2* genes, in what is otherwise a normal genome⁵². Later studies in the early-mid 2000s further corroborated these findings and discovered that additional material from other parts of chromosome

12 or other chromosomes may also be included in these supernumerary chromosomes, although this was mostly non-recurrent^{47, 48}. A portion of low-grade central osteosarcomas also harbour such aberrations^{46, 94}. It is not known whether any more genes are specifically targeted for amplification in low-grade osteosarcomas, and more recent reports have shown that about 10% of high-grade conventional osteosarcomas also harbour amplifications of *MDM2* and/or *CDK4*⁴⁹. Whether these amplifications occur due to the same underlying mechanisms in high-grade and low-grade tumours or are the result of different processes is not known. We therefore set out, in **Article V**, to analyse a series of osteosarcomas with 12q-amplifications using high-resolution copy number and genome sequencing in search for more potential target genes and whether there truly are mechanistic differences behind how the chromosome arm 12q amplifications arise in low-grade versus high-grade osteosarcomas.

Novel investigations into primary osteogenic bone tumours

Aim

The overarching aim of this thesis was to elucidate new genetic mechanisms underlying the development of a variety of primary osteogenic skeletal tumours, and thereby identify novel diagnostic alterations that are potentially clinically targetable in hopes of improving treatment outcome for patients. More specifically, the individual articles discussed in this thesis aimed to:

Article I – determine the target gene(s) in chromosome band 22q12 affected by deletions in a subset of osteoblastoma.

Article II – genetically characterise a very rare chondroblastoma-like osteosarcoma/malignant phosphaturic mesenchymal tumour of bone.

Article III – assess the prevalence and potential functionality of *NTRK* gene fusions in conventional osteosarcoma.

Article IV – determine the prevalence and importance of *TP53* promoter rearrangement events to the development of high-grade conventional osteosarcoma.

Article V – further genetically characterise and subdivide low- and high-grade osteosarcomas with amplification of *MDM2* and/or *CDK4* in chromosome arm 12q.

Materials and Methods

Patients and tumour samples

All tumour samples analysed in this thesis were obtained as fresh-frozen tumour biopsies from the following sarcoma centres in Europe: the Skåne University Hospital and the Karolinska Hospital in Sweden, the University Hospital Basel in Switzerland, the Klinikum rechts der Isar of the Technical University of Munich in Germany, the Leiden University Medical Centre in the Netherlands, and the Royal Orthopaedic Hospital in Birmingham, United Kingdom. All samples were obtained after informed consent, and the studies were approved by the appropriate ethical review committees of the institutions involved. In total, one or more samples from 29 osteoblastoma and 168 osteosarcoma patients were analysed in **Articles I, II, IV and V**. An additional osteosarcoma cohort was included in **Article III** together with a part of the 168 osteosarcomas studied in the other articles.

DNA and RNA extraction

DNA and RNA were extracted from fresh-frozen tumour biopsies. A small piece of the tumour (approximately 200-300 mm³) was cut out, pulverised and then split in two halves, one for DNA extraction and one for RNA extraction. This was done whenever possible to obtain DNA and RNA that are representative of the same part of the tumour, as different sections of the tumour may have significant differences on the genomic and transcriptomic levels due to intratumour heterogeneity. Having DNA and RNA representing the same piece of tumour is helpful in confirming gene fusions or promoter swapping events on both the genomic and transcriptomic levels, and few studies on bone tumours have included matched DNA and RNA information on many samples. The powder, containing cells from the tumour, is then incubated with different lysate buffers depending on whether DNA or RNA is being extracted. For DNA extraction, the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) was used, while for RNA extraction, the RNeasy Lipid Tissue kit (Qiagen) was used for extractions from tumour biopsies and the RNeasy Micro kit (Qiagen) was used for extractions from cell cultures.

Polymerase chain reaction

Polymerase chain reaction (PCR) is a technique used to amplify a certain nucleotide sequence of interest. The starting material can be DNA (genomic PCR) or RNA that is reverse transcribed using a 'reverse transcriptase' enzyme into complementary DNA and then amplified using PCR (RT-PCR). Predesigned oligonucleotides called primers are used to bind 5' and 3' of the region of interest, and then a DNA polymerase enzyme amplifies the target region during a series of thermal cycles. In real-time quantitative PCR (denoted RT-qPCR), probes specific for the gene of interest are used in addition to the primers after reverse transcription to assess the relative expression level of a gene. The probes hybridise to their complementary sequences and are degraded by the DNA polymerase enzyme during the PCR cycles, emitting signals as a result. These can be quantified and then used to compare the relative expression levels between different samples.

Sanger sequencing

Sanger sequencing is a technique developed in the late 1970s by Frederick Sanger and colleagues based on the selective use of chain-terminating radioactive nucleotides during *in vitro* DNA replication to sequence amplified DNA fragments^{95,96}. Nowadays, fluorescent chain-terminating nucleotides are used instead. This technique, although quite old now, is still useful when validating gene fusions on a small scale, for example. Newer techniques, described below, allow for the detection of gene fusions but suffer from high false-positivity rates. RT-PCR followed by Sanger sequencing is therefore a way of validating that the fusion gene actually exists in an RNA sample from a tumour and is not just a technical artifact. To do so, primers are designed to bind to specific sequences 5' and 3' of the suspected fusion breakpoints, enabling amplification via PCR and subsequent Sanger sequencing.

Fluorescence *in situ* hybridisation

Fluorescence *in situ* hybridisation (FISH) is a method where fluorescently labelled probes are used to bind either specific complementary DNA sequences or whole chromosomes (whole chromosome paint). FISH can be used to detect copy number alterations affecting one or more genes or chromosomal regions of interest, and to detect chromosomal rearrangements. If cells are cultured first, FISH can be carried out on the cells that are in metaphase during mitosis, otherwise interphase FISH can be used. FISH analyses are targeted in nature and will only detect what they are designed to detect, albeit with a top resolution of around 10 kb. There can also be issues with

false positives, false negatives and background signals, especially when using non-commercial probes or the material at hand is of lower quality.

SNP array

A single nucleotide polymorphism (SNP) is defined as a SNV that occurs in at least 1% of the population. There are hundreds of millions of naturally occurring SNPs in the human genome, and a single individual may have up to between 4 and 5 million SNPs in their genome⁹⁷. These can be used as part of SNP arrays to determine copy number alterations and copy number neutral loss of heterozygosity in DNA samples. SNP arrays consist of millions of oligonucleotide probes attached onto chips. Fragmented DNA from the sample under study is then applied to the chip and allowed to hybridise, after which the hybridised probes emit signals. The signal intensities are recorded and normalised, and thereafter converted into \log_2 ratios and allele frequencies indicating copy number and allelic distribution, respectively. The majority of cases included in this thesis were analysed with CytoScan HD arrays (Thermo Fischer Scientific, Waltham, MA, USA), and a minority had already been analysed using the Human Omni-Quad BeadChips (Illumina, San Diego, CA, USA). CNVs were identified through visual inspection using the appropriate software for each platform, and by segmenting \log_2 ratios using the R package 'copynumber'.

Although very well equipped to detect larger copy number imbalances, SNP arrays have multiple limitations. Balanced chromosomal translocations do not lead to loss or gain of genetic material and will therefore go undetected. The copy number alterations observed are due to imbalanced rearrangements, but it is not possible to know how the chromosomes are rearranged. For example, amplifications on multiple chromosomes in osteosarcomas may reside together in a marker chromosome, but this would not be apparent using SNP array; they instead seem to be individual, unrelated amplicons. Normal cell content in tumour biopsies can also affect detection of copy number imbalances. In samples with high normal cell content, differentiating true imbalances from background noise can prove challenging. Additionally, subclonal events in heterogenous tumours such as osteosarcomas may be hard to confidently detect.

Next generation sequencing

Next generation sequencing (NGS), also referred to as deep or massively parallel sequencing, is a technique used to sequence DNA or RNA at extreme depths, opening a path of discovery into tumour genetics that was not previously possible. NGS revolutionised the field of cancer genomics over the last decade, offering unparalleled

coverage of the cancer genome and transcriptome. Although the technique itself is now over 10 years old and is essentially “current generation”, new applications and improved sequencing machines are being introduced regularly. NGS allows for a mostly unbiased exploration of tumour genetics, as DNA and RNA can be sequenced without using any targeted approaches (although such targeted approaches do exist). In this thesis, multiple NGS techniques were used, as described below.

In brief, libraries were prepared with a sample preparation kit specific to each sequencing technique (see below for details). The NextSeq 500 (Illumina) sequencing machine was used in most instances. Libraries were sequenced using high-throughput flow-cells that generate around 400 million reads per run.

RNA-sequencing

RNA-sequencing can be used to assess global gene expression profiles, to detect gene fusions and, in some instances, even promoter swapping events. The expression levels of specific genes can be compared between different cases, something that is especially helpful when studying gene fusions and promoter swapping events as one expects an upregulation of the 3' partner. The expression levels of the different exons of genes that have been disrupted by SVs leading to gene fusions or promoter swapping events can also be quantified. RNA-sequencing can even be employed as a method for detecting SNVs affecting exonic sequences. In this thesis, RNA-sequencing was used in all articles. Library preparation was carried out using the TruSeq RNA Sample Preparation Kit v2 (Illumina). Using this kit, we only capture RNA molecules with poly A-tails at their 3' end, meaning all ribosomal and many non-coding or immature RNA molecules will not be captured and sequenced. There are different RNA capture methods that can be used depending on the question at hand.

Whole-exome sequencing

Whole-exome sequencing (WES) is used to sequence exonic or coding DNA sequences at a very high depth. The Nextera Rapid Capture Exomes kit (Illumina) was used to prepare exome-libraries from two samples in **Article I**. WES allows for the confident detection of SNVs and small insertions and deletions (indels), and for the assessment of CNVs affecting exonic sequences, but it requires access to matched normal controls. This gives a relatively accurate picture of the genome-wide copy number profile. An obvious disadvantage with WES is that no information is obtained about intronic and other non-genic parts of the genome (approximately 99% of the genome is not sequenced at all actually). This hampers the detection of important SNVs and indels

located outside exons, and the detection of SVs is generally not possible, as they rarely affect exons. Most SV breakpoints are either in intronic sequences or outside of genes altogether.

Whole-genome mate pair sequencing

Whole-genome mate pair sequencing (simply referred to as mate pair sequencing from here on out) is a robust methodology used to detect both intra- and interchromosomal rearrangements. The technique utilises a clever circularisation step during library preparation, resulting in large insert sizes between sequencing reads and a reverse-forward read pair orientation. This offers a high spanning coverage of the human genome, despite having low depth. This technique was used in all articles discussed in this thesis and is the basis for most of the findings we have demonstrated. Mate pair libraries were prepared using the Nextera Mate Pair Library Preparation kit (Illumina). Mate pair sequencing allows for the detection of genome-wide chromosomal rearrangements and is not restricted to exonic sequences as WES is. However, it is not possible to detect SNVs and small indels with mate pair sequencing. Copy number analyses are not recommended either due to the low depth provided by the technique and lack of normal control. A normal control is not required for detection of SVs, as such rearrangements should not be present in normal cells. Having a normal control would, however, facilitate filtering of technical artifacts, although there are steps one can take to circumvent this limitation.

Whole-genome paired end sequencing

Whole-genome paired end sequencing (WGS) is one of the most widely used sequencing technologies nowadays. It allows for the detection of SNVs, indels, SVs and CNVs. It does however require access to matched normal controls to be a resilient method of mutation detection. Nonetheless, WGS can provide all the information the aforementioned methods do individually, with the exception of RNA-sequencing. The major downside to WGS is the high depth required to confidently be able to call CNVs, SVs and SNVs. Coupled with the need for a normal control, this translates into a very large amount of data that is both expensive to generate and store. Structural variant investigations using combined SNP array and mate pair sequencing can offer the same information WGS does at a lower cost, although small variants can be missed by the former two methods. WGS was carried out using the Complete Genomics platform (BGI Europe A/S, Copenhagen, Denmark) on one case in **Article I** and in collaboration

with a group at the University Hospital Basel in Switzerland on another five cases in **Article IV**.

Single cell whole-genome sequencing

Low-pass single cell WGS was used to study the copy number profiles of cryopreserved single cells from selected cases in **Articles I** and **IV**. With a coverage of only 0.01x over the human genome, the technique can only be used to detect CNVs affecting genomic regions around 1Mb in size. This was carried out in collaboration with a group at the European Research Institute for the Biology of Ageing in Groningen, the Netherlands using a modified single cell WGS library preparation protocol⁹⁸.

Nextera sequencing

Nextera sequencing is a deep sequencing method that specifically captures and amplifies a sequence of interest, such as a few exons of a gene that were artificially mutated using genome editing technologies. The Nextera XT DNA library preparation kit was used (Illumina) to investigate exactly how the *TP53* gene was knocked-out in a modified cell line (see *In vitro* cell models below).

Bioinformatics

The amount of data generated by NGS techniques is staggering, but with the help of bioinformatic programs and algorithms, this data can be analysed and interpreted. Sequencing machines generate raw base calls that can be transformed into fastq files. These files contain the unique name, sequence and quality scores of every single read pertaining to a certain sample. The lower the quality score, the more likely the read base-pair is erroneous. A quality control step is usually performed at this stage prior to downstream processing of the fastq files. Depending on the sequencing technique used, a variety of different programs are used to further analyse the raw data. In general, sequencing reads must be mapped to the human genome before further analyses such as detection of SNVs, SVs or CNVs can begin. The different mapping processes employed in articles in this thesis create BAM files that were then further processed using tools in the PICARD or Genome Analysis Tool Kit (GATK) suites before downstream analyses were carried out.

For RNA sequencing experiments in this thesis, fastq files were aligned to the hg19/GRCh37 reference human genome build using STAR⁹⁹. Cufflinks was used for quantification and FPKM normalisation of gene expression reads¹⁰⁰. FPKM stands for

fragments per kilobase of transcript per million reads and takes gene length as well as number of mapping reads into account, allowing for comparison of gene expression levels between samples. Normalised gene expression data was then analysed with the Qlucore software (Qlucore AB, Lund, Sweden). For the detection of fusion transcripts, both ChimeraScan^{101, 102} and FusionCatcher¹⁰³ were used initially, but ChimeraScan was replaced by STAR-Fusion¹⁰⁴ later on. Chimeric transcripts were considered true if identified by either of the algorithms used and had simultaneous support at the DNA level with another method, usually mate pair sequencing. In some instances, gene fusions detected via RNA-sequencing were validated using RT-PCR as described previously. For the detection of known disease-defining SNVs affecting selected genes in **Article II**, VarScan¹⁰⁵ was used following the best practice guidelines published by the GATK.

For DNA sequencing experiments in this thesis, fastq files were aligned to the hg19/GRCh37 reference human genome build using BWA-mem¹⁰⁶. For the detection of SVs, the structural variant callers TIDDIT¹⁰⁷ and Delly¹⁰⁸ were always used. In some instances, even Manta¹⁰⁹ was used. Given the complexity of rearrangements in osteosarcoma, multiple structural variant callers are needed to accurately represent a chromosomal rearrangement. As a rule of thumb, two of the three variant callers need to agree on how a variant is structured for it to be considered true. Visual inspection using the Integrative Genomics Viewer¹¹⁰ was always carried out to verify or refute findings. For the detection of CNVs from WGS and WES data, CNVkit¹¹¹ was used, using a matched normal control or combined normal control from multiple samples as a baseline for a diploid genome. Copy number analysis for cases with single-cell WGS was carried out using AneuFinder¹¹².

***In vitro* cell models**

The *TERT*-immortalised human foreskin fibroblast cell line Bj5ta (ATCC, LGC Standards, Middlesex, UK) was used for functional studies in **Article IV**. CRISPR/*Cas9*, a genome editing technology, was used to knockout the *TP53* gene in Bj5ta, using a guide RNA targeting exon 6 of *TP53*. Cells successfully transduced with Cas9 and the guide RNA will acquire antibiotic resistance, while cells that were not transduced will die. Surviving cells were then single-cell cloned and the specific *TP53* mutations confirmed by the Surveyor mutation detection kit (Integrated DNA Technologies, Inc., Coralville, IA, USA), Sanger sequencing, Nextera sequencing and SNP array analysis. One clone was finally chosen out of the validated ones as it had a 19 bp deletion affecting *TP53* exon 6 in one allele and the other *TP53* allele was lost along with the distal part of chromosome arm 17p. This translated into a complete

knockout of *TP53*. The *TP53* knockout cells are from here on out referred to as Bj5ta *TP53*^{-/-}.

As a proof-of principle experiment, the functional importance of a transferred *TP53* promoter region was modelled using a promoter-less vector containing the *TP53-ROR2* fusion gene that was transduced into Bj5ta *TP53*^{-/-} cells. The chromosomal region 2 000 bp upstream of *TP53* together with *TP53* exon 1 and 500 bp of intron 1 were used to embody the *TP53* promoter region. This promoter sequence was then attached to the last 500 bp of *ROR2* intron 1 and the coding exons 2-13 of *ROR2*. This chimeric sequence will be referred to simply as *TP53-ROR2* and contains the complete coding sequence of *ROR2* variant ENST000003757.1 under the regulation of the *TP53* promoter region. As controls, the same promoter-less vector as above containing no extra sequences (empty vector) or the vector containing only the *ROR2* sequences described above without the *TP53* promoter were also transduced into Bj5ta *TP53*^{-/-} cells.

Bj5ta *TP53*^{-/-} cells containing either the *TP53-ROR2* fusion or only *ROR2* were cultured either without or with varying concentrations of the DNA-damaging agent cisplatin. RNA was extracted and expression levels of *TP53* and *ROR2* were then analysed using RT-qPCR. Additionally, wild type Bj5ta cells, Bj5ta *TP53*^{-/-} cells and Bj5ta *TP53*^{-/-} cells containing either *TP53-ROR2* or the empty vector were subjected to RNA-sequencing to study the effects of introducing a *TP53*-fusion gene on the global gene expression profile relative to wild type Bj5ta and Bj5ta *TP53*^{-/-} cells.

Results

Our investigations into the genetic alterations underlying the development of primary osteogenic tumours have highlighted the role structural variation, or chromosome remodelling, plays in these tumours. Be it the creation of a gene fusion, a promoter swapping event or in gene deletion and amplification, a deregulation of cellular processes will occur and, in some instances, the cell's own checkpoint mechanisms can be turned to instead promote oncogenesis. A brief summary of the results of each individual article are presented below, after which a general discussion linking findings from the different articles follows.

Article I

Loss of *NF2* defines a genetic subgroup of non-*FOS*-rearranged osteoblastoma

In a series of 29 osteoblastoma cases, including five diagnosed as epithelioid osteoblastoma, we detected loss of whole or parts of chromosome arm 6q in five cases, of which four had concomitant loss of 22q12. This included homozygous loss of the *NF2* gene in all four 22q12 deleted cases and loss of *ZNRF3* in three of them, a finding mirrored in gene expression data. Notably, single cell WGS of an epithelioid osteoblastoma case showed that there is a large normal cell content but a copy number level between zero and one for parts of 22q12, indicative of a mixture of hemi- and homozygous losses, was nonetheless observed in neoplastic cells. Most other cases displayed no acquired copy number alterations. Mate pair sequencing revealed that 10/23 (43%) osteoblastomas with available structural variant data harbour balanced two- or three-way translocations affecting the terminal exon of the *FOS* gene, which is lower than what has previously been reported. No rearrangements affecting *FOSB* were detected in our series. None of the 6q and 22q12 deleted cases showed structural rearrangement of *FOS* or *FOSB*.

Article II

Genetic profiling of a chondroblastoma-like osteosarcoma/malignant phosphaturic mesenchymal tumor of bone reveals a homozygous deletion of *CDKN2A*, intragenic deletion of *DMD*, and a targetable *FNI-FGFR1* gene fusion

Copy number investigations on a chondroblastoma-like osteosarcoma revealed that multiple chromosomal regions were lost, including homozygous loss of *CDKN2A* in 9p21 and an intragenic deletion affecting *DMD* in Xp21. Copy number gains were also detected. Sequencing analyses identified an *FNI-FGFR1* gene fusion and an

upregulation of the *FGFR1* exons involved in addition to the *FGF23* gene, encoding a ligand of FGFR1. Defining gain-of-function hotspot SNVs were screened for in select genes to eliminate other morphologically overlapping diagnoses. The chondroblastoma-like osteosarcoma case under study showed no mutations in the studied genes nor any promoter swapping events known to occur in other types of bone tumours. However, since it harboured an *FN1-FGFR1* gene fusion, it cannot be excluded that this case is instead a phosphaturic mesenchymal tumour of bone.

Article III

***NTRK* fusions in osteosarcoma are rare and non-functional events**

In an RNA-sequencing screen of 113 conventional osteosarcomas, only three (3%) harboured detectable *NTRK* fusions. None of the identified fusions were deemed to be functional, let alone expressed, as the *NTRK* genes showed extremely low expression levels and the chimeric coding sequences either lacked a start codon, or had premature stop codons prior to the tyrosine kinase domain of the respective NTRK protein. Immunohistochemistry using a pan-Trk antibody showed no immunoreactivity in any of the three cases, likely due to lack of translated NTRK proteins.

Article IV

Oncogenes hijack a constitutively active *TP53* promoter in osteosarcoma

Combined copy number and structural variant analyses demonstrated that *TP53* is disrupted in around 40% of conventional osteosarcomas, resulting in separation of the promoter region from the coding parts of the gene which are often deleted. The *TP53* promoter region is however not lost but retained in most cases. In 15% of cases, the promoter is either gained or amplified. In such cases, a peculiar copy number profile was seen on chromosome arm 17p, something we have termed ‘*TP53* promoter gain’ and described in detail in the manuscript of **Article IV**. Interestingly, both the presence of breakpoints in *TP53* and ‘*TP53* promoter gain’ were non-randomly associated with a young age of diagnosis.

Analysis of structural variant data revealed that relocation of the *TP53* promoter occurred via two separate patterns. This was either by simple inter- or intrachromosomal rearrangements that did not further compromise chromosome stability, or through an initial interchromosomal rearrangement that created a dicentric derivative chromosome. The *TP53* SV was likely an early event that sparked a series of chromosomal rearrangements, often involving multiple chromosomes and leading to

copy number amplifications of its direct partner region and other regions of interest. There was no evidence of chromothripsis affecting chromosome 17 in cases with ‘*TP53* promoter gain’ which instead seem to display characteristics of breakage-fusion-bridge cycles. SV analyses of *TP53*-rearranged multi-sampled osteosarcomas using WGS showed that all primary tumour, local recurrence and/or metastatic biopsies harboured their respective *TP53* fusion. Single cell WGS of two *TP53*-rearranged cases also showed that all neoplastic cells harboured the respective *TP53* fusion found in the bulk DNA of the primary tumour. Additionally, transcriptomic analyses of *TP53*-rearranged osteosarcomas where a 3′ partner gene could be confidently identified, recurrently showed a higher expression level of *TP53* exon 1, relative to the coding exons of the gene dislocated from the promoter, and an upregulation of either the entirety of the 3′ partner or the exons of said partner that are involved in the fusion. The *TP53* promoter was thus active *in vivo*, *i.e.*, in the tumour cells directly prior to them being fresh-frozen.

To model these findings *in vitro*, Bj5ta *TP53*^{-/-} cells containing either a *TP53-ROR2* fusion or only *ROR2* sequences without an adjoined *TP53* promoter were exposed to varying concentrations of cisplatin, a DNA damaging agent used as part of the chemotherapy regimen osteosarcoma patients receive^{23,113}. *ROR2* expression levels were significantly higher in Bj5ta *TP53*^{-/-} cells containing *TP53-ROR2* than in Bj5ta *TP53*^{-/-} cells containing only *ROR2* even without any cisplatin. Increasing concentrations of cisplatin lead to incrementally higher *ROR2* expression. This demonstrated that the *TP53* promoter is constitutively active in a *TP53*^{-/-} background and can thereby drive the expression of an oncogene translocated into its vicinity. Its effect can be strengthened by additional DNA damage, as normally occurs in tumour cells. RNA-sequencing analyses of the Bj5ta *TP53*^{-/-} cell models revealed that around 3 000 genes were significantly downregulated due to *TP53* loss compared to wild type cells. Introduction of a *TP53-ROR2* fusion rescues the expression of approximately half of these genes, reverting the global gene expression profile back towards that of wild type Bj5ta cells.

Article V

Genetic rearrangements in osteosarcoma result in ectopic gene expression through transposition of regulatory elements and selective amplification

Our investigation into 25 osteosarcomas with 12q-amplifications elucidated four major subgroups based on integrated genomic copy number and structural alterations in addition to transcriptomic data as described next. The first group (Group A) consisted of five conventional osteosarcomas with amplification of *CDK4*, but not *MDM2*, and extensive genome-wide rearrangements. All five cases in this group harboured loss-of-

function mutations affecting *TP53*, of which four displayed ‘*TP53* promoter gain’. Amplification of *CDK4* was intertwined with rearrangement of *TP53*. The fifth case had a homozygous deletion of *TP53*. The second group (Group B) consisted of seven conventional osteosarcomas with amplification of both *CDK4* and *MDM2* in addition to extensive genome-wide copy number and structural alterations, mostly in the form of copy number amplifications. The third group (Group C) consisted of ten high- and low-grade osteosarcomas with amplification of both *CDK4* and *MDM2* in addition to a relatively small number of alterations outside of chromosome 12. Chromosome banding and FISH analyses of some of the low-grade osteosarcomas showed that amplified material was located to ring or giant rod marker chromosomes. The fourth and final group (Group D) consisted of three low-grade osteosarcomas with amplification of *CDK4* and *MDM2* along with other regions in chromosome 12. No genomic alterations outside of chromosome 12 were observed in these cases. Detailed follow-up information on these patients can be found in the manuscript of **Article V**.

Cases in Groups B and C showed recurrent rearrangement of the *PLEKHA5* and *FRS2* loci in 12p12 and 12q15, respectively, in a similar manner as to how the *TP53* locus was rearranged in cases in Group A and **Article IV**. Due to the complex rearrangements affecting chromosome 12, a three-way *ALDH2-PLEKHA5-ATF7* and a canonical two-way *PLEKHA5-EPS8* gene fusion were detected in two cases. Upregulation of the 3' partner gene was observed in both *PLEKHA5*-rearranged cases, which showed particularly aggressive clinical behaviours. Another nine cases in Groups B and C displayed SVs that separated the *FRS2* promoter from its coding sequences. A 3' partner was confidently determined in five *FRS2*-rearranged cases which were low-grade osteosarcomas or tumours with a low-grade origin. *FRS2* was rearranged with the *ADAM32*, *SSPN*, *ELF1*, *FAM60A* and *ANKRD62* genes, all of which, except for *SSPN*, showed the highest expression level in the respective case.

The four discerned groups had multiple differences and similarities between them. All cases in this study had amplification of *CDK4*. Cases in Group A had *TP53* gene fusions or homozygous loss, while cases in the other three groups had *MDM2* amplification instead and most cases in groups B and C showed promoter swapping events involving either the *PLEKHA5* or *FRS2* regulatory regions. The similarities and differences were also evident in global gene expression analyses, where groups A and C formed distinct ‘expression profile clusters’, and the few cases from groups B and D which had available RNA for sequencing grouped close to groups A and C, respectively. Copy number information on multi-sampled cases showed that metastases shared many similarities with their respective primary tumours but had some significant differences as well. A persistent disease sample showed an identical copy number profile to the primary tumour, as did core-needle biopsies and their respective primary tumours.

General Discussion

***FOS* stabilisation versus *NF2* loss in osteblastoma: two genetic subgroups of the same disease**

In osteblastoma, SVs that stabilise the *FOS* mRNA and FOS protein likely lead to a prolonged effect of the AP-1 transcription factor complex, of which FOS is a part, that regulates the expression of numerous genes^{14, 114}. Since the functional effect of the SVs are believed to influence *FOS* and not its partner genes, the latter are thought to be non-recurrent. This might however not be the case, as a recent study found a *FOS-ANKH* fusion, also reported by Fittal et al.¹⁴, and a *FOS-RUNX2* fusion¹¹⁵. We have reported a *FOS-KIAA1199* fusion similar to the one reported by Fittal et al.¹⁴ One of our *FOS*-rearranged cases might generate a *FOS-RUNX2* fusion, although the specific three-way translocation could not be resolved fully. As not many osteblastomas have been subjected to genomic and/or transcriptomic sequencing, it may be that there indeed are recurrent *FOS* rearrangement patterns that break-apart FISH probes cannot inform us about. Such probes only detect whether *FOS* is rearranged or not, and do not give information about the partner gene/region. Nonetheless, all reported *FOS*-fusions involve the placement of extra nucleotides to the *FOS* reading frame that likely lead to stabilisation on the mRNA and protein level instead of the creation of chimeric proteins.

Of the four cases in our series showing complete loss of *NF2*, three were classified as epithelioid and one was a conventional osteblastoma. None showed rearrangement of *FOS* or *FOSB*. We propose that the phenotypic effect of a *FOS* rearrangement can be achieved in osteblastomas with homozygous loss of *NF2* without rearranging *FOS*. *NF2* loss can namely decrease the activation of the Hippo-signalling pathway and indirectly lead to increased AP-1 activity just like FOS stabilisation might¹¹⁶. Other genes in 22q12 initially suspected to be targets for deletions by Nord et al.¹⁵ were homozygously lost in one or two cases only in our series. This further solidified the hypothesis that *NF2* is the true target of the 22q12 deletions but does not rule out the potential importance of the other homozygously lost genes. More importantly, osteblastoma cases negative for *FOS* or *FOSB* rearrangements, especially those classified as epithelioid osteblastoma, could instead have *NF2* loss, a finding that should be further validated. *NF2* loss may however not be usable as a diagnostic biomarker, as some conventional osteosarcomas also show chromosomal breaks affecting *NF2*, and *NF2*-mutated mice are known to develop osteosarcomas^{24, 30, 117}. This could be a potential explanation as to why some osteblastomas can be morphologically similar to osteoblastic osteosarcoma.

Bone sarcomas are rare, their subtypes even more so

The main issue with basic research into rare tumour types is just that: they are rare and hard to come by. Having access to fresh-frozen biopsy tissue where matched DNA and RNA can be obtained is not a trivial task. Building a large enough series of a certain tumour type is therefore a profound challenge. In our osteoblastoma series, it was not always possible to obtain both DNA and RNA from the same tumour, something that hindered studying *FOS* gene expression levels in some rearranged cases, for example. The perfect example of rare tumour types in this thesis is, however, the chondroblastoma-like osteosarcoma case discussed in **Article II**. Given that this subtype represents less than 1% of conventional osteosarcomas^{84, 85}, which in turn represent less than 0.2% of diagnosed cancers¹⁰, one can appreciate the value of any genetic information we can obtain about these entities. Since a number of other bone tumours overlap morphologically with chondroblastoma-like osteosarcoma⁸⁵, we tested for the presence of defining hot-spot mutations or promoter swapping events seen in other bone tumours to exclude a misdiagnosis of this chondroblastoma-like osteosarcoma^{38, 68, 69, 118-122}. The case under study had none of the defining mutations we screened for.

The *FNI-FGFR1* gene fusion we detected, or in some instances *FNI-FGFI*, are found in around half of reported phosphaturic mesenchymal tumour cases, a rare soft tissue or bone neoplasm^{123, 124}. The chondroblastoma-like osteosarcoma under study did not show features, such as osteomalacia, typical of phosphaturic mesenchymal tumour. However, it cannot be ruled out that the current case is actually a phosphaturic mesenchymal tumour of bone, although with a rare clinical behaviour¹²⁵, as it metastasised to the lungs of the patient. Homozygous loss of *CDKN2A*, known to occur in chondrosarcoma and in some cases of conventional osteosarcoma^{29, 43}, and loss of *DMD* seen in aggressive myogenic sarcomas and sclerosing epithelioid fibrosarcoma^{126, 127}, coupled to the *FNI-FGFR1* fusion finding solidified the argument that case under study seemed to have a unique combination of genetic features amongst bone tumours. Whether these features are unique to chondroblastoma-like osteosarcoma, phosphaturic mesenchymal tumour or present in both is difficult to say as there are no reports on genomic alterations in phosphaturic mesenchymal tumour besides the existence of *FNI-FGFR1* or *FNI-FGFI* gene fusions. The tumour was surgically removed as no signs of disseminated disease were detected at diagnosis. However, a lung metastasis was detected two months after diagnosis and the patient was not responding to chemotherapy as of last follow-up. More knowledge must be gathered about such rare tumours to be able to say whether this was indeed a chondroblastoma-like osteosarcoma or a malignant phosphaturic mesenchymal tumour of bone. Regardless, when a case with similar morphology is encountered in the future and genetically analysed, the findings should be compared with ours.

Targeting genomic alterations as part of future personalised medicine practices: any considerations?

With the advent of NGS techniques, the concept of personalised medicine has been gaining ground^{128, 129}. This has opened up the possibility of mapping the exact genomic alterations a patient's tumour has acquired and attempting to specifically target them in hopes of curing the patient. For example, the *FN1-FGFR1* fusion in the chondroblastoma-like osteosarcoma case can in theory be targeted as multiple TKIs, including dovitinib which can target FGFR1, have shown promising results *in vitro*¹³⁰. However, it is empirical to assess the expression level and functionality of a gene fusion prior to using an inhibitor. For example, the three *NTRK* fusions we have reported in conventional osteosarcomas were deemed non-functional and at best passenger mutations that are a symptom of the extensive genomic rearrangement or catastrophic chromosomal events seen in osteosarcoma. It is extremely unlikely that the *NTRK*-rearranged cases would have responded to an approved pan-NTRK TKI. This highlights the importance of having matched genomic and transcriptomic information, as the existence of a chromosomal rearrangement need not imply a functional chimeric oncoprotein. The most important concept to keep in mind is that when we sequence DNA or RNA from a tumour, we only capture a snapshot of its developmental path and not necessarily the whole story.

Hijacked regulatory elements are a central theme in osteosarcoma

Promoter swapping is by now a well-established phenomenon in neoplasia. It was documented for the first time in human cancers with the discovery of rearrangements of the immunoglobulin loci and the *MYC* gene leading to its deregulation in Burkitt lymphoma, a B-cell malignancy^{82, 131, 132}. Similar findings were seen in some T-cell malignancies where the T-cell receptor loci can be rearranged with a variety of oncogenes leading to their deregulation¹³³⁻¹³⁶. Further research showed that promoter swapping events also occur in, amongst others, the soft tissue tumour dermatofibrosarcoma protuberans upregulating the *PDGFB* gene¹³⁷, in the bone tumour chondromyxoid fibroma upregulating the *GRM1* gene⁶⁸, and in prostate carcinoma where the non-coding regions of the *TMPRSS2* gene are rearranged with a member of the *ETS*-family of transcription factors¹³⁸. Despite the massive amount of rearrangements in osteosarcoma, there have been, until now, no reports demonstrating whether promoter swapping indeed occurs in this disease.

We have shown that not only do SVs affecting *TP53* intron 1 separate the regulatory sequences of the gene from the coding ones and abrogate its expression, they also place either known or putative oncogenes in the vicinity of the relocated *TP53* promoter. In

some instances, these illegitimately recombined sequences are even amplified. One can reason that cells that have completely lost *TP53* will have a constant need to produce more p53 protein as DNA damage will still occur and an attempt to stop the cell cycle and fix the error must be made. We have shown that the *TP53* promoter is indeed active, both *in vivo* and *in vitro*, and can respond to ongoing DNA damage, although the direct consequence would be the upregulation of a potential oncogene. This gene then initiates its own signalling cascade, leading to an erroneous cellular response to the DNA damage instead of an actual *TP53*-mediated response that would normally halt the ongoing cell cycle. We observe *TP53* rearrangement in about 40% of conventional osteosarcoma cases and although the consequences of such rearrangements are not always decipherable, promoter swapping is likely to occur in many of them, especially in the 15% of cases that display ‘*TP53* promoter gain’. Even though none of the identified *TP53*-partners were recurrent, we noted that some of them are putative members of a *TP53*-regulated signalling network^{139, 140}. It could be that the deregulation of the entire *TP53*-mediated network may not be beneficial to tumour cells unless there is a simultaneous rescue of parts of it via a hijacked *TP53* promoter. Importantly, *TP53* promoter swapping was an early event in tumour development and was present in all investigated samples and single cells from affected tumours. Targeting the *TP53* promoter may thereby be a potential new tool to be added to the treatment repertoire of osteosarcoma patients in the future. One may of course raise questions as to how this can be done to begin with and how to specifically target tumour cells to avoid toxic side effects. Delineating an early and recurrent mechanism in osteosarcoma is a much needed first step, nonetheless.

Additionally, our investigations into 12q-amplified low- and high-grade osteosarcomas have also turned up several instances of potential promoter swapping. We describe recurrent rearrangements of the non-coding regions of the *FRS2* and *PLEKHA5* loci that lead to upregulation of a variety of 3' partner genes. These phenomena were preferentially detected in high-grade cases or dedifferentiated cases with a low-grade origin. How or if the *FRS2* and *PLEKHA5* promoter swapping events contribute to disease progression is still unknown. From the nearly 170 osteosarcomas analysed in this thesis, only one gene, *ELF1*, was recurrently identified as a 3' partner gene in a promoter swapping event; first as a partner to the *TP53* promoter in a conventional osteosarcoma and then as a partner to the *FRS2* promoter in a dedifferentiated parosteal osteosarcoma. Further sequencing studies of osteosarcomas will be necessary to find more recurrently involved 3' partner genes, if there are any, or if certain signalling pathways are preferentially deregulated via such rearrangement events.

Interestingly, hijacking of regulatory elements need not be limited to *TP53*, *FRS2* and *PLEKHA5* in osteosarcoma. A recent study has shown that non-randomly clustered

enhancer elements are positively selected for in metastatic osteosarcoma samples compared to primary tumours and between non-metastatic and lung-metastatic osteosarcoma cell lines¹⁴¹. Aberrant enhancer activity has been previously documented in neoplasia¹⁴²⁻¹⁴⁵ and can in some instances even be inhibited to downregulate oncogenes such as *MYC* in multiple myeloma or to prevent metastases of osteosarcoma cell lines^{141, 145}. Enhancer regions can even be co-amplified with the oncogenes they normally regulate, such as those regulating the amplified *EGFR* gene in glioblastoma. They are colocalised with the amplified *EGFR* copies in extrachromosomal circular structures¹⁴⁶, something that could also be likely in 12q-amplified cases harbouring ring or giant rod chromosomes. Given the complexity of the osteosarcoma genome, enhancers are usually not located in their habitual spot on a chromosome but can be relocated and amplified and therefore come to control the expression of numerous other genes that normally would not have been affected by those enhancers. One case in our osteoblastoma series demonstrated a rearrangement translocating the *WNT5A* enhancer, itself induced by *FOS*¹⁴⁷, to the vicinity of the *FOS* gene instead of disrupting the terminal exon. This case showed a positive immunoreactivity for *FOS* protein and enhancer hijacking may be a yet under-investigated phenomenon in osteoblastoma. Further research into enhancer elements and their role in osteosarcoma and other malignancies is warranted as they may be targetable entities whose inhibition may have a profound effect on a tumour's ability to continue surviving and proliferating.

Conclusions

This thesis highlights the presence of novel genetic subgroups of the bone tumours osteoblastoma and osteosarcoma. It also elucidates new genetic mechanisms operating in such tumours. Briefly, the articles included in this thesis demonstrated the following:

- i. Structural variations lead to homozygous loss of *NF2* in non-*FOS*-rearranged osteoblastomas.
- ii. A chondroblastoma-like osteosarcoma case (or malignant phosphaturic mesenchymal tumour of bone) harbours an *FN1-FGFR1* gene fusion and homozygous loss of *CDKN2A* and *DMD*.
- iii. *NTRK* fusions are rare in conventional osteosarcomas and likely non-functional.
- iv. *TP53* promoter rearrangements represent an early-occurring and recurrent gain-of-function mechanism in conventional osteosarcomas and are predominantly found in younger patients.
- v. 12q-amplified osteosarcomas are genetically diverse and the *FRS2* and *PLEKHA5* promoters can be rearranged and induce partner genes in a similar fashion to the *TP53* promoter.

The importance of studying such genetic mechanisms lies in generating a greater understanding of how these tumours form and develop. Given such an increased level of knowledge, we can start to better diagnose tumours and to target them in more clever ways. For instance, loss of *NF2* can be found in both a subgroup of osteoblastoma and in some osteosarcomas. However, just because a suspected osteoblastoma is negative for *FOS/FOSB* rearrangement should not exclude the diagnosis as it may have *NF2* loss instead and still be a locally aggressive osteoblastoma and not a malignant osteosarcoma. Furthermore, while some malignant bone tumours harbour gene fusions that likely can be targeted in the future, others harbour non-functional and thus non-targetable gene fusions. To determine gene fusion functionality, analyses of matched DNA and RNA are crucial.

The delineation of an early and recurrent event in a subset of conventional osteosarcoma represents a crucial step forward in the fight against this cancer. Osteosarcoma patients, if they survive, do so with scars that negatively affect quality of life. Further research into *TP53*-rearranged osteosarcoma is therefore warranted, especially if there is treatment potential. Since promoter swapping seems to be widespread in osteosarcoma, a deeper dive into 12q-amplified tumours to study the effects of relocated *FRS2* and *PLEKHA5* promoters is also warranted. Finally, it is imperative to remember that one gene is never the answer in cancer. A multifaceted out-of-the-box approach is the way forward.

Popular science summary

Our organs and tissues are made up of cells whose proliferation rate is a tightly controlled process. This process, known as the “cell cycle”, makes sure that everything goes according to plan before a cell decides to replicate its hereditary material (DNA) and split into two identical daughter cells. For a tumour to arise, mutations or changes in the DNA are required that either turn off or over-activate certain parts of this cell cycle in an erroneous fashion. The cell loses its inbuilt control checkpoints and starts replicating itself endlessly without regard to its local environment.

DNA is made up of four different “letters” or base-pairs. These are ordered in such a way that they either constitute genes or regulatory sequences, such as promoters and enhancers, that regulate and control the activity levels of our genes. The DNA molecules are in turn organised into structures called chromosomes. Mutations in the DNA can take many forms such as: (i) base-pairs that are altered or lost, (ii) changes in the number of chromosomes that result in the cell having less or more than the normal number and/or (iii) structural rearrangements that physically move parts of the same or different chromosomes and wrongly stitch them together, leading to chromosomal remodelling. In some instances, these rearrangements can lead to the fusion of two normal genes, resulting in a gene fusion with novel characteristics. In this thesis, the genetic mutations that give rise to the osteogenic, or bone-producing, skeletal tumours osteoblastoma and osteosarcoma were investigated using the latest deep sequencing technologies and bioinformatic approaches.

Osteoblastoma is a tumour of bone that can grow into and disrupt its surrounding local environment but does not spread to other parts of the body. It can be cured by surgical removal in most cases. However, it can sometimes be difficult to distinguish between osteoblastoma and osteosarcoma using standard clinical practices. This is crucial as osteosarcoma is a malignant tumour that requires a much more aggressive and comprehensive treatment that has a negative impact on quality of life. Approximately a third of osteosarcoma patients do not survive their disease and no new treatment options that significantly improve survival rates have been introduced during the last three decades. In order to improve diagnostics and treatment of osteoblastoma and

osteosarcoma, a better understanding of the genetic changes that cause these diseases is imperative.

Our investigations uncovered several previously unknown mutations and genetic mechanisms that give rise to osteoblastoma and osteosarcoma. First, we identified a subgroup of osteoblastomas that display complete loss of the *NF2* gene, which is involved in bone growth. Second, we genetically characterised a rare subtype of osteosarcoma and found genetic changes similar to those described in another tumour type. Whether this means that two different tumour types have the same underlying genetics or if the tumours should be classified within the same category is yet unknown. Third, we show that osteosarcomas generally do not harbour functional gene fusions involving the *NTRK* family of genes. This is of importance as some genetic analysis methods can give a false sense of hope that targeting such *NTRK* fusions will be useful during treatment of osteosarcoma. Our findings show that all-encompassing genetic analyses are required to determine whether such targeted treatment will be successful or not in specific cases. Fourth, we encountered a seemingly paradoxical mechanism in mostly younger osteosarcoma patients where the promoter of the *TP53* gene, also known as the guardian of the genome, can promote tumour growth instead of stopping the cell cycle as it normally should. Finally, we show that similar mechanisms occur in osteosarcomas affecting older patients, although the genes involved differ depending on patient age.

Populärvetenskaplig sammanfattning

Våra organ och vävnader består av celler som ständigt förökas och dör genom noggrant kontrollerade processer. Den så kallad cellcykeln ser till att allt går enligt plan innan en cell bestämmer sig för att kopiera sin arvs massa (DNA) och dela sig i två identiska dotterceller. För att en tumör ska uppstå krävs förändringar i DNA som antingen stänger av eller aktiverar vissa delar av cellcykeln på ett felaktigt sätt. Cellen förlorar då sina inbyggda kontrollmekanismer och börjar föröka sig utan hänsyn till sin omgivning.

DNA består av fyra olika "bokstäver" eller baspar. Dessa är ordnade på ett sådant sätt att de antingen utgör gener eller sekvenser som på olika sätt reglerar genernas aktivitet. DNA-molekylerna är i sin tur ordnade i strukturer som kallas kromosomer. Skador på DNA (mutationer) kan ske på många olika sätt, bland annat genom att: (i) enskilda baspar byts ut eller förloras, (ii) antalet kopior av DNA-molekylen förändras så att cellen antingen får fler eller färre kopior än normalt och/eller (iii) kromosomerna går sönder och lagas på ett sådant sätt att delar som normalt inte hör ihop felaktigt läks samman. I vissa fall kan detta leda till att två normala gener slås samman och bildar en så kallad genfusion med helt nya egenskaper. I denna avhandling studerades genetiska förändringar som ger upphov till osteoblastom och osteosarkom, vilka är tumörer som uppstår i skelettet. De studerades med några av de mest högupplösande genetiska analysmetoder som finns tillgängliga idag.

Osteoblastom är en lokalt aggressiv tumör. Det innebär att den förstör skelettet där den växer, men den sprider sig inte vidare till andra delar av kroppen. I de flesta fall botas patienten med kirurgi. Ibland kan det dock vara svårt att skilja osteoblastom från osteosarkom i klinisk diagnostik. Detta är viktigt eftersom osteosarkom är en elakartad tumör som kräver en betydligt mer omfattande behandling. En tredjedel av patienterna dör av sin sjukdom och varken överlevnad eller behandlingsstrategier har förbättrats märkbart under de senaste 30 åren. För att kunna förbättra diagnostik och behandling av osteoblastom och osteosarkom krävs en ökad förståelse för de genetiska förändringarna som orsakar dessa sjukdomar.

Genom våra studier lyckades vi identifiera flera tidigare okända mutationer och genetiska mekanismer som orsakar osteoblastom och osteosarkom. I en undergrupp av osteoblastom hittade vi förlust av *NF2*-genen, vilken bland annat är involverad i

skelettillväxt. I en sällsynt variant av osteosarkom fann vi förändringar som liknar de som beskrivits för en annan tumörtyp. Om detta betyder att två olika tumörtyper har samma bakomliggande genetik eller om tumörerna snarare borde klassificeras inom en och samma kategori kan vi idag inte svara på. Vidare visar vi att osteosarkom sällan eller aldrig har funktionella genfusioner som involverar *NTRK*-familjen av gener. Detta är viktigt att känna till eftersom vissa genetiska analysmetoder kan ge falska förhoppningar om att målstyrd terapi mot just sådana *NTRK*-förändringar skulle kunna vara av nytta vid behandling av osteosarkom. Enligt våra fynd krävs det omfattande analyser för att avgöra om sådan behandling kan vara framgångsrik i enskilda fall. Därtill har vi identifierat en ny mekanism där osteosarkom, framförallt hos unga individer, utnyttjar specifika DNA-sekvenser som normalt ska förhindra tumörutveckling till att istället främja sådan. Slutligen visar vi att liknande mekanismer förekommer även i osteosarkom som drabbar äldre, men de gener som är involverade skiljer sig mellan osteosarkom hos unga respektive äldre patienter.

Acknowledgements

I would like to begin by thanking everyone who actually read the whole thesis first and did not directly skip to the acknowledgements section. But on a more serious tone, I am grateful for the technical, scientific and emotional support I received during the past five years at the Division of Clinical Genetics! In particular, I would like to acknowledge the following people:

My supervisor, **Karolin Hansén Nord**. I am beyond grateful that I had the chance to pursue my PhD in your group, working with some of the most intriguingly messed up tumours there are! I am thankful for all the expertise, the advice, the “I am just going to go over and tell her something I found” meetings that end up being life lessons. You have taught me to always think outside the box, be critical of what I read, that people have different opinions on where the word *also* should be in a sentence, and also that more than five scoops of coffee powder is not considered a normal amount. I will make sure to continue documenting my future research on hundreds of post-it notes as we have over the past four years!

My co-supervisors, **Emelie Styring** and **Johan Staaf**, for sharing their knowledge and advice with me. Specifically, I thank Emelie for all the hours we have spent looking up patient data and the reminders that what we do is for the benefit of the patients at the end of the day. I would like to thank Johan for the bioinformatic tips and tricks I learned.

My **co-authors** and **collaborators**, thank you for sharing your expertise, samples and knowledge for the common goals we have. This thesis would not have been possible without you.

Our previous post-doc **Louise Cornmark**. I still look at the little piece of paper you made after my half-time defence that explains how our favourite genes actually interact with each other when I need the reminder!

Our new PhD student **Valeria Difilippo**. I wish you all the luck in your studies and I am looking forward to learning some more bioinformatics with you!

JennyLinda, or is it **LindaJenny**? No one really knows... You are our group's structural foundation. I am beyond thankful for all the jokes, DNA extractions, library prepping,

and all the other lab stuff you do! I am also grateful for the skånska/Lomma words I have learned (and promptly forgotten) and for all the absurd conversations we had around the fika table at exactly 09:30 every morning. Linda, I promise never to forget the DNA when transforming bacteria! Jenny, I promise never to ask you to start up GenomeStudio for me ever again! This thesis would have been impossible without both of you!

The ‘gods of cytogenetics’ as they are known amongst some of us, **Felix Mitelman** and **Nils Mandahl**. Thank you for still being engaged in research after all these years! I specifically thank Felix for checking that the three-way translocations I wrote were actually correct, and Nisse for always being interested in what the sarcomas are up to at our Wednesday meetings. It gives a sense of accomplishment to receive positive feedback from both of you!

The PIs **David Gisselsson Nord**, **Fredrik Mertens**, **Thoas Fioretos**, **Marcus Järås**, **Anna Hagström**, **Kajsa Paulsson** and **Bertil Johansson** for providing us with a creative and successful working environment with some of the best people you could ever meet!

Anette Welin, thank you for all the help navigating the bureaucracy that is Lund University!

Marianne, the all-knowing, good thing you predicted all those years ago that I will stay at Clinical Genetics for a long time, you were of course correct. A special thanks to **Geoffroy**, **Carro**, **Andrea** and **Helena S** too!

All current and previous PhD students and post-docs for all the best discussions, group emotional breakdowns, and the best (pre-COVID-19) afterworks on the planet! A special thanks to my unofficial C13 parents **Jakob** and **Elsa** for introducing me to all the best sarcoma/genetics conferences to go to. The Hofvander family should reunite at that AirBnB in Maui one sunny post-corona day! I would like to thank **Hanna**, **Louise** and **Natalie** for the best roommates one can ask for! I wish you all good luck with your PhD studies! Thanks to **Larissa** and **Katrin** for all the banana breaks, waving through the door window, weird looks and the potential diagnoses people have given us for the way we simply are! A shout out to the best BBQ-organiser **Ton** and to **Giuseppe**, never have I seen anyone take what goes on pasta so personally. A thanks to **Saskia**, **Naveen**, **Ram**, **Hiroaki**, **Mattias**, **Marion**, **Somadri**, **Shamik**, **Bahar**, **Sofia**, **Niklas**, **Mia**, **Axel**, **Anders**, **Maria**, **Pablo**, **Noelia**, **Qirui**, **Jan**, **Anna**, **Gustav**, and **Subhayan** is also in order!

We are starting to become quite many at C13, but if your picture is or was on the wall, then a thank you is directed your way! Thanks to **Linda H** and **Jenny K** (not to be

confused with JennyLinda or LindaJenny), **Stina, Rebeqa, Kishan, Minjun, Helena Å, Calle, Christina, Henrik, Ellie, Rasmus and Catharina.**

The protein people **Mamma Rita, Oktawia, Sibel, Hanna, Niloofar, Raminta, Ping, Peng, Julia and Charlie,** you guys are great too!

My family back home for all their support throughout the years! Words cannot describe how grateful I am for my parents **Hala and Hussam** and my sister **Yara.** You made sure that nothing stood in the way of my dreams and were always supportive of my decisions! Thank you for all the times we've been through as a family and for the coffee on the balcony watching random people on the street (this is more interesting than it sounds once you factor in the chaos of a big city on the Mediterranean). This thesis is dedicated to all of you. A special thanks to the rest of the family, including all the **cousins, aunts, uncles and grandparents!**

Even though you cannot read this, I want to thank my favourite anti-depressant **Zeytoun** for always being happy and wagging his tail when I arrive from the airport. I am looking forward to meeting the new dog **Figla!** Thanks to the kitties **Milo and Maya** for being the special little furballs they are!

My friends in Lebanon (not that most of you live there anymore). A shout out to **Hussam, Joanna and Nadine** for the best years at AUB and being a constant reminder that pre-med is more stressful than the chill research option. **May and Elissa,** thanks for the best times wherever in Europe we were and back in Hamra while having beer and catching up. As for my friends in Sweden, a thanks to my boys **Oskar, Jimmie, André, Erik and Hannes** for all the good (not always remembered) times! My travel buddy **Kate,** looking forward for the next brewery we are hitting up! Thanks to **Mimmi** and **Emma** for all the good times back at the nation! Thanks to **Ludivine** for letting me crash at her place whenever I needed to do embassy stuff in Stockholm! Thanks to **Cassy** for the once-a-year catch ups that take a year to plan! A shout out to the **Kämmnärs 71E** and **Kristianstads 2S** corridors for all the good times and memories!

Lexi. You are the reason I have not gone crazy(ier) the past few months. Thank you for always being there, through the good and the bad. Thank you for being my weirdo. I look forward to seeing your life and professional career blossom.

Finally, I would like to thank my opponent **Adrienne M. Flanagan** and the examination board members **Gustaf Ljungman, Björn Nilsson and Catarina Lundin** for reviewing and assessing my thesis. Thank you to **Kajsa Paulsson** for agreeing to chair my defence.

This work was supported by the **Swedish Childhood Cancer Fund**, the **Swedish Cancer Society**, the **Swedish Research Council**, the **Faculty of Medicine at Lund University**, the **Åke Wiberg Foundation**, the **Royal Physiographic Society** (Lund, Sweden), and the **Crafoord Foundation**.

For members of the public that may find this just a bit too much, and in hopes of continuing an idea and maybe starting a tradition, please enjoy this (mostly) genetics-themed word search. Good luck!



Words are hidden vertically, horizontally and diagonally.

AMPLIFICATION

MALIGNANT

CHROMOSOME

MITOSIS

CHROMOTHRIPSIS

MUTATION

DELETION

OSTEOBLASTOMA

DNA

OSTEOSARCOMA

DUPLICATION

PROMOTER

HOMOZYGOUS

TRANSLOCATION

INVERSION

TUMOUR

References

1. Boveri T. (1914) Zur Frage der Entstehung Maligner Tumoren (Gustav Fischer Verlag, Jena).
2. Avery OT, Macleod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of Pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. *J Exp Med.* 1944; **79**: 137-158.
3. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953; **171**: 737-738.
4. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A.* 1971; **68**: 820-823.
5. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature.* 2009; **458**: 719-724.
6. Baca SC, Prandi D, Lawrence MS, et al. Punctuated evolution of prostate cancer genomes. *Cell.* 2013; **153**: 666-677.
7. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell.* 2011; **144**: 27-40.
8. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000; **100**: 57-70.
9. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; **144**: 646-674.
10. Howlader N, Noone AM, Krapcho M, et al. (eds). SEER Cancer Statistics Review, 1975-2017, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2017/, based on November 2019 SEER data submission, posted to the SEER web site, April 2020.
11. *Soft tissue and bone tumours* (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020.
12. Flanagan AM, Blay JY, Bovée JVMG, et al. Bone tumours: Introduction. In *Soft tissue and bone tumours*. (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 340-344.

13. Amary F, Bredella MA, Horvai AE, et al. Osteoblastoma. In *Soft tissue and bone tumours*. (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 397-399.
14. Fittall MW, Mifsud W, Pillay N, et al. Recurrent rearrangements of *FOS* and *FOSB* define osteoblastoma. *Nat Commun*. 2018; 9: 2150.
15. Nord KH, Nilsson J, Arbajian E, et al. Recurrent chromosome 22 deletions in osteoblastoma affect inhibitors of the Wnt/beta-catenin signaling pathway. *PLoS one*. 2013; 8: e80725.
16. Gianferante DM, Mirabello L, Savage SA. Germline and somatic genetics of osteosarcoma - connecting aetiology, biology and therapy. *Nat Rev Endocrinol*. 2017; 13: 480-491.
17. Bone sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2014; 25 Suppl 3: iii113-123.
18. Goorin AM, Schwartzentruber DJ, Devidas M, et al. Presurgical chemotherapy compared with immediate surgery and adjuvant chemotherapy for nonmetastatic osteosarcoma: Pediatric Oncology Group Study POG-8651. *J Clin Oncol*. 2003; 21: 1574-1580.
19. Kempf-Bielack B, Bielack SS, Jürgens H, et al. Osteosarcoma relapse after combined modality therapy: an analysis of unselected patients in the Cooperative Osteosarcoma Study Group (COSS). *J Clin Oncol*. 2005; 23: 559-568.
20. Meyers PA, Heller G, Healey J, et al. Chemotherapy for nonmetastatic osteogenic sarcoma: the Memorial Sloan-Kettering experience. *J Clin Oncol*. 1992; 10: 5-15.
21. Baumhoer D, Böhling TO, Cates JMM, et al. Osteosarcoma. In *Soft tissue and bone tumours*. (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 403-409.
22. Janeway KA, Grier HE. Sequelae of osteosarcoma medical therapy: a review of rare acute toxicities and late effects. *Lancet Oncol*. 2010; 11: 670-678.
23. Whelan JS, Bielack SS, Marina N, et al. EURAMOS-1, an international randomised study for osteosarcoma: results from pre-randomisation treatment. *Ann Oncol*. 2015; 26: 407-414.
24. Mertens F, Mandahl N, Orndal C, et al. Cytogenetic findings in 33 osteosarcomas. *Int J Cancer*. 1993; 55: 44-50.
25. Fletcher JA, Gebhardt MC, Kozakewich HP. Cytogenetic aberrations in osteosarcomas. Nonrandom deletions, rings, and double-minute chromosomes. *Cancer Genet Cytogenet*. 1994; 77: 81-88.

26. Behjati S, Tarpey PS, Haase K, et al. Recurrent mutation of *IGF* signalling genes and distinct patterns of genomic rearrangement in osteosarcoma. *Nat Commun.* 2017; **8**: 15936-15936.
27. Chen X, Bahrami A, Pappo A, et al. Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. *Cell reports.* 2014; **7**: 104-112.
28. Kovac M, Blattmann C, Ribi S, et al. Exome sequencing of osteosarcoma reveals mutation signatures reminiscent of *BRCA* deficiency. *Nat Commun.* 2015; **6**: 8940-8940.
29. Lorenz S, Barøy T, Sun J, et al. Unscrambling the genomic chaos of osteosarcoma reveals extensive transcript fusion, recurrent rearrangements and frequent novel *TP53* aberrations. *Oncotarget.* 2016; **7**: 5273-5288.
30. Smida J, Xu H, Zhang Y, et al. Genome-wide analysis of somatic copy number alterations and chromosomal breakages in osteosarcoma. *Int J Cancer.* 2017; **141**: 816-828.
31. Negri GL, Grande BM, Delaidelli A, et al. Integrative genomic analysis of matched primary and metastatic pediatric osteosarcoma. *J Pathol.* 2019; **249**: 319-331.
32. Ribi S, Baumhoer D, Lee K, et al. *TP53* intron 1 hotspot rearrangements are specific to sporadic osteosarcoma and can cause Li-Fraumeni syndrome. *Oncotarget.* 2015; **6**: 7727-7740.
33. Wang J, Nord KH, O'Donnell PG, et al. Parosteal osteosarcoma. In *Soft tissue and bone tumours.* (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 410-413.
34. Yoshida A, Bredella MA, Gambarotti M, et al. Low-grade central osteosarcoma. In *Soft tissue and bone tumours.* (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 400-402.
35. Bertoni F, Bacchini P, Staals EL, et al. Dedifferentiated parosteal osteosarcoma: the experience of the Rizzoli Institute. *Cancer.* 2005; **103**: 2373-2382.
36. Choong PF, Pritchard DJ, Rock MG, et al. Low grade central osteogenic sarcoma. A long-term followup of 20 patients. *Clin Orthop Relat Res.* 1996; 198-206.
37. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet.* 2018; **59**: 253-268.
38. Behjati S, Tarpey PS, Presneau N, et al. Distinct *H3F3A* and *H3F3B* driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet.* 2013; **45**: 1479-1482.
39. Killela PJ, Reitman ZJ, Jiao Y, et al. *TERT* promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc Natl Acad Sci U S A.* 2013; **110**: 6021-6026.

40. Lin Y, Seger N, Chen Y, et al. *hTERT* promoter mutations in chondrosarcomas associate with progression and disease-related mortality. *Mod Pathol.* 2018; **31**: 1834-1841.
41. Horn S, Figl A, Rachakonda PS, et al. *TERT* promoter mutations in familial and sporadic melanoma. *Science.* 2013; **339**: 959-961.
42. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med.* 2017; **23**: 703-713.
43. Hallor KH, Staaf J, Bovée JVMG, et al. Genomic profiling of chondrosarcoma: chromosomal patterns in central and peripheral tumors. *Clin Cancer Res.* 2009; **15**: 2685-2694.
44. Mohseny AB, Szuhai K, Romeo S, et al. Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of *Cdkn2*. *J Pathol.* 2009; **219**: 294-305.
45. Perry JA, Kiezun A, Tonzi P, et al. Complementary genomic approaches highlight the PI3K/mTOR pathway as a common vulnerability in osteosarcoma. *Proc Natl Acad Sci U S A.* 2014; **111**: E5564-5573.
46. Dujardin F, Binh MB, Bouvier C, et al. MDM2 and CDK4 immunohistochemistry is a valuable tool in the differential diagnosis of low-grade osteosarcomas and other primary fibro-osseous lesions of the bone. *Mod Pathol.* 2011; **24**: 624-637.
47. Gisselsson D, Pålsson E, Höglund M, et al. Differentially amplified chromosome 12 sequences in low- and high-grade osteosarcoma. *Genes Chromosomes Cancer.* 2002; **33**: 133-140.
48. Heidenblad M, Hallor KH, Staaf J, et al. Genomic profiling of bone and soft tissue tumors with supernumerary ring chromosomes using tiling resolution bacterial artificial chromosome microarrays. *Oncogene.* 2006; **25**: 7106-7116.
49. Mejia-Guerrero S, Quejada M, Gokgoz N, et al. Characterization of the 12q15 *MDM2* and 12q13-14 *CDK4* amplicons and clinical correlations in osteosarcoma. *Genes Chromosomes Cancer.* 2010; **49**: 518-525.
50. Pedetour F, Forus A, Coindre JM, et al. Structure of the supernumerary ring and giant rod chromosomes in adipose tissue tumors. *Genes Chromosomes Cancer.* 1999; **24**: 30-41.
51. Pedetour F, Suijkerbuijk RF, Forus A, et al. Complex composition and co-amplification of *SAS* and *MDM2* in ring and giant rod marker chromosomes in well-differentiated liposarcoma. *Genes Chromosomes Cancer.* 1994; **10**: 85-94.

52. Szymanska J, Mandahl N, Mertens F, et al. Ring chromosomes in parosteal osteosarcoma contain sequences from 12q13-15: a combined cytogenetic and comparative genomic hybridization study. *Genes Chromosomes Cancer*. 1996; **16**: 31-34.
53. Lau CC, Harris CP, Lu XY, et al. Frequent amplification and rearrangement of chromosomal bands 6p12-p21 and 17p11.2 in osteosarcoma. *Genes Chromosomes Cancer*. 2004; **39**: 11-21.
54. Henriksen J, Aagesen TH, Maeldandsmo GM, et al. Amplification and overexpression of *COPS3* in osteosarcomas potentially target *TP53* for proteasome-mediated degradation. *Oncogene*. 2003; **22**: 5358-5361.
55. Squire JA, Pei J, Marrano P, et al. High-resolution mapping of amplifications and deletions in pediatric osteosarcoma by use of CGH analysis of cDNA microarrays. *Genes Chromosomes Cancer*. 2003; **38**: 215-225.
56. Man TK, Lu XY, Jaeweon K, et al. Genome-wide array comparative genomic hybridization analysis reveals distinct amplifications in osteosarcoma. *BMC Cancer*. 2004; **4**: 45.
57. McClintock B. *Genetics*. 1938; **23**: 215-376.
58. McClintock B. *Genetics*. 1940; **26**: 234-282.
59. Gisselsson D, Pettersson L, Höglund M, et al. Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. *Proc Natl Acad Sci U S A*. 2000; **97**: 5357-5362.
60. Korbel JO, Campbell PJ. Criteria for inference of chromothripsis in cancer genomes. *Cell*. 2013; **152**: 1226-1236.
61. Cortés-Ciriano I, Lee JJ, Xi R, et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nat Genet*. 2020; **52**: 331-341.
62. Anderson ND, de Borja R, Young MD, et al. Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors. *Science*. 2018; **361**.
63. Nowell PC, Hungerford DA. A minute chromosome in human chronic granulocytic leukemia. *Science*. 1960; **132**: 1497.
64. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*. 1973; **243**: 290-293.
65. Stam K, Heisterkamp N, Grosveld G, et al. Evidence of a new chimeric *bcr/c-abl* mRNA in patients with chronic myelocytic leukemia and the Philadelphia chromosome. *N Engl J Med*. 1985; **313**: 1429-1433.
66. Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2020). Mitelman F, Johansson B and Mertens F (Eds.), <https://mitelmandatabase.isb-cgc.org>

67. de Álava E, Lessnick SL, and Stamenkovic I. Ewing sarcoma. In *Soft tissue and bone tumours*. (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 322-325.
68. Nord KH, Lilljebjörn H, Vezzi F, et al. *GRM1* is upregulated through gene fusion and promoter swapping in chondromyxoid fibroma. *Nat Genet*. 2014; **46**: 474-477.
69. Oliveira AM, Hsi BL, Weremowicz S, et al. *USP6 (Tre2)* fusion oncogenes in aneurysmal bone cyst. *Cancer Res*. 2004; **64**: 1920-1923.
70. Siebert, R & Aukema, SM. In *Cancer cytogenetics: chromosomal and molecular genetic aberrations of tumor cells* (4th edn), edited by Heim S, Mitelman F. John Wiley & Sons, Ltd: 2015; 252-331.
71. Heuckmann JM, Pauwels P, Thunnissen E. Comprehensive Hybrid Capture-Based Next-Generation Sequencing Identifies a Double *ALK* Gene Fusion in a Patient Previously Identified to Be False-Negative by FISH. *J Thorac Oncol*. 2017; **12**: e22-e24.
72. Lim G, Cho EH, Cho SY, et al. A novel *PML-ADAMTS17-RARA* gene rearrangement in a patient with pregnancy-related acute promyelocytic leukemia. *Leuk Res*. 2011; **35**: e106-110.
73. McCarron SL, Kelly J, Coen N, et al. A novel e8a2 *BCR-ABL1* fusion with insertion of *RALGPS1* exon 8 in a patient with relapsed Philadelphia chromosome-positive acute lymphoblastic leukemia. *Leuk Lymphoma*. 2011; **52**: 919-921.
74. Pflueger D, Terry S, Sboner A, et al. Discovery of non-*ETS* gene fusions in human prostate cancer using next-generation RNA sequencing. *Genome Res*. 2011; **21**: 56-67.
75. Walz C, Grimwade D, Saussele S, et al. Atypical mRNA fusions in *PML-RARA* positive, *RARA-PML* negative acute promyelocytic leukemia. *Genes Chromosomes Cancer*. 2010; **49**: 471-479.
76. van Ijzendoorn DGP, Forghany Z, Liebelt F, et al. Functional analyses of a human vascular tumor FOS variant identify a novel degradation mechanism and a link to tumorigenesis. *J Biol Chem*. 2017; **292**: 21282-21290.
77. Amary F, Markert E, Berisha F, et al. FOS Expression in Osteoid Osteoma and Osteoblastoma: A Valuable Ancillary Diagnostic Tool. *Am J Surg Pathol*. 2019; **43**: 1661-1667.
78. Lam SW, Cleven AHG, Kroon HM, et al. Utility of FOS as diagnostic marker for osteoid osteoma and osteoblastoma. *Virchows Arch*. 2020; **476**: 455-463
79. Lucas DR, Unni KK, McLeod RA, et al. Osteoblastoma: clinicopathologic study of 306 cases. *Hum Pathol*. 1994; **25**: 117-134.
80. Malkin D, Li FP, Strong LC, et al. Germ line *p53* mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*. 1990; **250**: 1233-1238.

81. Ognjanovic S, Olivier M, Bergemann TL, et al. Sarcomas in *TP53* germline mutation carriers: a review of the IARC *TP53* database. *Cancer*. 2012; **118**: 1387-1396.
82. Draper GJ, Sanders BM, Kingston JE. Second primary neoplasms in patients with retinoblastoma. *Br J Cancer*. 1986; **53**: 661-671.
83. Martin JW, Yoshimoto M, Ludkovski O, et al. Analysis of segmental duplications, mouse genome synteny and recurrent cancer-associated amplicons in human chromosome 6p21-p12. *Cytogenet Genome Res*. 2010; **128**: 199-213.
84. Aycan OE, Vanel D, Righi A, et al. Chondroblastoma-like osteosarcoma: a case report and review. *Skeletal Radiol*. 2015; **44**: 869-873.
85. Hmada YA, Bernieh A, Morris RW, et al. Chondroblastoma-like Osteosarcoma. *Arch Pathol Lab Med*. 2020; **144**: 15-17.
86. Cocco E, Scaltriti M, Drilon A. *NTRK* fusion-positive cancers and TRK inhibitor therapy. *Nat Rev Clin Oncol*. 2018; **15**: 731-747.
87. Pulciani S, Santos E, Lauver AV, et al. Oncogenes in solid human tumours. *Nature*. 1982; **300**: 539-542.
88. Vaishnavi A, Le AT, Doebele RC. TRKking down an old oncogene in a new era of targeted therapy. *Cancer Discov*. 2015; **5**: 25-34.
89. Penault-Llorca F, Rudzinski ER, Sepulveda AR. Testing algorithm for identification of patients with TRK fusion cancer. *J Clin Pathol*. 2019; **72**: 460-467.
90. Drilon A, Laetsch TW, Kummar S, et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N Engl J Med*. 2018; **378**: 731-739.
91. Drilon A, Siena S, Ou SI, et al. Safety and Antitumor Activity of the Multitargeted Pan-TRK, ROS1, and ALK Inhibitor Entrectinib: Combined Results from Two Phase I Trials (ALKA-372-001 and STARTRK-1). *Cancer Discov*. 2017; **7**: 400-409.
92. Masuda H, Miller C, Koeffler HP, et al. Rearrangement of the *p53* gene in human osteogenic sarcomas. *Proc Natl Acad Sci U S A*. 1987; **84**: 7716-7719.
93. Miller CW, Aslo A, Tsay C, et al. Frequency and structure of *p53* rearrangements in human osteosarcoma. *Cancer Res*. 1990; **50**: 7950-7954.
94. Salinas-Souza C, De Andrea C, Bihl M, et al. *GNAS* mutations are not detected in parosteal and low-grade central osteosarcomas. *Mod Pathol*. 2015; **28**: 1336-1342.
95. Sanger F, Air GM, Barrell BG, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*. 1977; **265**: 687-695.
96. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977; **74**: 5463-5467.

97. Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. *Nature*. 2015; **526**: 68-74.
98. van den Bos H, Bakker B, Taudt A, et al. Quantification of Aneuploidy in Mammalian Systems. *Method Mol Biol*. 2019; **1896**: 159-190.
99. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; **29**: 15-21.
100. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010; **28**: 511-515.
101. Maher CA, Kumar-Sinha C, Cao X, et al. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009; **458**: 97-101.
102. Maher CA, Palanisamy N, Brenner JC, et al. Chimeric transcript discovery by paired-end transcriptome sequencing. *Proc Natl Acad Sci U S A*. 2009; **106**: 12353-12358.
103. Nicorici D, Satalan M, Edgren H, et al. FusionCatcher – a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014: 011650.
104. Haas BJ, Dobin A, Li B, et al. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol*. 2019; **20**: 213.
105. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012; **22**: 568-576.
106. Li H. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*. 2014; **30**: 2843-2851.
107. Eisfeldt J, Vezzi F, Olason P, et al. TIDDIT, an efficient and comprehensive structural variant caller for massive parallel sequencing data. *F1000Research*. 2017; **6**: 664.
108. Rausch T, Zichner T, Schlattl A, et al. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012; **28**: i333-i339.
109. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2016; **32**: 1220-1222.
110. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013; **14**: 178-192.
111. Talevich E, Shain AH, Botton T, et al. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol*. 2016; **12**: e1004873.

112. Bakker B, Taudt A, Belderbos ME, et al. Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. *Genome Biol.* 2016; **17**: 115.
113. Meyers PA, Schwartz CL, Krailo M, et al. Osteosarcoma: a randomized, prospective trial of the addition of ifosfamide and/or muramyl tripeptide to cisplatin, doxorubicin, and high-dose methotrexate. *J Clin Oncol.* 2005; **23**: 2004-2011.
114. Curran T, Franza BR, Jr. Fos and Jun: the AP-1 connection. *Cell.* 1988; **55**: 395-397.
115. Panagopoulos I, Gorunova L, Lobmaier I, et al. *FOS-ANKH* and *FOS-RUNX2* Fusion Genes in Osteoblastoma. *Cancer Genomics Proteomics.* 2020; **17**: 161-168.
116. Yang W, Han W, Qin A, et al. The emerging role of Hippo signaling pathway in regulating osteoclast formation. *J Cell Physiol.* 2018; **233**: 4606-4617.
117. McClatchey AL, Saotome I, Mercer K, et al. Mice heterozygous for a mutation at the *Nf2* tumor suppressor locus develop a range of highly metastatic tumors. *Genes Dev.* 1998; **12**: 1121-1133.
118. Amary MF, Bacsı K, Maggiani F, et al. *IDH1* and *IDH2* mutations are frequent events in central chondrosarcoma and central and periosteal chondromas but not in other mesenchymal tumours. *J Pathol.* 2011; **224**: 334-343.
119. Badalian-Very G, Vergilio JA, Degar BA, et al. Recurrent *BRAF* mutations in Langerhans cell histiocytosis. *Blood.* 2010; **116**: 1919-1923.
120. Kurek KC, Pansuriya TC, van Ruler MA, et al. R132C *IDH1* mutations are found in spindle cell hemangiomas and not in other vascular tumors or malformations. *Am J Pathol.* 2013; **182**: 1494-1500.
121. Lee SE, Lee EH, Park H, et al. The diagnostic utility of the *GNAS* mutation in patients with fibrous dysplasia: meta-analysis of 168 sporadic cases. *Hum Pathol.* 2012; **43**: 1234-1242.
122. Pansuriya TC, van Eijk R, d'Adamo P, et al. Somatic mosaic *IDH1* and *IDH2* mutations are associated with enchondroma and spindle cell hemangioma in Ollier disease and Maffucci syndrome. *Nat Genet.* 2011; **43**: 1256-1261.
123. Lee JC, Jeng YM, Su SY, et al. Identification of a novel *FN1-FGFR1* genetic fusion as a frequent event in phosphaturic mesenchymal tumour. *J Pathol.* 2015; **235**: 539-545.
124. Lee JC, Su SY, Changou CA, et al. Characterization of *FN1-FGFR1* and novel *FN1-FGF1* fusion genes in a large series of phosphaturic mesenchymal tumors. *Mod Pathol.* 2016; **29**: 1335-1346.
125. Lee JC and Folpe AL. Phosphaturic mesenchymal tumour. In *Soft tissue and bone tumours*. (5th edn), WHO Classification of Tumours Editorial Board. International Agency for Research on Cancer (IARC): Lyon, 2020; 284-286.

126. Arbajian E, Puls F, Antonescu CR, et al. In-depth Genetic Analysis of Sclerosing Epithelioid Fibrosarcoma Reveals Recurrent Genomic Alterations and Potential Treatment Targets. *Clin Cancer Res.* 2017; **23**: 7426-7434.
127. Wang Y, Marino-Enriquez A, Bennett RR, et al. Dystrophin is a tumor suppressor in human cancers with myogenic programs. *Nat Genet.* 2014; **46**: 601-606.
128. Morganti S, Tarantino P, Ferraro E, et al. Next Generation Sequencing (NGS): A Revolutionary Technology in Pharmacogenomics and Personalized Medicine in Cancer. *Adv Exp Med Biol.* 2019; **1168**: 9-30.
129. Nakagawa H, Fujita M. Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer Sci.* 2018; **109**: 513-522.
130. Landberg N, Dreimane A, Rissler M, et al. Primary cells in *BCR/FGFR1*-positive 8p11 myeloproliferative syndrome are sensitive to dovitinib, ponatinib, and dasatinib. *Eur J Haematol.* 2017; **99**: 442-448.
131. Leder P, Battey J, Lenoir G, et al. Translocations among antibody genes in human cancer. *Science.* 1983; **222**: 765-771.
132. Croce CM, Nowell PC. Molecular basis of human B cell neoplasia. *Blood.* 1985; **65**: 1-7.
133. Erikson J, Finger L, Sun L, et al. Deregulation of *c-myc* by translocation of the alpha-locus of the T-cell receptor in T-cell leukemias. *Science.* 1986; **232**: 884-886.
134. Hayashi Y, Yamamoto K, Kojima S. T-cell acute lymphoblastic leukemias with a t(8;14) possibly involving a *c-myc* locus and T-cell-receptor alpha-chain genes. *N Engl J Med.* 1986; **314**: 650-651.
135. Mathieu-Mahul D, Sigaux F, Zhu C, et al. A t(8;14)(q24;q11) translocation in a T-cell leukemia (L1-ALL) with *c-myc* and TcR-alpha chain locus rearrangements. *Int J Cancer.* 1986; **38**: 835-840.
136. McKeithan TW, Shima EA, Le Beau MM, et al. Molecular cloning of the breakpoint junction of a human chromosomal 8;14 translocation involving the T-cell receptor alpha-chain gene and sequences on the 3' side of *MYC*. *Proc Natl Acad Sci U S A.* 1986; **83**: 6636-6640.
137. Simon MP, Pedeutour F, Sirvent N, et al. Deregulation of the platelet-derived growth factor B-chain gene via fusion with collagen gene *COL1A1* in dermatofibrosarcoma protuberans and giant-cell fibroblastoma. *Nat Genet.* 1997; **15**: 95-98.
138. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of *TMPRSS2* and *ETS* transcription factor genes in prostate cancer. *Science.* 2005; **310**: 644-648.
139. Li Y, Liu J, McLaughlin N, et al. Genome-wide analysis of the *p53* gene regulatory network in the developing mouse kidney. *Physiol Genomics.* 2013; **45**: 948-964.

140. Miled C, Pontoglio M, Garbay S, et al. A genomic map of p53 binding sites identifies novel p53 targets involved in an apoptotic network. *Cancer Res.* 2005; **65**: 5096-5104.
141. Morrow JJ, Bayles I, Funnell APW, et al. Positively selected enhancer elements endow osteosarcoma cells with metastatic competence. *Nat Med.* 2018; **24**: 176-185.
142. Akhtar-Zaidi B, Cowper-Sal-lari R, Corradin O, et al. Epigenomic enhancer profiling defines a signature of colon cancer. *Science.* 2012; **336**: 736-739.
143. Cohen AJ, Saiakhova A, Corradin O, et al. Hotspots of aberrant enhancer activity punctuate the colorectal cancer epigenome. *Nat Commun.* 2017; **8**: 14400.
144. Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease. *Cell.* 2013; **155**: 934-947.
145. Lovén J, Hoke HA, Lin CY, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell.* 2013; **153**: 320-334.
146. Morton AR, Dogan-Artun N, Faber ZJ, et al. Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell.* 2019; **179**: 1330-1341.e1313.
147. Zhong Z, Ethen NJ, Williams BO. WNT signaling in bone development and homeostasis. *Wiley Interdiscip Rev Dev Biol.* 2014; **3**: 489-500.