

Universiteit
Leiden
The Netherlands

A systems biology approach to study high-grade osteosarcoma

Kuijjer, M.L.

Citation
Kuijjer, M. L. (2013, June 26). *A systems biology approach to study high-grade osteosarcoma*. Retrieved from https://hdl.handle.net/1887/21043

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/21043

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle http://hdl.handle.net/1887/21043 holds various files of this Leiden University dissertation.

Author: Kuijjer, Marieke Lydia
Title: A systems biology approach to study high-grade osteosarcoma
Issue Date: 2013-06-26

A Systems Biology Approach to Study High-grade Osteosarcoma

The work presented in this thesis was financially supported by the Dutch Cancer Society (2008-4060) and EuroBoNeT, a European Commission granted Network of Excellence for studying the pathology and genetics of bone tumors (LSHC-CT-2006-018814).

Publication of this thesis was financially supported by the Dutch Cancer Society and by the Department of Pathology, Leiden University Medical Center.

Cover art: M.L. Kuijjer, photographs of the contemporary artwork 'Ritterschlacht' by Martin Honert, taken at the exhibition 'Kinderkreuzzug', Hamburger Bahnhof, Berlin, Germany.

Printed by: Wöhrmann Print Service, Zutphen, the Netherlands.

A Systems Biology Approach to Study High-grade Osteosarcoma

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 26 juni 2013
klokke 16.15 uur

door

Marieke Lydia Kuijjer
geboren te Zaanstad
in 1982

Promotiecommissie

Promotor:	Prof. dr. P.C.W. Hogendoorn
Co-promotor:	Dr. A.M. Cleton-Jansen
Overige leden:	Prof. dr. C.J. Cornelisse
	Prof. dr. J. Kirpensteijn (Utrecht University, Utrecht)
	Prof. dr. O. Myklebost (Norwegian Radium Hospital, Oslo, Norway)
	Dr. J.M. Boer (Erasmus Medical Center, Rotterdam)

"Qui o altrove, tenente, siamo tutti da qualche parte per sbaglio"

Il deserto dei Tartari, Dino Buzzati / Valerio Zurlini

Contents

1 General introduction 9

2 Genome-wide analyses on high-grade osteosarcoma: making sense of a genomically most unstable tumor 29

3 mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts 53

4 Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents 71

5 IR/IGF1R signaling as potential target for treatment of high-grade osteosarcoma 89

6 Kinome and mRNA expression profiling of osteosarcoma identifies genomic instability, and reveals Akt as potential target for treatment 105

7 Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data 127

8 Frequent loss of heterozygosity and amplification in high-grade osteosarcoma: analysis of recurrent tumor suppressor genes 147

9 Concluding remarks and future perspectives 165

10 Nederlandse samenvatting 175

Curriculum Vitae 180

List of publications 181

Acknowledgments 183

Chapter 1

General introduction

Cancer genomics

Cancer develops through the acquisition of genomic alterations, *i.e.* changes in the DNA sequence and chromosomal numerical content of a cell, such as point mutations, insertions, deletions, amplifications, and translocations. Such alterations may alter protein expression and/or function of oncogenes—genes promoting cancer—and tumor suppressors—genes protecting from cancer. Other factors, caused by mechanisms which do not change the underlying DNA sequence (*e.g.* DNA methylation and histone modification), may also alter the expression of genes, and may thereby play a role in cancer as well. As different biological processes, the so-called hallmarks of cancer (1, 2), need to be deregulated before a cancer can develop, a combination of events is needed to change a normal cell into a cancerous cell. Cancer is thought to arise through the stepwise acquisition of such events (3), although it has become clear that different alterations may be caused by a single event (4, 5), and particularly oncogene-activating translocations seem to be sufficient for oncogenesis in some types of leukemias, lymphomas, and sarcomas (6).

In cancer genomics, germline and somatic aberrations, *i.e.* aberrations present in the germline of the patient and acquired aberrations, are studied in order to identify genes and biological processes which are important in the development and progression of cancer. Determining aberrations that are crucial for a cancer cell to survive, identifying defective tumor suppressors, and identifying biological processes which facilitate tumor progression is tremendously important for diagnostics and prognostics, and for the identification of targeted treatments. In the late 1990s, high-throughput methods have been developed which can be utilized in studying cancer genomics—so-called microarrays. In this thesis, we have used these high-throughput techniques in order to study high-grade osteosarcoma genomics, aiming to learn more on osteosarcoma biology, and to identify possible targets for treatment.

High-grade osteosarcoma

High-grade osteosarcoma is a primary malignant tumor consisting of mesenchymal tumor cells producing osteoid. The tumor is rare, with an approximate incidence of 5–6 patients in a population of one million per year. The incidence is higher in adolescents and young adults, and shows a second peak at middle age (7). Osteosarcoma developing later in adult life is thought to be partially secondary, and may be caused by previous treatment with radiation or by an underlying Paget's disease of bone. Males are more often affected by osteosarcoma than females (with a ratio of 3:2). High-grade osteosarcoma most frequently develops in the long bones of patients, with the metaphysis as the most frequent (91%), and the diaphysis as the second most frequent site ($< 9\%$). Most often, the tumor develops in the region around the knee (distal femur and proximal tibia), followed by the proximal humerus (8). Osteosarcoma is rarely seen in the axial bones of the patient. The

incidence pattern of osteosarcoma suggests a link between the development of the disease and growth (9) (this will be further discussed in Chapter 5).

High-grade osteosarcoma is a very aggressive tumor. Patients are usually treated with several series of neoadjuvant chemotherapy consisting of a combination of different chemotherapeutic drugs, especially cisplatin, doxorubicin, and high-dose methotrexate (8). The tumor is then removed by limb-salvage surgery, although sometimes amputation is needed. Afterwards, a second series of adjuvant therapy is given to the patient. Despite this intensive treatment schedule, a significant number of patients die due to the development of distant metastases, which are most often pulmonary. The tumor metastasizes in approximately 45% of all patients (10). Overall survival of patients with resectable metastatic disease is roughly 20% (11). Neoadjuvant treatment was introduced in the 1970s, and improved overall survival from 10–20% to approximately 60%. However, except for macrophage-activating and recruiting agents, such as L-MTP-PE (discussed in Chapter 4 of this thesis), no new treatment options have been developed that can raise survival significantly. The many caveats and challenges hampering osteosarcoma research, which might explain why osteosarcoma patients still have no other treatment options, are discussed in Chapter 2.

Known genes involved in osteosarcomagenesis have essential roles in cell cycle progression (12). The tumor suppressor *TP53*, which can induce cell cycle arrest or apoptosis in response to cellular stress, such as DNA damage, is mutated in approximately 20% of high-grade osteosarcomas and also often present in regions of copy number loss. *MDM2*, which targets the p53 protein for degradation, is amplified in 6–15% of the tumors. *TP53* and *MDM2* aberrations have been described to be mutually exclusive (13), although in our dataset, one sample (the osteosarcoma cell line HAL) had copy number loss of *TP53* and gain of *MDM2*. Inactivating somatic mutations of *RB1*, a negative regulator of the cell cycle, are also often found in osteosarcoma, and this gene is present in regions of copy number loss in over 60% of osteosarcomas (14, 15). Other players of the Rb pathway have been described in osteosarcoma as well, for instance *CDKN2A* deletions, which present homozygously and occur in approximately 25% of all patients (16). *TP53* and *RB1* mutations are not always somatic—a small percentage of osteosarcoma is hereditary, with mutations present in the germline of patients. The associated hereditary syndromes, Li-Fraumeni and Retinoblastoma for mutations in *TP53* and *RB1*, respectively, give a strong predisposition to develop osteosarcoma. A third hereditary syndrome that is thought to predispose to osteosarcoma is Rothmund-Thomson syndrome, where *REQL4*, a gene encoding for a DNA helicase, is mutated (17), however, in contrast to *TP53* and *RB1*, the gene is not a frequent target for sporadic mutations in osteosarcoma (18).

The EuroBoNeT high-grade osteosarcoma database

The aim of this thesis was to study osteosarcomagenesis by bioinformatics analysis of a high-throughput dataset consisting of microarray data from high-grade osteosarcoma specimens. A relatively large cohort of naïve, preoperative diagnostic osteosarcoma biopsies was collected as a collaborative effort by EuroBoNeT, a European Network of Excellence for studying primary bone tumors. This clinically well defined cohort consisted of samples from 84 patients. For most of these patients, clinical data were available on patient sex, age at diagnosis of the primary tumor (in months), tumor location, histological subtype of the tumor, and response to neoadjuvant chemotherapy (Huvos grade) (19). Follow-up data (metastasis-free survival and overall survival, measured in months from diagnosis) was available for 83/84 patients. Clinical characteristics of this cohort can be found in Table 7.1. In addition to the clinical samples, we used data from two osteosarcoma model systems—osteosarcoma cell lines (characterized and published by Ottaviano *et al.* (20)) and xenografts (21), see Table 3.1 for clinical characteristics of the original tumors of these model systems. The entire osteosarcoma database consisted of data obtained from three different microarray platforms—genome-wide gene expression data, data obtained with a kinome screen, and Single Nucleotide Polymorphism (SNP) microarrays. Table 1.1 illustrates the different data types, numbers of osteosarcoma and control samples, and the different comparative analyses which are described in this thesis. Raw and processed data are deposited in online databases (22, 23).

Data type	mRNA	Kinome	SNP
Company	Illumina	PamGene	Affymetrix
Array	Human-6 v2.0	Ser/Thr kinase PamChip	Genome-wide Human SNP Array 6.0
Software	Bioconductor	BioNavigator, Bioconductor	Genotyping Console, Nexus Copy Number
OS samples	84 diagnostic biopsies, 19 cell lines, 18 xenografts	2 cell lines	32 diagnostic biopsies, 12 cell lines
Control samples	12 MSC cultures, 3 osteoblast cultures	12 MSC cultures	27 normal samples
Analysis methods	*LIMMA*, *pamr*	*LIMMA*	Cut-off for aberration frequency
Comparative analyses	Clinical parameters, Tumor *vs* controls	Tumor *vs* controls	Clinical parameters, Tumor *vs* controls

Table 1.1: Layout of the high-grade osteosarcoma database. MSC: mesenchymal stem cell.

High-throughput platforms to study osteosarcoma

Genome-wide gene/mRNA expression profiling can be performed using RNA isolated from a sample, such as tumor tissue or a cell culture. Generally, cDNA or cRNA is prepared

from the RNA and is labeled with a fluorescent dye. This is then hybridized to a microarray chip containing oligonucleotide probes, which are short sequences of DNA, complementary to most or all specific transcripts, capable of binding the labeled cDNA/cRNA. For measuring genome-wide gene expression, single- and dual channel microarrays are available. With dual channel microarrays, samples, *e.g.* paired tumor samples and normal tissues, can be directly compared on one chip, by labeling the cDNA/cRNA with two different fluorescent dyes. For the research described in this thesis (Chapters 3–8), single channel microarrays were used, which means that control samples were hybridized on different chips. We used Illumina Human-6 v2.0 BeadChips (Illumina, San Diego, CA). These microarrays contain over 48,000 probes, of which approximately half are recognized by well-annotated Reference Sequence (RefSeq) genes (24). Illumina BeadChips have a special structure: probes are present on beads, which are randomly arranged on the chip. Every bead type is replicated on each chip with a mean of approximately 35–40 times (25, 26) (see Figure 1.1A). Both the random position and the high amount of repli-

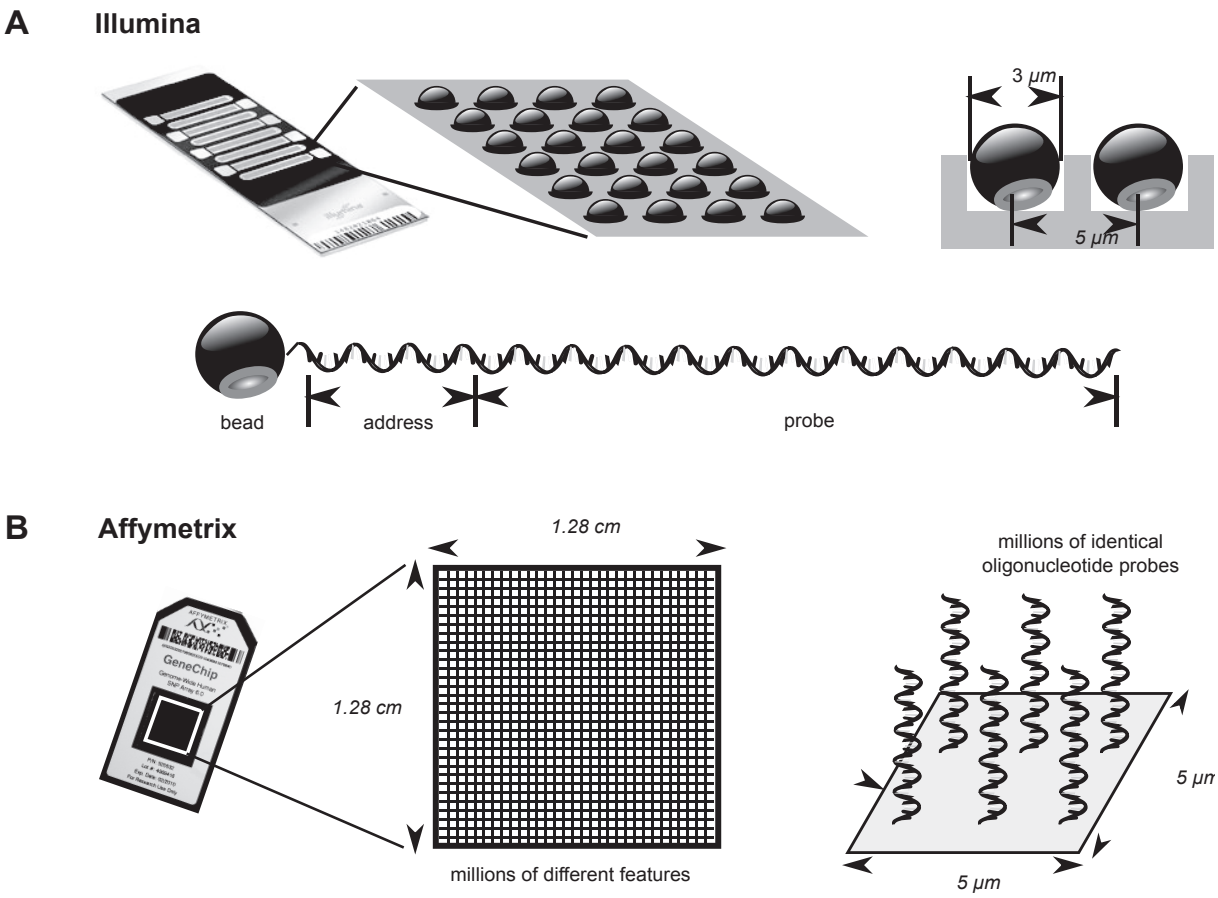


Figure 1.1: Schematic overview of *A*, the Illumina BeadChip and *B*, the Affymetrix SNP 6.0 array. Figure adapted from Hupé, P., `http://commons.wikimedia.org`.

cated beads make robust measurements possible (27). The software designed by Illumina for data analysis, BeadStudio, does not take advantage of the large number of replications of beads present on these chips. Therefore, various methods have been specifically

developed for analyzing Illumina BeadChips, such as Bioconductor (28) packages *beadarray* (27), *beadarraySNP* (29) (specifically for Illumina SNP data), and *lumi* (30), which will be described in the next section.

Peptide microarrays can be used for studying kinase activity in a sample. For the research performed in Chapter 6 of this thesis, we used PamGene® serine/threonine (Ser/Thr) PamChips (PamGene, 's-Hertogenbosch, the Netherlands). These chips consist of porous membranes, which contain 142 different peptides derived from phosphorylation sites for Ser/Thr kinases of the human proteome. Cell or tissue lysates are supplemented with ATP and subsequently pumped through these membranes, so that kinases in the lysates have access to, and can phosphorylate the peptides on the chip. Phosphorylation is measured over a time span of 30 to 60 minutes by the detection of light emitted by fluorescently-labeled, phospho-specific antibodies. Figure 1.2 gives an overview of the experimental workflow of PamGene.

Single Nucleotide Polymorphisms (SNPs) are genetic changes or variations of a single base pair, which occur in at least 1% of the population (31). SNP microarrays contain so-called allele-specific oligonucleotide probes (Figure 1.1B), which are used to discriminate between specific SNPs in the sample, because of the different binding properties of the sample DNA, which is again labeled with a fluorescent dye. SNP microarrays can be employed to genotype a sample, which is used to identify small variations between genomes (to determine *e.g.* disease susceptibility), but can also be utilized to infer copy number aberrations and allelic states of regions in the genome. The SNP microarrays used in this thesis (Chapters 7–8) are Affymetrix Genome-Wide Human SNP Array 6.0 chips (Affymetrix, Santa Clara, CA). These high-density chips contain over $900,000$ SNPs and over $900,000$ probes for the detection of copy number variation.

Microarray data preprocessing

The three different platforms described above have in common that, after hybridization of DNA/cDNA/cRNA to the chip, or after phosphorylation of peptides on the microarray, a fluorescent signal is emitted, which is measured by a scanner. The image files that are returned by the scanner can be utilized for deducing intensity signals and the location of the specific spots/beads. This is usually performed directly by the software provided by the company which distributes the arrays, and generally overlays a grid and returns median intensity signals for each spot/bead. Alternatively raw image files can be analyzed (for example using *beadarray* (27)), thereby allowing additional methods of data processing. In the following paragraphs, we will discuss data preprocessing and subsequent data analysis of data generated with the above described microarray chips.

Preprocessing of microarray data is performed in order to correct for experimental bias and to reduce the signal to noise ratio. Numerous methods of microarray data preprocessing exist, and specific methods may differ per data type and platform. Preprocessing

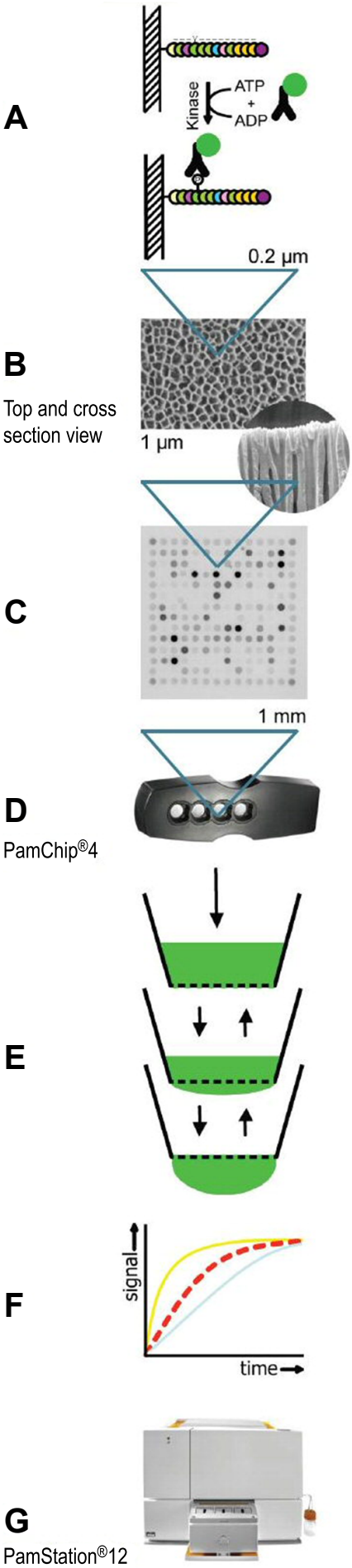


Figure 1.2: Peptides can serve as substrates for kinases present in the sample. Phosphorylation is detected by fluorescently labeled phopho-specific antibodies (*A*). The microarrays consist of a porous ceramic membrane (*B*), on which 142 different peptide substrates are present (*C*). Four arrays are combined into one chip (*D*). The phosphorylation reaction occurs by an up and down movement of the sample solution through the array, giving the kinases maximal opportunity to phosphorylate the peptides on each array (*E*). When the solution is underneath the array, the CCD camera in the workstation takes an image of each array, which is later used by the software to generate kinetic data curves (*F*). The incubation, washing, dispensing of reagents and imaging of the arrays is done in fully automated workstations (*G*). Figure adapted from PamGene®.

of microarray data can be performed using the software provided by the company that produced the arrays, or can be analyzed with open source programs, such as the statistical software R (32), for which several packages have been made available in the Bioconductor (28) framework to specifically analyze the raw data of various microarray platforms.

An optional start of preprocessing the raw data is a global or local background subtraction step. This can eliminate signals due to nonspecific binding, thereby reducing noise in the data. However, when applying this step, probes of low signal will be discarded, resulting in missing values. Some researchers convert these missing values into zero expression. Illumina's scanner software, BeadScan, automatically subtracts local background measures from the foreground intensities to generate bead level text files—files including intensities and location information obtained from the original .tiff files produced by the scanner software. These *bead level files* can be used for downstream data analysis. The standard local background subtraction method provided by Illumina results in a very low estimate of the background, which is thought to be mostly related to the optical properties of the array surface (33). Additional background subtracting methods can be applied, such as *background normalization* in BeadStudio, which subtracts the mean intensity of negative control beads from the foreground intensities. This method increases variability, and also introduces a significant number of negative values (33). Especially for small sample sizes it is crucial to achieve a homogeneous variance, and thus, as background subtraction introduces additional variation in the data, this may not be beneficial for the detection of differences between two or more groups (34). Apart from the local background subtraction by BeadScan (for mRNA expression data), we did not use other background subtraction methods in the preprocessing of our microarray data.

Data transformation is necessary because of the complicated error structure of microarray data, which is intensity-dependent and nonlinear (35). Often, a simple log transformation is used, but other methods exist that are milder in transforming signals near background, which are inflated by standard log transformations. Examples of such methods are variance stabilizing normalization (*vsn*) (36), which both transforms the data and performs normalization of the data between the different arrays, and variance stabilizing transformation (*vst*) (37), a method similar to *vsn*, specifically developed for preprocessing Illumina BeadChips. *vst* has been shown to be advantageous over log transformation when large changes in expression are expected (38, 39). Normalization of the data is applied to reduce bias that may arise due to differences in sample preparation, and production (batch effects) and processing of the arrays. Various normalization methods exist, of which complete data methods, such as quantile normalization, are preferred over methods that use a baseline array in order to normalize the data (40). We used *vst* and robust spline normalization (*rsn*), a normalization method specifically designed to normalize variance stabilization transformed data, on mRNA expression data (Chapters 3–8). Transformation and normalization of peptide chips (Chapter 5) was performed using *vsn*, while SNP microarray data were log_2 transformed and quantile normalized. SNP microarray data

(Chapters 7–8) were further corrected for the guanine-cytosine (GC) content, as different percentages in GC content can cause waviness in the log_2 ratio data, which can increase false positive and false negative segment calls. We used the Regional GC correction algorithm in Genotyping Console to correct for this waviness (41).

Quality control

A very important microarray data preprocessing step is outlier detection. When correctly performed, this step can significantly improve data quality and thereby improve the outcome of the experiment, *e.g.* the detection of differential expression (42, 43). Defective probes from Affymetrix chips can be detected and subsequently removed (44). In Illumina data, spatial artifacts can be detected and removed using BeadArray Subversion of Harshlight, or *BASH* (45). Although the detection of large spatial artifacts may be helpful for determining whole outlier chips, the *BASH* algorithm only improves results very mildly. This can be described to the extremely robust structure of the Illumina BeadChips (tested for Human-6 and GoldenGate BeadChips, Kuijjer *et al.*, *unpublished results*). The more recently developed HumanHT-12 Expression BeadChips contain fewer replicates per bead type, and this preprocessing step may therefore be valuable for removing outliers in these newer chips. Other artifacts in Illumina data have been reported, such as particularly bright beads showing a bleed over effect on neighboring beads, raising their associated values (46). One can adjust for such spatial artifacts by masking affected beads using the *beadarray* package (27).

Regularly, it is necessary to remove entire chips of poor quality, since such chips can impair overall statistical and biological significance (43). Poor quality chips can be identified by visually checking the scanner images, the distribution of both raw and normalized data (*e.g.* by plotting density plots, boxplots, and MA-plots), and by performing unsupervised hierarchical clustering or visualizing the data using principal components analysis (PCA, reducing the data dimensionality to *e.g.* its first two or three principle components). Such methods can for example be applied using Bioconductor package *arrayQualityMetrics* (47) (used for quality control of mRNA and kinome profiling in this thesis) or using quality control functions in the package *affy* (48). Another method to control the influence of poor quality chips is assigning weights to all chips, so that arrays of better quality will have a higher influence on the analysis than poor quality arrays (*arrayWeights* (49)). Such an approach is, however, not intended to replace a quality check identifying catastrophically poor quality chips, and these should still be discarded. In a comparative study of removing poor quality chips with *arrayQualityMetrics*, assigning *arrayWeights* to the data, or applying both methods on the *LIMMA* analysis described in Chapter 4, we determined more differentially expressed probes at a false-discovery rate (FDR, see next section for an explanation) of 0.05 without assigning weights, but this depended on the FDR (for $0.05 < \text{FDR} \leq 0.1$ *arrayWeights* or a combination of both methods performed slightly

better, Kuijjer *et al.*, *unpublished results*).

In SNP microarray quality control, one can determine the ability of an experiment to resolve SNP signals into three genotype clusters (AA, AB, BB). The Affymetrix Genotyping Console Contrast Quality Control test metric is a measure for this ability (50), and was used in this thesis (Chapters 7–8). This test uses 10,000 random SNPs to measure the difference between peaks in the distributions of homozygote genotypes (AA and BB), and the valleys these distributions share with the heterozygote peak (AB). When this difference approaches zero, the experiment poorly distinguishes between homozygous and heterozygous genotypes. Such chips should be removed from further data analysis.

Microarray data analysis

After having performed the preprocessing steps necessary for the specific type and platform of microarray data, the actual data analysis can be performed.

Unsupervised hierarchical clustering of microarray data may not only be used as a quality check (as described in the previous section), but can also be applied to detect different subgroups of samples, which may be associated with a clinical feature. In a supervised approach, differences between groups of samples can be determined using a moderated t-test, such as the *LIMMA* analysis (used in this thesis for detection of differential expression and phosphorylation) (51). Important to note is that with the testing of multiple hypotheses, the amount of true null hypotheses that are rejected will increase. In microarray experiments, often large numbers of probes/peptides are tested for differential expression or phosphorylation, and therefore, an excessive amount of false-positives may be returned from conventional statistical tests. Hence, a correction for multiple testing should be performed (42). Examples of such methods are conservative familywise error rate procedures, such as the Bonferroni method (52), or the less stringent false discovery rate (FDR) controlling methods, *e.g.* the Benjamini and Hochberg (53), and Benjamini and Yekutieli (54) approaches. Other methods use permutations to estimate the FDR, such as Significant Analysis of Microarrays (*SAM*) (55).

SNP data is analyzed in a different manner. Genotyping can be performed by specific genotyping algorithms, such as the Birdseed v2 algorithm in Genotyping Console, which uses unsupervised learning to fit the data, producing genotype calls and returning confidence scores for each SNP (56). Copy number data analysis is performed by comparing the intensity signals for each marker and each sample against a reference genome, which usually consists of a set of in-house or publicly available control samples. A cut-off for gains and losses is used to determine whether probes are present in a region of amplification or deletion (in this thesis, an absolute $\log_2$ ratio cut-off of 0.2, equivalent to an absolute fold change of approximately 1.15, was used). Using the genotyping information, calls can also be made for allelic ratios. In Nexus Copy Number software, this is done by determining the B-allele frequency. Regions on the genome which show LOH will not

reveal any AB signals (a B-allele frequency of 0.5, at least in theory, if there are no normal cells present in the tumor tissue). This also makes the identification of allelic imbalance possible, which, over a genomic region, will show multiple B-allele frequencies in between 0 and 1, depending on the amounts of copy number of each allele.

A drawback in SNP data analysis is that copy number changes are detected relative to the overall DNA content in the sample (57). In addition, normal cell populations, such as stromal and inflammatory cells, and heterogeneity within the tumor itself can further impede the detection of the true copy number alterations in the tumor cell. In epithelial tumors, a DNA index can be determined by flow-sorting tumor cells, which can separate these from mesenchymal cells, and which can identify subpopulations of tumor cells with different chromosomal aberrations. To infer true copy numbers and allelic states, the algorithm lesser allele intensity ratio (LAIR, included in *beadarraySNP* (29)) integrates the DNA index in the analysis of SNP data (58). Unfortunately, this approach can not be applied to SNP data analysis of high-grade osteosarcoma samples, as osteosarcoma is a mesenchymal tumor for which no specific markers are available. However, the amount of stroma in osteosarcoma is not as extensive as in epithelial tumors, and the percentage of stroma as determined by the pathologist could in principle be used in order to approximate the DNA index of these tumors.

SNP microarray data show a high degree of noise, and not all markers reflect the true copy number of the region. Segmentation is performed in order to identify the chromosomal segments with actual copy number aberrations. Most frequently used algorithms for segmentation are Circular binary segmentation (CBS)-based (59) or Hidden Markov Model (HMM)-based methods. CBS-based methods divide the genome into always smaller segments until no region can be further segmented, taking into account a minimum amount of probes per segment. The SNPRank segmentation algorithm in Nexus Copy Number Software is CBS-based, and ranks log ratio probe values and B-allele frequencies in a segment. If the distribution of these probe ranks is significantly different from those of an adjacent segment, the region is segmented out, meaning the region probably has a different median copy number than that of the adjacent segment. HMM-based methods, such as the SNP-FASST segmentation algorithm in Nexus Copy Number software, perform faster than CBS-based methods, but require an estimate of signal–copy number relationship, as it works with integer copy numbers. Because of the heterogeneity present in tumor samples, this is probably not an optimal way to segment tumor data (60). We used SNPRank segmentation to segment the copy number data, with a minimum of 5 probes per segment. After segmentation, a cut-off for frequency of copy number changes can be set, so that the most recurrent alterations will be detected. One can also specifically look for focal or broad events, as is described in Chapter 9 of this thesis. As with the analysis of other microarray data types, permutations can be used to determine whether there are significant differences in copy number or LOH profiles of groups with different features (*e.g.* in Nexus Copy Number software).

Downstream data analysis

Deducing a biological interpretation from large lists of significant genes may be challenging, and validation of all significant genes is often very labor intensive. Several methods have been developed which determine whether specific signal transduction pathways, biological processes, or other groups of genes with similar functions, are affected. Genes making up such pathways or processes are often taken from public databases, such as the Gene Ontology (GO) (61) or the Kyoto Encyclopedia of Genes and Genomes (KEGG) (62), or are available as commercial software, such as Ingenuity Pathways Analysis (IPA, Ingenuity Systems), which is manually curated. The hypergeometric test (a one-tailed Fisher's exact test) is most often used to obtain information on the enrichment of significant genes in specific pathways or biological processes. This test determines whether there is more overlap between the list of significant genes and the set of genes of interest (*e.g.* the pathway) than would be expected by chance. The hypergeometric test can be applied on microarray data in IPA (used in Chapter 6) and in the Bioconductor *topGO* package (used in Chapters 3 and 7) (63). A disadvantage of this simple test is that it requires a hard definition of significance (*e.g.* a p-value cut-off), and discards information on the exact p-values of the genes tested. The hypergeometric test also assumes independence of genes, which is not accurately representing the biology of a cell, since the expression of functionally related genes is often correlated. Because of this assumption, the hypergeometric test may understate the true p-values. It is therefore recommended to use a very low p-value (*e.g.* 0.001 or lower) as cut-off for significance when applying this test. Another problem of the hypergeometric test is that it assumes independence of categories. GO terms are certainly not independent, as these terms are set up in a hierarchical structure of nodes, with parent terms representing a broader GO term, and child terms a more specific subset of its parent terms (61, 64). Algorithms which can identify the GO term which better represents the biological situation (significantly affected genes) than other terms from its neighborhood have been developed, such as the *weight* algorithm in the *topGO* package (63).

A method which takes into account a continuous measure of significance is gene set enrichment analysis (*GSEA*) (65). This method ranks genes based on their associated p-values and subsequently determines an enrichment score based on the rank of the genes present and not present in a specific pathway or category. The significance of this enrichment score is subsequently tested by permuting phenotype labels to determine the null distribution of the enrichment score.

Another approach to determine which biological pathways are significantly affected is the *globaltest* (used in Chapter 5). Based on a logistic regression model, this test determines whether a prespecified group of genes is differentially expressed, and thus tests groups of genes instead of single genes (66). This test is particularly intended for identifying gene sets for which many genes are associated with a phenotype in a small way.

Using this approach may be especially fruitful in case no overall differential expression is detected due to small sample sizes, as this approach significantly reduces the multiple testing problem (67). The *globaltest* has much more power than self-contained tests (tests which compare a gene set with its complement), such as the hypergeometric test (68). To apply the *globaltest* on GO terms, Goeman *et al.* also developed a method that preserves the specific graph structure of the Gene Ontology (69). In addition, this algorithm can be used in combination with follow-up data (67).

A final method to extract biological information from lists of significantly affected genes is performing network analysis. Networks are assembled *de novo*, based on connectivity (*e.g.* binding or functional properties) between affected molecules. In IPA, networks are assembled using decreasingly connected molecules from the significant genes in the dataset which is analyzed, and are annotated with functional categories, which are manually curated. In contrast to pathway analysis, these IPA networks do not have directionality (but network analysis methods which include directionality between molecules also exist). We used network analysis to interpret differential gene expression between various histological subtypes of osteosarcoma (Chapter 3).

Supervised learning

Generating a prediction profile which can classify tumors based on mRNA expression or specific copy number aberrations may also be used in microarray analysis of a cancer dataset. Classification may for example help to diagnose a tumor based on its microarray data profile, or may predict event-free or overall survival of patients. Some examples of supervised learning approaches are nearest shrunken centroids classification (*e.g.* available in Bioconductor package *pamr* (70)), support vector machine (SVM) learning (*e.g.* available in R package *e1701* (71)), and random forest classification (*e.g.* available in R package *varSelRF* (72)).

In this thesis, we used nearest shrunken centroids classification to develop a classifier of the main histological subtype of conventional osteosarcoma. We validated this classifier on an independent dataset, and applied it on data obtained from osteosarcoma model systems (Chapter 3). Nearest centroids classification determines centroids for each class by dividing average expression of a gene signature by the standard deviation. New samples are classified to that specific class, of which the centroid is closest—in squared distance—to the expression of the genes in the prediction profile. Nearest shrunken centroids is an adaptation of this method—it shrinks each centroid toward the overall centroid for all classes by a certain threshold. This shrinkage automatically selects genes and reduces the effect of noisy genes. The profile with the lowest prediction error is then selected as the final classifier. Internal cross-validation, which divides the training set in different parts, is subsequently used to compute a cross-validated error. This approach, however, leads to an underestimation of the error rate, as the same data is used to select features and to

estimate the error rate. An extra external cross-validation step would thus be appropriate, or, given that there is often only a limited number of samples available for training, the feature selection (genes to include in the profile) should be newly computed for each separate cross-validation step (73, 74). External cross-validation is performed in order to correct for overfitting of the data by the model. This can be done on an independent cross-validation set, or by using methods such as one-leave-out cross-validation (75). Also regularization may be used to prevent overfitting, but this is not often used in microarray data analysis, and is therefore beyond the scope of this thesis.

In prediction profiling, the way the distance between the actual sample and the class is calculated may be very different, and this has important consequences for biological interpretation of the profile. In a prediction profile where the magnitude (*e.g.* of gene expression) is important, Euclidian distance is best used, while correlation (*e.g.* Pearson or Spearman) coefficients are more useful when the way the genes depend on each other, so the pattern of expression, is important for the specific gene list (76). This may be one of the reasons why the CINSARC profile, a gene expression signature which was generated on sarcomas (77) and which uses Spearman correlation as a measurement for distance, did not show significant results on our osteosarcoma dataset (centroids for classifier needed to be retrained, because we used data of a different platform than the original CINSARC signature, Kuijjer *et al.*, *unpublished results*), while the Carter signature (78), which classifies data based on average expression of genomic instability genes, could predict for metastasis-free survival in our data (as shown in Chapter 7).

Data integration

As explained in the next chapter, the integration of different data types is particularly relevant when studying a highly genomically unstable tumor. We used superimposed integration of mRNA expression and kinome profiling data in Chapter 6. This approach was taken, because kinase activity usually does not have a direct downstream effect on mRNA expression (generally, there are several intermediate molecules which confer signaling), and the other way around. It may therefore be more relevant to determine how these data complement each other, instead of identifying only overlapping genes.

For integration of copy number and gene expression (Chapters 7–8) data, we identified genes with aberrations occurring in both data types, as the copy number state of a gene can have a direct effect on its expression. We specifically chose to identify cooccurrence and not correlation of copy number and expression signals, because these signals do not have to show a linear correlation, *i.e.* correlation will miss our genes which are also regulated at other dimensions, such as epigenetics and feedback mechanisms.

A conservative approach was taken—only genes which were significantly differentially expressed between osteosarcoma tumors and presumed osteosarcoma progenitors were analyzed, and the cut-off for recurrence was set to 35%. We tested this approach in a

paired and nonpaired way to determine cooccurrence of copy number aberrations and differential expression in Chapter 7, and used paired analysis of cooccurrence of LOH, copy number gains, and differential expression in Chapter 8.

Aims and outline of this thesis

In this thesis, a systems biology approach to study high-grade osteosarcoma is described. Chapter 1 starts with an introduction on cancer genomics and high-grade osteosarcoma, and introduces the EuroBoNeT high-grade osteosarcoma database, on which the research in the following chapters is based. In addition, different platforms used in this thesis are described, and different types of high-throughput data analyses are explained (this chapter).

In Chapter 2, published literature on microarray studies on high-grade osteosarcoma is reviewed. This review also discusses challenges in high-throughput data analysis of osteosarcoma and introduces different model systems which have been used in osteosarcoma research. In addition, information on different comparative analyses and a rationale for integrating different data types are given. The review concludes with a section on how bioinformatics can be translated into functional studies.

The following six chapters of the thesis describe the work which has been performed to answer different research questions regarding osteosarcoma biology and possible targets for therapy. Specifically, we aimed to study molecular differences between clinically different tumors, such as tumors of different histological subtypes, and of tumors with different metastasis-free survival profiles. These research questions are answered in Chapters 3 and 4, respectively. In addition, in Chapter 3, a histological subtype-specific gene expression profile is tested on osteosarcoma model systems. High-grade osteosarcoma is also compared with controls, in order to detect what signal transduction pathways may be targeted in osteosarcoma to identify potential adjuvant drugs for treatment of this aggressive tumor (Chapters 5 and 6). Chapter 5 reports on the analysis of gene expression data, while Chapter 6 determines active pathways based on kinome profiling, and integrates gene expression data with kinome profiling results. Finally, we performed integrative data analysis of SNP and gene expression data, to detect osteosarcoma driver genes (Chapters 7 and 8). In Chapter 7, copy number aberrations are integrated with overexpression and downregulation, while in Chapter 8 we specifically look at the combination of Loss of Heterozygosity (LOH), DNA copy number gain, and differential mRNA expression.

In Chapter 9, results described in Chapters 3 to 8 are discussed and future perspectives for high-throughput data analysis on high-grade osteosarcoma are given. Chapter 10 includes a Dutch summary, Curriculum Vitae, and a list of publications.

References

[1] Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*, 100(1):57–70, 2000.

[2] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674, 2011.

[3] Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*, 61(5):759–767, 1990.

[4] Stephens PJ, Greenman CD, Fu B, Yang F, *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*, 144(1):27–40, 2011.

[5] Yates LR, Campbell PJ. Evolution of the cancer genome. *Nature Reviews Genetics*, 13(11):795–806, 2012.

[6] Knudson AG. Human Cancer Genetics. In Keynes M, Edwards AWF, Peel R, editors, A century of Mendelism in human genetics, 133–144. CRC Press, 2004.

[7] Mirabello L, Troisi RJ, Savage SA. Osteosarcoma incidence and survival rates from 1973 to 2004. *Cancer*, 115(7):1531–1543, 2009.

[8] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[9] Mirabello L, Pfeiffer R, Murphy G, Daw NC, *et al.* Height at diagnosis and birth-weight as risk factors for osteosarcoma. *Cancer Causes and Control*, 22(6):899–908, 2011.

[10] Pakos EE, Nearchou AD, Grimer RJ, Koumoullis HD, *et al.* Prognostic factors and outcomes for osteosarcoma: an international collaboration. *European Journal of Cancer*, 45(13):2367–2375, 2009.

[11] Buddingh EP, Anninga JK, Versteegh MIM, Taminiau AHM, *et al.* Prognostic factors in pulmonary metastasized high-grade osteosarcoma. *Pediatric Blood & Cancer*, 54(2):216–221, 2010.

[12] Cleton-Jansen AM, Buerger H, Hogendoorn PCW. Central high-grade osteosarcoma of bone: diagnostic and genetic considerations. *Current Diagnostic Pathology*, 11(6):390–399, 2005.

[13] Overholtzer M, Rao PH, Favis R, Lu XY, *et al.* The presence of p53 mutations in human osteosarcomas correlates with high levels of genomic instability. *Proceedings of the National Academy of Sciences*, 100(20):11547–11552, 2003.

[14] Thomas DM, Yang HS, Alexander K, Hinds PW. Role of the retinoblastoma protein in differentiation and senescence. *Cancer Biology & Therapy*, 2(2):123–129, 2003.

[15] Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, *et al.* Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes, Chromosomes and Cancer*, 51(7):696–706, 2012.

[16] Mohseny AB, Tieken C, van der Velden PA, Szuhai K, *et al.* Small deletions but not methylation underlie CDKN2A/p16 loss of expression in conventional osteosarcoma. *Genes, Chromosomes and Cancer*, 49(12):1095–1103, 2010.

[17] Calvert GT, Randall RL, Jones KB, Cannon-Albright L, *et al.* At-risk populations for osteosarcoma: the syndromes and beyond. *Sarcoma*, 2012:152382, 2012.

[18] Nishijo K, Nakayama T, Aoyama T, Okamoto T, *et al.* Mutation analysis of the RECQL4 gene in sporadic osteosarcomas. *International Journal of Cancer*, 111(3):367–372, 2004.

[19] Huvos AG. Bone tumors: diagnosis, treatment and prognosis. WB Saunders Company, 1991.

[20] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[21] Mayordomo E, Machado I, Giner F, Kresse SH, *et al.* A tissue microarray study of osteosarcoma: histopathologic and immunohistochemical validation of xenotransplanted tumors as preclinical models. *Applied Immunohistochemistry & Molecular Morphology*, 18(5):453–461, 2010.

[22] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1):207–210, 2002.

[23] R2: microarray analysis and visualization platform. `r2.amc.nl`.

[24] Pruitt K, Brown G, Tatusova T, Maglott D. The Reference Sequence (RefSeq) database. In McEntyre J, Ostell J, editors, The NCBI handbook [internet]. National Center for Biotechnology Information, 2002.

[25] Oliphant A, Barker DL, Stuelpnagel JR, Chee MS. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques*, 32(6):56–58, 2002.

[26] Barbosa-Morais NL, Dunning MJ, Samarajiwa SA, Darot JFJ, *et al.* A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. *Nucleic Acids Research*, 38(3):e17, 2010.

[27] Dunning MJ, Smith ML, Ritchie ME, Tavaré S. beadarray: R classes and methods for Illumina bead-based data. *Bioinformatics*, 23(16):2183–2184, 2007.

[28] Gentleman RC, Carey VJ, Bates DM, Bolstad B, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004.

[29] Oosting J. beadarraySNP: normalization and reporting of Illumina SNP bead arrays. Bioconductor. Available from: `www.bioconductor.org/packages/2.11/bioc/html/beadarraySNP.html`, 2010. Accessed: 02/10/2013.

[30] Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics*, 24(13):1547–1548, 2008.

[31] Gibbs RA, Belmont JW, Hardenbol P, Willis TD, *et al.* The international HapMap project. *Nature*, 426(6968):789–796, 2003.

[32] Team RDC. R: a language and environment for statistical computing, reference index version 2.15.0. R Foundation for Statistical Computing, 2011.

[33] Dunning MJ, Barbosa-Morais NL, Lynch AG, Tavaré S, *et al.* Statistical issues in the analysis of Illumina data. *BMC Bioinformatics*, 9(1):85, 2008.

[34] Schmid R, Baum P, Ittrich C, Fundel-Clemens K, *et al.* Comparison of normalization methods for Illumina BeadChip HumanHT-12 v3. *BMC Genomics*, 11(1):349, 2010.

[35] Durbin BP, Hardin JS, Hawkins DM, Rocke DM. A variance-stabilizing transformation for gene-expression microarray data. *Bioinformatics*, 18(suppl 1):S105–S110, 2002.

[36] Huber W, Von Heydebreck A, Sültmann H, Poustka A, *et al.* Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, 18(suppl 1):S96–S104, 2002.

[37] Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Research*, 36(2):e11, 2008.

[38] Dunning MJ, Ritchie ME, Barbosa-Morais NL, Tavaré S, *et al.* Spike-in validation of an Illumina-specific variance-stabilizing transformation. *BMC Research Notes*, 1(1):18, 2008.

[39] Du P, Lin S, Huber W, Kibbe WA. Evaluation of VST algorithm in lumi package. Bioconductor. Available from: `www.bioconductor.org/packages/2.2/bioc/html/lumi.html`, 2010. Accessed: 02/10/2013.

[40] Bolstad BM, Irizarry RA, Åstrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2):185–193, 2003.

[41] Affymetrix. Copy number algorithm with built-in GC waviness correction in Genotyping Console software. Available from: `www.affymetrix.com`.

[42] Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews Genetics*, 7(1):55–65, 2006.

[43] Kauffmann A, Huber W. Microarray data quality control improves the detection of differentially expressed genes. *Genomics*, 95(3):138–142, 2010.

[44] Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proceedings of the National Academy of Sciences*, 98(1):31–36, 2001.

[45] Cairns JM, Dunning MJ, Ritchie ME, Russell R, *et al.* BASH: a tool for managing BeadArray spatial artefacts. *Bioinformatics*, 24(24):2921–2922, 2008.

[46] Smith ML, Dunning MJ, Tavaré S, Lynch AG. Identification and correction of previously unreported spatial phenomena using raw Illumina BeadArray data. *BMC Bioinformatics*, 11(1):208, 2010.

[47] Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics*, 25(3):415–416, 2009.

[48] Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, 20(3):307–315, 2004.

[49] Ritchie ME, Diyagama D, Neilson J, van Laar R, *et al.* Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics*, 7(1):261, 2006.

[50] Affymetrix. Quality control assessment in Genotyping Console. Available from: `www.affymetrix.com`.

[51] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):3, 2004.

[52] Weisstein EW. Bonferroni correction. From MathWorld—a Wolfram web resource. Available from: `mathworld.wolfram.com/BonferroniCorrection.html`, 2006. Accessed: 02/10/2013.

[53] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57(1):289–300, 1995.

[54] Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4):1165–1188, 2001.

[55] Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*, 98(9):5116–5121, 2001.

[56] Affymetrix. Affymetrix Genotyping Console 4.0 user manual.

[57] Attiyeh EF, Diskin SJ, Attiyeh MA, Mossé YP, *et al.* Genomic copy number determination in cancer cells from single nucleotide polymorphism microarrays based on quantitative genotyping corrected for aneuploidy. *Genome Research*, 19(2):276–283, 2009.

[58] Corver WE, Middeldorp A, ter Haar NT, Jordanova ES, *et al.* Genome-wide allelic state analysis on flow-sorted tumor fractions provides an accurate measure of chromosomal aberrations. *Cancer Research*, 68(24):10333–10340, 2008.

[59] Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, 5(4):557–572, 2004.

[60] Rasmussen M, Sundström M, Kultima HG, Botling J, *et al.* Allele-specific copy number analysis of tumor samples with aneuploidy and tumor heterogeneity. *Genome Biology*, 12(10):R108, 2011.

[61] Ashburner M, Ball CA, Blake JA, Botstein D, *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25, 2000.

[62] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27–30, 2000.

[63] Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–1607, 2006.

[64] Rhee SY, Wood V, Dolinski K, Draghici S. Use and misuse of the gene ontology annotations. *Nature Reviews Genetics*, 9(7):509–515, 2008.

[65] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005.

[66] Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics*, 20(1):93–99, 2004.

[67] Goeman JJ, Oosting J, Cleton-Jansen AM, Anninga JK, *et al.* Testing association of a pathway with survival using gene expression data. *Bioinformatics*, 21(9):1950–1957, 2005.

[68] Goeman JJ, Bühlmann P. Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics*, 23(8):980–987, 2007.

[69] Goeman JJ, Mansmann U. Multiple testing on the directed acyclic graph of gene ontology. *Bioinformatics*, 24(4):537–544, 2008.

[70] Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proceedings of the National Academy of Sciences*, 99(10):6567–6572, 2002.

[71] Dimitriadou E, Hornik K, Leisch F, Meyer D, *et al.* Misc functions of the Department of Statistics (e1071), TU Wien. CRAN R-project. Available from: `cran.r-project.org/web/packages/e1071/index.html`, 2008. Verified: 04/06/2009.

[72] Diaz-Uriarte R. GeneSrF and varSelRF: a web-based tool and R package for gene selection and classification using random forest. *BMC Bioinformatics*, 8(1):328, 2007.

[73] Wood IA, Visscher PM, Mengersen KL. Classification based upon gene expression data: bias and precision of error rates. *Bioinformatics*, 23(11):1363–1370, 2007.

[74] Ambroise C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Sciences*, 99(10):6562–6566, 2002.

[75] Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *Journal of the National Cancer Institute*, 95(1):14–18, 2003.

[76] Quackenbush J. Microarray analysis and tumor classification. *New England Journal of Medicine*, 354(23):2463–2472, 2006.

[77] Chibon F, Lagarde P, Salas S, Pérot G, *et al.* Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nature Medicine*, 16(7):781–787, 2010.

[78] Carter SL, Eklund AC, Kohane IS, Harris LN, *et al.* A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nature Genetics*, 38(9):1043–1048, 2006.

Chapter 2

Genome-wide analyses on high-grade osteosarcoma: making sense of a genomically most unstable tumor

This chapter is based on the review: Kuijjer ML, Hogendoorn PCW, Cleton-Jansen AM. *Int J Cancer.* 2013 Feb 22;Epub

Abstract

High-grade osteosarcoma is an extremely genomically unstable tumor. This, together with other challenges, such as the heterogeneity within and between tumor samples, and the rarity of the disease, renders it difficult to study this tumor on a genome-wide level. Now that most laboratories change from genome-wide microarray experiments to Next-Generation Sequencing it is important to discuss the lessons we have learned from microarray studies. In this review, we discuss the challenges of high-grade osteosarcoma data analysis. We give an overview of microarray studies that have been conducted so far on both osteosarcoma tissue samples and cell lines. We discuss recent findings from integration of different data types, which is particularly relevant in a tumor with such a complex genomic profile. Finally, we elaborate on the translation of results obtained with bioinformatics into functional studies, which has lead to valuable findings, especially when keeping in mind that no new therapies with a significant impact on survival have been developed in the past decades.

Introduction

High-grade osteosarcoma, a rare, genomically complex and unstable tumor

High-grade osteosarcoma is the most prevalent primary malignant bone tumor. The disease occurs most often in children and adolescents and is the sixth leading cause of death in children under the age of 15 years. Notwithstanding, osteosarcoma is a rare disease, with an incidence of five to ten new cases per $1,000,000$ per year (1, 2). Osteosarcoma is composed of extremely genomically complex and unstable mesenchymal tumor cells, generally exhibiting both complex clonal and numerous nonclonal aberrations (1), which are characterized by the direct production of osteoid (2, 3). The tumor is highly aggressive, with distant metastases developing in approximately 45% of all patients (4) although patients are treated with intensive neoadjuvant treatment consisting of high doses of multiple chemotherapeutic drugs. Better surgery has improved survival slightly but no other significant improvement has been made since decades, and increasing dose or the administration of more than three chemotherapeutic regimens does not increase overall survival (5–7). Hence, new therapeutics are seriously needed. Studying the tumor biology and pathology in a systematic manner can result in a better understanding of osteosarcomagenesis and can potentially identify new targets for treatment.

Caveats and challenges

Several challenges and caveats are encountered when studying a rare, highly genomically unstable tumor on a genome-wide level. The first challenge is apparent when collecting osteosarcoma tumor samples. Osteosarcoma is a rare disease and therefore often large interinstitutional efforts have to be achieved to collect the substantial amount of samples that is needed for analyses in computational biology. For most purposes, studying osteosarcoma pretreatment diagnostic biopsies is preferred over using resection material of the primary tumor. Presurgery chemotherapy causes substantial necrosis, even in poor responders, thereby rendering the tissue unsuitable for high quality nucleic acid retrieval. Moreover, biopsies are more representative of the state of the tumor before any treatment as chemotherapy changes the distribution of subclones present in the primary tumor, and can cause clonal evolution (8). Biopsies are taken to establish a histopathological diagnosis, and are unfortunately often very small and not always available for research. In addition, material is often collected retrospectively, which can introduce heterogeneity owing to, for example, different treatment procedures, unless patients are collected who have been enrolled in the same clinical trial. Thus, the collection of clinical data and the grouping of clinical parameters have to be carried out very carefully. For a rare entity such as osteosarcoma, collaborations are indispensable to collect significant cohorts, an example of this being the European Network of Excellence EuroBoNeT, in which various European institutes collaborated to collect a large, homogeneous set of, among other bone tumors, high-grade osteosarcoma biopsies.

Primary osteosarcoma is subdivided into numerous different low- and high-grade subtypes (9). In this review, we concentrate on high-grade conventional osteosarcoma, which is by far the most prevalent variant. Although there is often intratumor heterogeneity, high-grade conventional osteosarcoma can be grouped into various histological subtypes, based on the produced extracellular matrix of the tumor (9). Osteoblastic, chondroblastic and fibroblastic osteosarcoma are the most common histological subtypes of high-grade conventional osteosarcoma. Some correlation of the distinct histological subtypes to specific clinical outcomes has been observed (10, 11) and it may thus be difficult to collect a homogeneous set of samples. In fact, often it is not clearly described which exact histological subtypes are used in a specific study, and in what percentages these subtypes are present in the data set. In addition, the subclassification is hindered by the occurrence of mixed cases containing two different matrix types. Nonetheless, a concordance of 98% has been found between the histological subtype of osteosarcoma biopsies and the corresponding resections (10).

A general problem in studying tumor cell biology is that the true cell of origin is often not defined, rendering it difficult to select a representative control tissue or control cells. Osteosarcoma cells are osteoblast-like cells of mesenchymal origin. Of the different histological subtypes that exist, multiple subtypes can be present within a single tumor.

Considering the differentiation capacity of the mesenchymal stem cell (MSC), this cell type is the most probable candidate for being the osteosarcoma progenitor (12, 13). It was recently found that osteosarcoma tumors can be spontaneously formed when mouse MSCs are transferred into mice (14, 15) and zebrafish (16, 17). This does, however, not exclude osteoblasts as putative progenitor cells, as osteoblasts might redifferentiate into the primitive osteoblast-like tumor cells of osteosarcoma.

Osteosarcoma models

As collecting fresh frozen osteosarcoma tumor samples can be a challenge, performing analyses on data derived from osteosarcoma cell lines or xenografts may be a good alternative (17). Osteosarcoma cell lines are frequently used in biological studies, because they generally grow fast and are easy to maintain in culture and hence osteosarcoma cell lines are easily available. One caveat of using cell cultures is that slight differences in culture conditions, for example the percentage of cells in the culture dish or flask, or the medium that is used, can lead to significant differences in protein expression or signal transduction pathway activities, and these specific conditions may differ per cell line. Using a large panel of cell lines cultured under standard settings can overcome this problem. Cell culture may furthermore introduce additional mutations and genomic aberrations in the cell genome, because of selection based on the *in vitro* conditions (18), but in general, cell lines are reported to adequately represent the tumor from which they are derived. *In vitro*, they preserve the genetic aberrations of the parent tumor, while acquiring additional locus-specific alterations (19).

A panel of 19 osteosarcoma cell lines was recently characterized genetically by MLPA on 38 tumor suppressor gene loci (20). A screen for *TP53* mutations, *MDM2* amplification, *CDKN2A/B* deletion and genomic deletions of 38 additional tumor suppressor genes was performed on these cell lines. As three cell lines of this panel—HOS, 143B and MNNG-HOS— have common ancestry, we report the following percentages based on 17 cell lines. Homozygous deletion of the *CDKN2A/B* locus was detected in 35%, whereas hemizygous deletion of this locus was found in 24% of osteosarcoma cell lines. An additional homozygous deletion was found for *TP73* in one cell line. Mutation in *TP53* was detected in 41%, whereas *MDM2* amplification was detected in 17% of cell lines. These percentages are higher than those in osteosarcoma tumor tissues that are reported in the previously published literature (21), which may be explained by an advantage for primary tumor cells harboring such mutations to be effectively immortalized, or by the acquisition of additional mutations owing to long-term culture. *MDM2* amplification and *TP53* mutations were mutually exclusive in this cell line panel. This has also been observed in osteosarcoma tumor data (22). The differentiation capacity of this cell line panel has been determined as well (23). All 19 cell lines were able to differentiate toward at least one of the three tested—osteoblastic, chondroblastic and adipocytic—lineages. Most cell

lines (14/19) could differentiate to at least two lineages, whereas 3/19 cell lines had full differentiation capacity.

In vivo osteosarcoma model systems include transplantation of a human tumor in mice (24, 25), subcutaneous or orthotopically injections of osteosarcoma cells or late-passage transformed MSCs into mice (15) or zebrafish (16). Transgenic mouse models of osteosarcoma can be developed by overexpression of *c-fos* (26), or conditional inactivation of *TP53* and *RB1* (27). These different models have been shown to resemble osteosarcoma phenotypically (15, 16, 23–28). For example, subcutaneous and intramuscular injection of osteosarcoma cells in nude mice resulted in high-grade sarcoma, resembling tumors which produced osteoid (23) for 8/19 cells from the above-described panel. The *in vivo* lineage-specific differentiation capacity of these cells, however, was limited, reflecting the importance of stomal or microenvironmental stimulation for this process.

As with cell lines, xenograft tumor cells may acquire additional changes owing to selection, and often, xenografts lose matrix after several passages (24). This will probably not have a significant effect on genomic profiles, but does influence expression and methylation patterns. High-resolution microassay-based array comparative genomic hybridization (aCGH) including nine osteosarcoma patient–xenograft pairs showed that genomes of human tumors transplanted into immunodeficient mice, which were repeatedly passaged in new mice, where comparable to genomes of their tumor of origin, with the acquisition of only a small number additional significant changes in the xenograft genomes (25). Different microarray studies have shown that osteosarcoma cell lines and xenografts resemble the primary tumor from which they are derived. Gene expression profiling of a subset of the EuroBoNeT cell line panel, for which the original histological subtype of the primary tumor was known, and of osteosarcoma xenografts and pretreatment biopsies showed that, despite the lower amounts of matrix, histological subtypespecific mRNA signatures are retained in these model systems, and therefore may be a useful tool for expression analysis (Chapter 3, (28)). Despite the similarities between genome and expression profiles of the model systems described above and the tumors of origin, the absence (cell lines) or lower amounts (xenografts) of stromal cells and extracellular matrix, the absence of interaction with the immune system (cell lines and some xenograft models) and the higher degree of clonality remain important limitations for studying tumor biology using these model systems.

Genome wide profiling to study osteosarcoma

In the next sections, we describe different methods to analyze specific types of microarray data, and give examples of how results from bioinformatics can be translated into functional studies. This review is not aiming to give a comprehensive overview of all genome-wide studies on osteosarcoma, but rather illustrates and summarizes the major findings on DNA/RNA microarray reports. A summary of these findings is provided in

Analysis	Data type	Study	Osteosarcoma samples	Comparison	Pathway/genes
Single-way	mRNA	Kuijjer *et al.*[28]	76 B, 13 X, 18 C	Histological subtypes	NFκB in fibroblastic, chondroid-matrix-associated genes in chondroblastic osteosarcoma. Primary tumor expression signatures are preserved in model systems
		Buddingh *et al.*[30]	53 B	Metastasis-free survival	Macrophage-associated genes correlate with better MFS
		Su *et al.*[31]	3 C, 5 X	Capacity to metastasize	*IGFBP5* downregulation correlates with metastasis
		Namløs *et al.*[33]	12 B/T, 11 M	Tumor sample type	Immunological processes and chemokine pattern upregulated in metastases
		Cleton-Jansen *et al.*[32]	25 B	Response to chemotherapy	No significant differential expression
		Kuijjer *et al.*[28]	69 B		
		Cleton-Jansen *et al.*[32]	25 B	Control samples (osteoblastoma, MSC, osteoblast)	Cell-cycle regulation, DNA replication pathways
		Sadikovic *et al.*[42]	6 B	Control sample (osteoblast)	DNA replication network
		Kuijjer *et al.*[43]	84 B	Control samples (MSC, osteoblast)	Apoptosis, signal transduction
		Kansara *et al.*[44]	5 C	Treatment with demethylating agent	WIF1 methylation and downregulation
	miRNA	Jones *et al.*[45]	18 B	Control samples (normal bone)	miR-16 Downregulation, miR-27a association with metastasis
	CN	Kresse *et al.*[25]	9 T/M and their derived xenografts	Tumor sample type	Xenografts are representative for primary tumors although some additional aberrations are observed
		Squire *et al.*[50]	9B	Control samples	Overall high level of aneuploidy, which seems nonrandom. Regions described by three or more studies are gains on 1p, 6p, 8q, 12q and 17p and losses on 2q, 3q, 6q, 10, 13q and 17p
		Man *et al.*[51]	48 B/T/M		
		Atiye *et al.*[52]	22 C/TS/R		
		Yang *et al.*[53]	20 B		
		Kresse *et al.*[54]	36 TS/M/X, 20 C		
		Kuijjer *et al.*[43]	32 B		
		Lockwood *et al.*[55]	22 TS		
		Yen *et al.*[56]	42 TS/R/M/C		
		Smida *et al.*[58]	45 B		
		Pasic *et al.*[59]	27 B		

		Kuijjer *et al.*[43]	32 B	Metastasis/event-free survival	Genomic alterations are prognostic predictors
		Smida *et al.*[58]	45 B		
		Yen *et al.*[56]	23 TS, 14 R/M	Tumor sample type	Identified deletions/amplifications which differ between TS and R/M
		Kresse *et al.*[54]	36 TS/M/X, 20 C	Control samples	Frequent deletion of *LSAMP*
		Yen *et al.*[56]	42 TS/R/M/C		
		Pasic *et al.*[59]	27 B		
		Yang *et al.*[53]	20 B	Control samples	Enrichment of VEGF pathway
Integrative	CN, mRNA	Kuijjer *et al.*[43]	29 B	Control samples (MSC, osteoblast)	Set of 31 candidate drivers enriched in genes with a role in genomic instability
		Lockwood *et al.*[55]	22 TS, 8 X	Control samples (normal tissues)	Amplification and overexpression of cyclin E3
	miRNA, mRNA	Jones *et al.*[45]	14 B	Control samples (normal bone)	Transcriptional regulation, cell cycle control and cancer signaling
		Namløs *et al.*[46]	19 C	Control samples (normal bone)	Pairs of miRNAs with 26 mRNAs
	CN, mRNA, methylation	Sadikovic *et al.*[48]	2 C	Control sample (osteoblast)	Hypomethylation of genes connected to c-Myc
		Sadikovic *et al.*[42]	5 B	Control sample (osteoblast)	RUNX2 amplification and overexpression, *DOCK5* and *TNFRSF10A/D* loss and underexpression, hypomethylation, gain and overexpression of histone cluster 2 genes
		Kresse *et al.*[49]	19 C	Control samples (osteoblast, normal bone)	350 genes with two aberration types, including *RUNX2* and *DLX5* amplification and overexpression

Table 2.1: Overview of genome-wide data analyses in high-grade osteosarcoma. The table gives an overview of single-way and integrative analyses described in this review. For each study, the sample type and sample size is given under Osteosarcoma samples column and the comparison which is made in the bioinformatics analysis, for example, comparison with control tissue, is shown in the Comparison column. Several studies used different sample types in one group. When this was done, these sets are shown in the table as combined into one group as well. Groups of different sample types which have been used in separate analyses are shown as different groups. Not always, it is clear whether naïve tumor biopsies, untreated primary tumor resections or resections of treated primary tumors were used. For such studies, we have used the abbreviation TS (for tumor sample). B: naïve tumor biopsies, T: resections of primary tumors, R: resections of recurrences, M: metastatic resections, C: cell lines, X: xenografts.

Table 2.1. With the purpose to review bioinformatic analyses on osteosarcoma, we only review studies where at least three samples were included, and only refer to articles where robust statistical analyses have been applied.

Single platform analyses of osteosarcoma genome-wide data

Different approaches for single-way analyses

In a typical supervised genome-wide data analysis, significant differences, for example significantly differentially expressed genes/miRNAs or differential methylation, are determined between two or more groups of samples. These groups can exist of different clinical parameters, of tumor samples and their nontumorigenic counterpart or of experimentally induced and noninduced samples as shown in Figure 2.1. Copy number profiling data are analyzed somewhat differently, as copy number profiles of tumor samples do not necessarily have to be compared to their specific nontumorigenic counterparts, but can be compared to, for example, a public reference set, such as HapMap samples (29). Usually, a cutoff for frequency is used to determine whether an amplification or deletion is recurrently present in a specific region. Unsupervised analysis, on the other hand, can give information on quality of the data, and on whether there are certain subgroups within the tumor samples that behave differently.

Each of these distinct ways to analyze genome-wide data has been applied to high-grade osteosarcoma data sets. An overview of these different approaches in osteosarcoma on gene expression, microRNA (miRNA), methylation and copy number data is given in the following paragraphs. Functional verification of the results obtained with these studies will be discussed in a later section of this review.

Genome-wide gene expression data, comparison of clinical parameters

Comparisons between different clinical subgroups of osteosarcoma have resulted in a prediction profile that can classify the main histological subtypes of conventional high-grade osteosarcoma in biopsy material, but also in cell lines and in osteosarcoma xenografts (Chapter 3, (28)). Protein interaction networks illustrated that chondroid matrix-associated proteins were overexpressed in chondroblastic osteosarcoma, whereas NFκB–STAT5 signaling showed higher expression in fibroblastic osteosarcoma. The absence of a specific network for osteoblastic osteosarcoma indicates that the features of the main osteoblast-like cell and of the osteoid matrix are present in tumors of all three main histological subtypes.

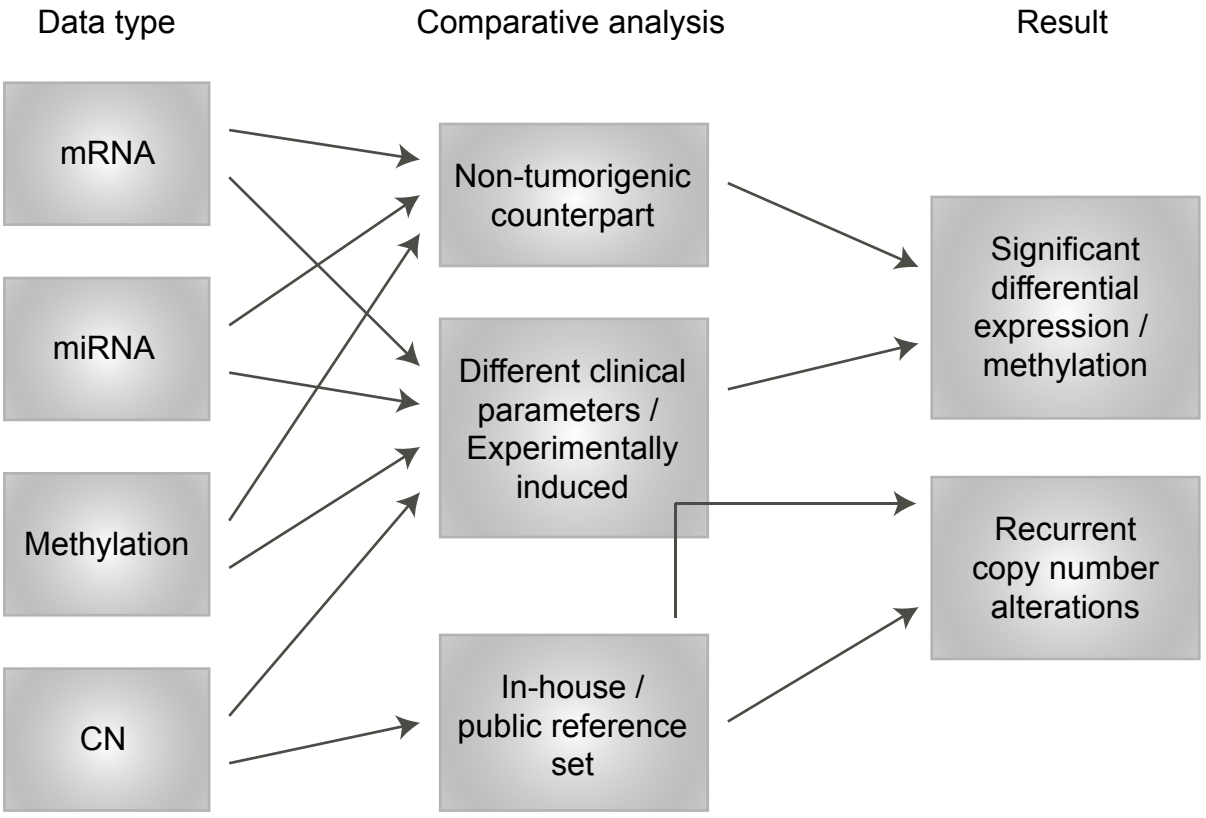


Figure 2.1: Different supervised comparisons in genome-wide data analysis. Flow chart describing single-way bioinformatic analyses that are most typically performed on genome-wide data. For mRNA, miRNA and methylation data analysis, the comparative analysis usually exists of tumor samples versus nontumorigenic counterparts, of different groups of tumor samples, defined by clinical parameters or samples which are experimentally altered compared to samples which are not, although tumor samples of a specific group are also sometimes compared to a pool of all samples (not illustrated in this figure). Copy number data are most often compared to a reference set, which may be an in-house, or a public reference set, and which does not have to consist of the nontumorigenic counterpart of the tumor that is studied. Additional comparative analyses may determine the differences between different subgroups within the samples that are studied. Although for mRNA, miRNA and methylation data, often significant differential expression/methylation is returned by statistical tests, for copy number data researchers mostly look at frequency of the aberration in the studied groups.

A second example of a comparison between different clinical parameters is the comparison of samples with different outcomes in event-free survival or overall survival. It is important to note that when designing an analysis for such a study, a uniform set of clinical follow-up parameters should be employed, instead of directly comparing patients with or without metastases, or patients who are alive or deceased. In one study, differential expression was determined between biopsy material of patients developing metastases within 5 years and patients who did not develop metastases within this time frame. This study demonstrated that, in osteosarcoma, an expression profile associated with macrophages correlated with better overall survival (Chapter 4, (30)). To identify genes playing a role in metastasis, comparisons between osteosarcoma cell lines that can or cannot metastasize upon passaging into mice have also been made. A recent study identified downregulation of *IGFBP5*, or insulin-like growth factor binding protein 5, in the metastatic cell line MG63.2 and in tumors derived from this cell line (31). Interestingly, this gene was also significantly downregulated in our analysis, comparing osteosarcoma biopsies with control tissues (32)). Metastasis progression can be studied by comparing metastatic resections to the primary tumor. This has been performed in one study, where higher expression of genes involved in immunological processes was detected in the metastasis samples (33). This may correlate with our findings that more $CD14^+$ cells are present in metastatic lesions than in pretreatment biopsies (30).

Another important clinical parameter that has been studied in human osteosarcoma is response to chemotherapy, which is predictive for overall survival (34–36). Differentially expressed genes discriminating between good and poor responders to chemotherapy have been detected by different groups, but with little consensus in the gene lists. Most studies did not use robust statistics with correction for multiple testing, a shortcoming that is too often seen in biomedical research (37). When differential expression was determined between poor and good responders in two studies where correction for multiple testing was applied (Chapter 3, (28), and (32)), no significant genes were detected although larger sample sizes and homogeneous data sets were used (17 poor *vs* 8 good responders in Cleton-Jansen *et al.* (32) and 36 poor *vs* 33 good responders in Kuijjer *et al.* (28)). Although these sample sizes are not comparable to what is often used for studying less rare tumor types, the distribution of the nonadjusted p-values did not show any trend for the lower p-values to be more prevalent (Additional Figure 2.1). This indicates that in a comparison between two groups no effect is detected, and increasing sample size will not lead to a significant increase in power (38). A major issue with comparing responders with nonresponders in gene expression analysis is that resistance to chemotherapy may be caused by the alteration of a single gene. A specific gene causing resistance in a subset of samples will not be picked up by a comparison of responders and nonresponders (39).

In human osteosarcoma xenografts, significant differential expression has been detected between good and poor responders to single chemotherapeutic agents (40). A pitfall of this study, however, was that the studied sample set included xenografts derived from

biopsies, resections as well as from metastases. Surviving cells of pretreated tumors are resistant to chemotherapy. Thus, the differences in gene expression between poor and good responders to these chemotherapeutic agents may actually reflect an effect of presurgery therapy. It was indeed demonstrated that xenografts of these implanted pretreated tumors often responded poorly to multiple chemotherapeutic agents (41).

Genome-wide gene expression data, comparison with control tissues

mRNA expression levels in osteosarcoma samples can also be compared to expression in control tissues. The control tissues that have been used for this purpose are normal bone, osteoblastoma, osteoblasts, MSCs, or, for example, a pool of different cell lines. One comparison of high-grade osteosarcoma biopsy specimens with control samples is described in Cleton-Jansen *et al.* (32), who made different comparisons of 25 osteosarcoma biopsies with five osteoblastomas, with five MSCs and with five osteoblast cultures. Gene set enrichment detected cell-cycle regulation and DNA replication pathways as the most significantly affected pathways in osteosarcoma. A DNA replication network was also identified in an analysis of gene expression microarrays of six osteosarcoma biopsies as compared to one osteoblast culture although a caveat of this study is the small sample size of the control set ($n = 1$) (42). A larger set of osteosarcoma biopsies ($n = 84$) was compared to 12 MSCs and separately with three osteoblast cultures (Chapter 7, (43)). Intersection of the differentially expressed genes in both analyses identified antigen processing and presentation as well as angiogenesis as significantly different between tumor samples and control cell lines, most probably because of the amount of stroma present in the tumor samples. In addition, altered apoptosis and signal transduction were detected.

Genome-wide gene expression data, experimentally induced differences

We give a final example of genome-wide gene expression analyses in osteosarcoma, which is experimentally induced differential expression. This is, for example, reported in the study by Kansara it et al. (44), who compared a set of five human osteosarcoma cells treated with a demethylating agent to untreated cells, after having shown that demethylating agents can induce growth arrest and differentiation in osteosarcoma. The list of candidate genes was then filtered for expression in human osteoblasts and loss of expression in primary osteosarcomas. This screen identified *WIF1*, a Wnt inhibitory factor, as a candidate tumor suppressor in osteosarcoma.

microRNA expression data

Several studies have been published, describing miRNA microarray data analysis on osteosarcoma tissues or cell lines as compared to osteoblasts or normal bone, but in most studies no robust statistics were applied. Jones *et al.* (45) and Namløs *et al.* (46) published the only miRNA microarray studies in which false discovery rate corrections were applied. In the article by Jones *et al.* (45), miRNA expression was compared between 18 osteosarcoma resections or biopsies and 12 normal bone samples, which lead to the detection of a downregulated tumor suppressive miRNA and of a prometastatic miRNA (these miRNAs will be discussed in the Integrative analyses section). Namløs *et al.* (46) compared miRNA expression in 19 osteosarcoma cell lines with expression in normal bone ($n = 4$) and integrated these results with mRNA expression data. Results from this study will therefore be discussed in the Integrative analyses section. Sarver *et al.* (47) published an online accessible Sarcoma miRNA expression database (S-MED), which includes 15 osteosarcoma samples and six normal bone samples.

Genome-wide methylation data

Only three studies have been published so far on genome-wide methylation in high-grade osteosarcoma (42, 48, 49). These studies describe an integrative analysis with different data types, without presenting conclusions on specific genes, or on results obtained with gene set enrichment on single-way methylation analyses although Kresse *et al.* (49) found overall more hypermethylation in osteosarcoma cell lines than hypomethylation. We will discuss the results from these studies under the Integrative analysis section of this review.

Genomic copy number data

The genomic instability of high-grade osteosarcoma, which is more pronounced in this tumor than in many other tumor types, hampers the identification of specific genomic regions. Several array comparative genomic hybridization (aCGH) studies (25, 50–55) and single-nucleotide polymorphism (SNP) microarray studies (43, 56–59) on osteosarcoma specimens have been published. Copy number profiles clearly show that high-grade osteosarcoma samples are characterized by a high level of aneuploidy, and that there is heterogeneity between different tumor samples. There is a general consensus about copy number alterations for some regions, such as gains on chromosome arms 6p, 8q and 17p, which have been detected by classical karyotyping and conventional CGH as well (2, 60), but it is difficult to directly compare studies as the definition of a recurrent alteration varies.

In three separate studies, a focal deletion of the region 3q13.31, which harbors a putative tumor suppressor gene, *LSAMP*, was detected (54, 56, 59). siRNA-mediated silencing of *LSAMP* promoted proliferation of normal osteoblasts (59), and low expression

of the gene was associated with poor overall survival in one of these studies (54). Gene set enrichment on an aCGH study showed an enrichment of amplified genes of the VEGF signaling pathway of which *VEGFA* amplification also correlated with poor prognosis, showing that gene set enrichment on copy number data can identify pathways associated with tumorigenesis (53).

Copy number profiles of osteosarcoma cell lines roughly resemble profiles of tumor biopsies, but show an increased overall aneuploidy (Kuijjer *et al.*, *unpublished data*) and increased expression of genomic instability genes (Chapter 7, (43)). Also, as described above, genomic profiles of xenografts are highly similar to primary tumors although some deviations may occur owing to additional genetic alterations during passaging, or owing to general tumor progression (25). In some data sets, specific copy number alterations have not been detected for different clinical groups of interest (43) although both a high degree of genomic alterations and a loss of heterozygosity were found to be associated with poor event-free survival (43, 58). Yen *et al.* (56) found that specific aberrations were more frequent in recurrences and metastases than in primary tumors—deletion of 6q14.1-q22.31 and 8p23.2-p12 and amplification of 8q21.12-q24.3 and 17p12—and vice versa—Xp11.22 gain and 13q31.3 deletion (56).

Integrative analyses

For high-grade osteosarcoma, integration of different data types is of specific importance. Integrative analysis can narrow down the large lists of significantly affected genes to a gene list containing the major tumor driver genes. An integrative approach on copy number and gene expression data, for example, typically returns a more specific list of driver genes because passenger- and tissue-specific genes will be largely eliminated (61). Different methods exist for the integration of different types of data. Figure 2.2 shows an overview of direct dependencies between copy number, methylation, miRNA and mRNA data. Comparison of data can be performed nonpaired or paired, and by determining

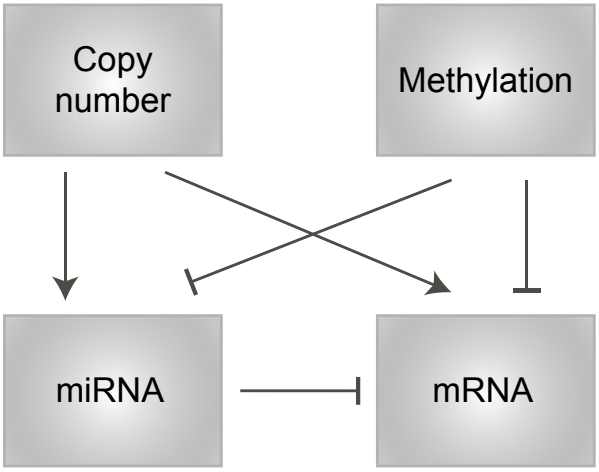


Figure 2.2: Flow chart showing direct dependencies between different data types, which can be utilized for the interpretation of integrative analyses. Arrow-headed and bar-headed lines show positive and negative influences, respectively. DNA copy number positively affects miRNA and mRNA copies, whereas miRNA expression can cause downregulation of target mRNAs, and DNA methylation can inhibit transcription.

correlation or cooccurrence.

Cooccurring genomic alterations and gene expression changes have been recently determined to identify putative driver genes in high-grade osteosarcoma (Chapter 7, (43)). A paired integrative analysis of 29 pretreatment biopsies returned a list of 31 genes with recurrence frequency of at least 35%, which showed an overall significant upregulation as compared to control cell lines in case of a gain, and downregulation in case of a deletion. Genes affecting genomic stability were overrepresented, which may point to a role of this process in osteosarcoma. Nonpaired analysis on the same series, but extended with more cases in both the SNP and the gene expression data sets resulted in a smaller set of significantly affected genes, with substantial overlap with the list of genes detected by the paired analysis, thereby showing that the paired analysis was more powerful on this data set. This is especially of interest for the data analysis of osteosarcoma pretreatment biopsies because these samples are rare. By performing a paired analysis, fewer samples can be used. Nonpaired integrative analysis of high-level amplifications in 22 osteosarcoma specimens with gene expression data of eight osteosarcoma xenografts as compared to 19 normal tissue controls identified 43 genes with high-level amplification and overexpression in osteosarcoma. *CCNE1*, the gene encoding for cyclin E1, showed correlation of copy number levels and gene expression in an additional panel of ten osteosarcoma cell lines, and therefore could play an oncogenic role in osteosarcoma (55).

miRNA expression data can be integrated with mRNA expression data to determine whether the miRNAs of interest affect mRNA expression of their target genes. This is generally performed by correlation of expression levels, as was performed by Baumhoer *et al.*, (62) Namløs *et al.* (46) and Jones *et al.* (45) (discussed above). The latter subsequently performed pathway analysis on target genes of the detected differentially expressed miRNAs, which illustrated the effects of these miRNAs on transcriptional regulation, cell-cycle control and known cancer signaling pathways (45). In the study by Namløs *et al.* (46), cell line miRNA data were integrated with mRNA targets which were significant in both osteosarcoma pretreatment biopsies and cell lines. Among the inversely correlated miRNA/mRNA pairs, miRNAs regulating *TGFBR2*, *IRS1*, *PTEN* and PI3K subunits were detected. Methylation data are also typically integrated with mRNA expression data to evaluate the effect of the methylation on gene expression, but few studies described two-way comparisons of methylation and mRNA microarray data in osteosarcoma. Kresse *et al.* (49) detected hypermethylation and underexpression of chemokine ligand 5 (*CXCL5*) by two-way comparison in both osteosarcoma cell lines and tumor samples.

Integration of more than two different data types is reported by Sadikovic *et al.* (48) in two articles, where copy number, methylation and gene expression data were integrated. In one of these articles, the authors described cooccurrent epigenetic, genomic and gene expression changes in two osteosarcoma cell lines as compared to an osteoblast culture, and detected a region of gain on chromosome 8q encompassing the *c-MYC* oncogene,

which was also detected in a network analysis, confirming overexpression and hypomethylation of genes connected to *c-MYC*. In the second article, the authors used the same integrative approach to perform a three-way analysis on five osteosarcoma pretreatment biopsies, and to compare gene regulation networks of single-way analyses including more samples. In this way, a number of candidate genes were characterized, including *RUNX2*, a transcription factor involved in osteoblastic differentiation (42). A shortcoming of both studies, however, is that as a control for methylation and mRNA expression in osteosarcoma, material from only one osteoblastic culture was used. Another integrative analysis on copy number, methylation and mRNA data reported 350 genes, showing two types of aberrations (*e.g.* gain and overexpression, or hypermethylation and underexpression). This set of genes was enriched in genes with a function in skeletal system development and extracellular matrix remodeling, such as *RUNX2* and *DLX5* (49).

Translating bioinformatics into functional studies

Functional validation of candidate genes

Several of the candidate tumor suppressor genes and oncogenes that have been identified with microarray studies have been functionally validated. *IGFBP5* was significantly downregulated in metastatic cell lines and derivative tumors as compared to nonmetastatic cell lines, and also showed lower protein expression in metastatic lesions than in primary tumor samples of osteosarcoma patients. The effects of overexpression or knockdown of *IGFBP5* on cell proliferation, migration, wound healing and invasion confirmed the role of this IGF-binding protein in preventing metastasis, which was furthermore validated in a xenograft model (31).

The candidate tumor suppressor gene *WIF1* was found to regulate differentiation and suppress cell growth *in vitro*. *WIF1* knockout mice developed radiation-induced osteosarcoma earlier than their littermate controls (44). From miRNA expression profiling studies, miR-16 was validated as a tumor suppressive miRNA, whereas miR-27a was validated as a prometastatic miRNA, using colony formation assays, and wound healing and invasion assays, respectively. Overexpression of these miRNAs *in vivo* resulted in smaller tumors for miR-16, and in higher numbers of pulmonary metastases for miR-27a (45).

Functional validation of pathway activity and enriched gene sets

Pathways important in the development of bone biology have been returned from gene expression analysis as compared to controls. Genes upstream canonical Wnt signaling were, for example, found to be downregulated as compared to osteoblasts (32). A subsequent functional study, where nuclear β-catenin staining was determined on osteosarcoma biopsies, and Wnt luciferase activity and mRNA expression of the specific downstream Wnt

target gene Axin2 were measured in cell lines, illustrated that canonical Wnt signaling is indeed often downregulated in osteosarcoma (63). Loss of canonical Wnt signaling causes failure to commit to differentiation of MSCs, as has been reported in malignant fibrous histiocytoma (also undifferentiated pleomorphic sarcoma), which could be reprogrammed by re-establishing Wnt signaling (64). Also in osteosarcoma, reactivation of the Wnt signaling pathway with a GSK3β inhibitor triggered a more differentiated phenotype, or a reduced proliferation capacity, depending on the osteosarcoma cell line (63).

These results seem contradictive to the finding that *WIF1* can inhibit cell growth and increase differentiation in osteosarcoma cells (44). A possible explanation for this discrepancy is that *WIF1* inhibits both canonical and noncanonical Wnt signaling (65), whereas GSK3β also plays a role in additional signal transduction pathways, such as NFκB signaling (66). However, the role of Wnt signaling remains contradictory, as this pathway was recently described to be active in multiple sarcoma subtypes, which also included osteosarcoma (67). The use of different methods to assess active Wnt signaling may be the cause for the discrepancies between these studies.

TGF-β/BMP signaling was found to be affected in osteosarcoma by pathway analysis on mRNA expression data. Activity of these pathways was validated by immunohistochemistry of phosphorylated Smad1 and Smad2, and nuclear staining of these intracellular effectors was detected in 70% of all osteosarcoma samples. Cases with very low or absent phosphorylated Smad2 had worse overall survival. *In vitro* pathway modulation did not affect proliferation or differentiation, but lower TGFβ/BMP activity might affect the prevention of metastasis in these patients (68).

The macrophage signature that was prominent upon comparing mRNA profiles of metastatic and nonmetastatic osteosarcoma was confirmed by qPCR and immunohistochemistry, and it was shown in additional cohorts that the sum of M1 and M2 types of macrophages was predictive for better overall survival (Chapter 4, (30)). Treating patients with macrophage-activating agents may reduce metastases of osteosarcoma (69). This is corroborated by clinical trials in dogs and humans, where treatment with mifamurtide, a macrophage-activating agent, has been reported to positively affect overall survival (70, 71).

Conclusions and future directions

In this review, we have presented and discussed the results of studies on high-grade osteosarcoma material using bioinformatic analysis on microarray data of three or more samples. Although studying such a very heterogeneous and genomically unstable tumor remains challenging, and sample sizes are often small owing to the rarity of the disease, structured microarray data analysis has provided interesting results and has given further insight into the biology and progression of osteosarcoma. This information could not have been obtained from functional studies only. Studying copy number aberrations,

differential expression, and epigenetics in a genome-wide manner and subsequent integration leads to new hypotheses regarding tumor development and progression, which can subsequently be validated in functional studies. This provides a motivation to take the study of high-grade osteosarcoma to the next level, and to analyze this tumor into further detail using Next Generation Sequencing methods, such as whole-genome, exome or transcriptome sequencing. Whole-genome sequencing has recently been performed in a study of different cancer types, which showed that a subset of osteosarcomas (three out of nine) undergo chromothripsis—a single catastrophic genomic instability event, resulting in hundreds of genomic rearrangements (72). This may explain the sudden onset of osteosarcoma and the complexity and heterogeneity of the osteosarcoma genome. Next Generation Sequencing will provide us with many forms of new information. In addition to copy number changes, mutations, translocations, unannotated genes, splicing variants, and so on, can be detected in a high-throughput manner. Transcriptome sequencing exhibits higher sensitivity and increased dynamic range than mRNA expression microarray data, thereby providing higher power for the detection of differential gene expression (73). Now that the first Next Generation Sequencing studies including large numbers of high-grade osteosarcoma are ongoing or being planned, it is important to reflect on the previous genome-wide studies in osteosarcoma. When keeping in mind the lessons we have learned on study design in microarray data analysis—using a sufficient amount of samples, defining homogeneous groups, and analyzing the data with robust statistics—we will be given new opportunities in unraveling the biology of this complex disease and in providing future clinical trials with robust data to incorporate into novel therapeutic strategies.

References

[1] Fletcher JA, Gebhardt MC, Kozakewich HP. Cytogenetic aberrations in osteosarcomas: nonrandom deletions, rings, and double-minute chromosomes. *Cancer Genetics and Cytogenetics*, 77(1):81–88, 1994.

[2] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[3] Helman LJ, Meltzer P. Mechanisms of sarcoma development. *Nature Reviews Cancer*, 3(9):685–694, 2003.

[4] Pakos EE, Nearchou AD, Grimer RJ, Koumoullis HD, *et al.* Prognostic factors and outcomes for osteosarcoma: an international collaboration. *European Journal of Cancer*, 45(13):2367–2375, 2009.

[5] Lewis IJ, Nooij MA, Whelan J, Sydes MR, *et al.* Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup. *Journal of the National Cancer Institute*, 99(2):112–128, 2007.

[6] Eselgrim M, Grunert H, Kühne T, Zoubek A, *et al.* Dose intensity of chemotherapy for osteosarcoma and outcome in the Cooperative Osteosarcoma Study Group (COSS) trials. *Pediatric Blood & Cancer*, 47(1):42–50, 2006.

[7] Anninga JK, Gelderblom H, Fiocco M, Kroep JR, *et al.* Chemotherapeutic adjuvant treatment for osteosarcoma: where do we stand? *European Journal of Cancer*, 47(16):2431–2445, 2011.

[8] Ding L, Ley TJ, Larson DE, Miller CA, *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, 481(7382):506–510, 2012.

[9] Mohseny AB. Bone: conventional osteosarcoma. `AtlasGeneticsOncology.org/Tumors/ConvOsteoID5344.html`. Accessed: 09/30/2010.

[10] Hauben EI, Weeden S, Pringle J, van Marck EA, *et al.* Does the histological subtype of high-grade central osteosarcoma influence the response to treatment with chemotherapy and does it affect overall survival? A study on 570 patients of two consecutive trials of the European Osteosarcoma Intergroup. *European Journal of Cancer*, 38(9):1218, 2002.

[11] Hauben EI, Bielack S, Grimer R, Jundt G, *et al.* Clinico-histologic parameters of osteosarcoma patients with late relapse. *European Journal of Cancer*, 42(4):460–466, 2006.

[12] Mohseny AB, Hogendoorn PCW. Concise review: mesenchymal tumors: when stem cells go mad. *Stem Cells*, 29(3):397–403, 2011.

[13] Hogendoorn PCW, Bovée JVMG, Karperien M, Cleton-Jansen AM. Skeletogenesis: genetics. In Cooper DN, editor, Nature encyclopedia of the human genome, 306–313. Nature Publishing Group, 2003.

[14] Tolar J, Nauta AJ, Osborn MJ, Panoskaltsis Mortari A, *et al.* Sarcoma derived from cultured mesenchymal stem cells. *Stem Cells*, 25(2):371–379, 2006.

[15] Mohseny AB, Szuhai K, Romeo S, Buddingh EP, *et al.* Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of Cdkn2. *The Journal of Pathology*, 219(3):294–305, 2009.

[16] Mohseny AB, Xiao W, Carvalho R, Spaink HP, *et al.* An osteosarcoma zebrafish model implicates Mmp-19 and Ets-1 as well as reduced host immune response in angiogenesis and migration. *The Journal of Pathology*, 227(2):245–253, 2012.

[17] Mohseny AB, Hogendoorn PCW, Cleton-Jansen AM. Osteosarcoma models: from cell lines to zebrafish. *Sarcoma*, 2012:417271, 2012.

[18] Weinberg RA. The cells forming cancer cell lines develop without heterotypic interactions and deviate from the behavior of cells within human tumors. In The Biology of Cancer, volume 255, 536–537. Garland Science, 2007.

[19] Greshock J, Nathanson K, Martin AM, Zhang L, *et al.* Cancer cell lines as genetic models of their parent histology: analyses based on array comparative genomic hybridization. *Cancer Research*, 67(8):3594–3600, 2007.

[20] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[21] Cleton-Jansen AM, Buerger H, Hogendoorn PCW. Central high-grade osteosarcoma of bone: diagnostic and genetic considerations. *Current Diagnostic Pathology*, 11(6):390–399, 2005.

[22] Overholtzer M, Rao PH, Favis R, Lu XY, *et al.* The presence of p53 mutations in human osteosarcomas correlates with high levels of genomic instability. *Proceedings of the National Academy of Sciences*, 100(20):11547–11552, 2003.

[23] Mohseny AB, Machado I, Cai Y, Schaefer KL, *et al.* Functional characterization of osteosarcoma cell lines provides representative models to study the human disease. *Laboratory Investigation*, 91(8):1195–1205, 2011.

[24] Mayordomo E, Machado I, Giner F, Kresse SH, *et al.* A tissue microarray study of osteosarcoma: histopathologic and immunohistochemical validation of xenotransplanted tumors as preclinical models. *Applied Immunohistochemistry & Molecular Morphology*, 18(5):453–461, 2010.

[25] Kresse SH, Meza-Zepeda LA, Machado I, Llombart-Bosch A, *et al.* Preclinical xenograft models of human sarcoma show nonrandom loss of aberrations. *Cancer*, 118(2):558–570, 2012.

[26] Rüther U, Komitowski D, Schubert FR, Wagner EF. c-fos expression induces bone tumors in transgenic mice. *Oncogene*, 4(7):861–865, 1989.

[27] Walkley CR, Qudsi R, Sankaran VG, Perry JA, *et al.* Conditional mouse osteosarcoma, dependent on p53 loss and potentiated by loss of Rb, mimics the human disease. *Genes & Development*, 22(12):1662–1676, 2008.

[28] Kuijjer ML, Namløs HM, Hauben EI, Machado I, *et al.* mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts. *BMC Medical Genomics*, 4(1):66, 2011.

[29] Gibbs RA, Belmont JW, Hardenbol P, Willis TD, *et al.* The international HapMap project. *Nature*, 426(6968):789–796, 2003.

[30] Buddingh EP, Kuijjer ML, Duim RAJ, Bürger H, *et al.* Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. *Clinical Cancer Research*, 17(8):2110–2119, 2011.

[31] Su Y, Wagner ER, Luo Q, Huang J, *et al.* Insulin-like growth factor binding protein 5 suppresses tumor growth and metastasis of human osteosarcoma. *Oncogene*, 30(37):3907–3917, 2011.

[32] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

[33] Namløs HM, Kresse SH, Müller CR, Henriksen J, *et al.* Global gene expression profiling of human osteosarcomas reveals metastasis-associated chemokine pattern. *Sarcoma*, 2012:639038, 2012.

[34] Bacci G, Ferrari S, Delepine N, Bertoni F, *et al.* Predictive factors of histologic response to primary chemotherapy in osteosarcoma of the extremity: study of 272 patients preoperatively treated with high-dose methotrexate, doxorubicin, and cisplatin. *Journal of Clinical Oncology*, 16(2):658–663, 1998.

[35] Huvos AG. Bone tumors: diagnosis, treatment and prognosis. WB Saunders Company, 1991.

[36] Bacci G, Longhi A, Versari M, Mercuri M, *et al.* Prognostic factors for osteosarcoma of the extremity treated with neoadjuvant chemotherapy. *Cancer*, 106(5):1154–1161, 2006.

[37] Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *Journal of the National Cancer Institute*, 99(2):147–157, 2007.

[38] van Iterson M, Pedotti P, Hooiveld GJEJ, den Dunnen JT, *et al.* Relative power and sample size analysis on gene expression profiling data. *BMC Genomics*, 10(1):439, 2009.

[39] Borst P, Wessels L. Do predictive signatures really predict response to cancer chemotherapy? *Cell Cycle*, 9(24):4836–4840, 2010.

[40] Bruheim S, Xi Y, Ju J, Fodstad O. Gene expression profiles classify human osteosarcoma xenografts according to sensitivity to doxorubicin, cisplatin, and ifosfamide. *Clinical Cancer Research*, 15(23):7161–7169, 2009.

[41] Bruheim S, Bruland OS, Breistol K, Maelandsmo GM, *et al.* Human osteosarcoma xenografts and their sensitivity to chemotherapy. *Pathology & Oncology Research*, 10(3):133–141, 2004.

[42] Sadikovic B, Yoshimoto M, Chilton-MacNeill S, Thorner P, *et al.* Identification of interactive networks of gene expression associated with osteosarcoma oncogenesis by integrated molecular profiling. *Human Molecular Genetics*, 18(11):1962–1975, 2009.

[43] Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, *et al.* Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes, Chromosomes and Cancer*, 51(7):696–706, 2012.

[44] Kansara M, Tsang M, Kodjabachian L, Sims NA, *et al.* Wnt inhibitory factor 1 is epigenetically silenced in human osteosarcoma, and targeted disruption accelerates osteosarcomagenesis in mice. *The Journal of Clinical Investigation*, 119(4):837–851, 2009.

[45] Jones KB, Salah Z, Del Mare S, Galasso M, *et al.* miRNA signatures associate with pathogenesis and progression of osteosarcoma. *Cancer Research*, 72(7):1865–1877, 2012.

[46] Namløs HM, Meza-Zepeda LA, Barøy T, Østensen IHG, *et al.* Modulation of the Osteosarcoma Expression Phenotype by MicroRNAs. *PloS One*, 7(10):e48086, 2012.

[47] Sarver AL, Phalak R, Thayanithy V, Subramanian S. S-MED: sarcoma microRNA expression database. *Laboratory Investigation*, 90(5):753–761, 2010.

[48] Sadikovic B, Yoshimoto M, Al-Romaih K, Maire G, *et al.* In vitro analysis of integrated global high-resolution DNA methylation profiling with genomic imbalance and gene expression in osteosarcoma. *PLoS One*, 3(7):e2834, 2008.

[49] Kresse SH, Rydbeck H, Skårn M, Namløs HM, *et al.* Integrative analysis reveals relationships of genetic and epigenetic alterations in osteosarcoma. *PloS One*, 7(11):e48262, 2012.

[50] Squire JA, Pei J, Marrano P, Beheshti B, *et al.* High-resolution mapping of amplifications and deletions in pediatric osteosarcoma by use of CGH analysis of cDNA microarrays. *Genes, Chromosomes and Cancer*, 38(3):215–225, 2003.

[51] Man TK, Lu XY, Jaeweon K, Perlaky L, *et al.* Genome-wide array comparative genomic hybridization analysis reveals distinct amplifications in osteosarcoma. *BMC Cancer*, 4(1):45, 2004.

[52] Atiye J, Wolf M, Kaur S, Monni O, *et al.* Gene amplifications in osteosarcoma-CGH microarray analysis. *Genes, Chromosomes and Cancer*, 42(2):158–163, 2005.

[53] Yang J, Yang D, Sun Y, Sun B, *et al.* Genetic amplification of the vascular endothelial growth factor (VEGF) pathway genes, including VEGFA, in human osteosarcoma. *Cancer*, 117(21):4925–4938, 2011.

[54] Kresse SH, Ohnstad HO, Paulsen EB, Bjerkehagen B, *et al.* LSAMP, a novel candidate tumor suppressor gene in human osteosarcomas, identified by array comparative genomic hybridization. *Genes, Chromosomes and Cancer*, 48(8):679–693, 2009.

[55] Lockwood WW, Stack D, Morris T, Grehan D, *et al.* Cyclin E1 is amplified and overexpressed in osteosarcoma. *The Journal of Molecular Diagnostics*, 13(3):289–296, 2011.

[56] Yen CC, Chen WM, Chen TH, Chen WYK, *et al.* Identification of chromosomal aberrations associated with disease progression and a novel 3q13.31 deletion involving LSAMP gene in osteosarcoma. *International Journal of Oncology*, 35(4):775–788, 2009.

[57] Kresse SH, Szuhai K, Barragan-Polania AH, Rydbeck H, *et al.* Evaluation of high-resolution microarray platforms for genomic profiling of bone tumours. *BMC Research Notes*, 3(1):223, 2010.

[58] Smida J, Baumhoer D, Rosemann M, Walch A, *et al.* Genomic alterations and allelic imbalances are strong prognostic predictors in osteosarcoma. *Clinical Cancer Research*, 16(16):4256–4267, 2010.

[59] Pasic I, Shlien A, Durbin AD, Stavropoulos DJ, *et al.* Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Research*, 70(1):160–171, 2010.

[60] Lau CC, Harris CP, Lu XY, Perlaky L, *et al.* Frequent amplification and rearrangement of chromosomal bands 6p12-p21 and 17p11.2 in osteosarcoma. *Genes, Chromosomes and Cancer*, 39(1):11–21, 2003.

[61] Lee H, Kong SW, Park PJ. Integrative analysis reveals the direct and indirect interactions between DNA copy number aberrations and gene expression changes. *Bioinformatics*, 24(7):889–896, 2008.

[62] Baumhoer D, Zillmer S, Unger K, Rosemann M, *et al.* MicroRNA profiling with correlation to gene expression revealed the oncogenic miR-17-92 cluster to be up-regulated in osteosarcoma. *Cancer Genetics*, 205(5):212–219, 2012.

[63] Cai Y, Mohseny AB, Karperien M, Hogendoorn PCW, *et al.* Inactive Wnt/β-catenin pathway in conventional high-grade osteosarcoma. *The Journal of Pathology*, 220(1):24–33, 2010.

[64] Matushansky I, Hernando E, Socci ND, Mills JE, *et al.* Derivation of sarcomas from mesenchymal stem cells via inactivation of the Wnt pathway. *Journal of Clinical Investigation*, 117(11):3248–3257, 2007.

[65] Malinauskas T, Aricescu AR, Lu W, Siebold C, *et al.* Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1. *Nature Structural & Molecular Biology*, 18(8):886–893, 2011.

[66] Tang QL, Xie XB, Wang J, Chen Q, *et al.* Glycogen synthase kinase-3β, NF-κB signaling, and tumorigenesis of human osteosarcoma. *Journal of the National Cancer Institute*, 104(10):749–763, 2012.

[67] Vijayakumar S, Liu G, Rus IA, Yao S, *et al.* High-frequency canonical Wnt activation in multiple sarcoma subtypes drives proliferation through a TCF/β-catenin target gene, CDC25A. *Cancer Cell*, 19(5):601–612, 2011.

[68] Mohseny AB, Cai Y, Kuijjer ML, Xiao W, *et al.* The activities of Smad and Gli mediated signalling pathways in high-grade conventional osteosarcoma. *European Journal of Cancer*, 48(18):3429–3438, 2012.

[69] Cleton-Jansen AM, Buddingh EP, Lankester AC. Immunotherapy: is it different for sarcomas? *Oncoimmunology*, 1(2):255–257, 2012.

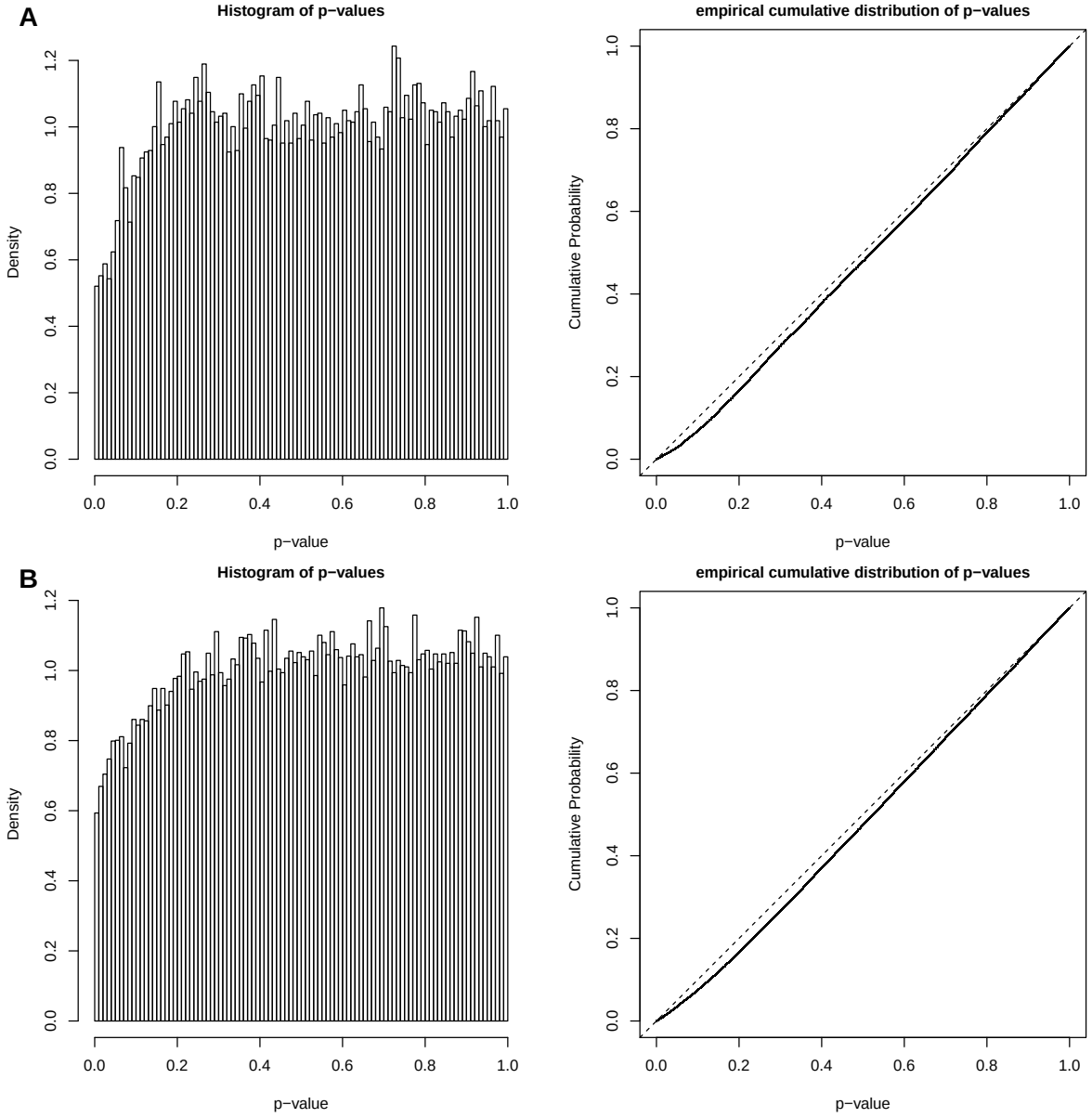
[70] Kurzman ID, MacEwen EG, Rosenthal RC, Fox LE, *et al.* Adjuvant therapy for osteosarcoma in dogs: results of randomized clinical trials using combined liposome-encapsulated muramyl tripeptide and cisplatin. *Clinical Cancer Research*, 1(12):1595–1601, 1995.

[71] Meyers PA, Schwartz CL, Krailo MD, Healey JH, *et al.* Osteosarcoma: the addition of muramyl tripeptide to chemotherapy improves overall survival—a report from the Children's Oncology Group. *Journal of Clinical Oncology*, 26(4):633–638, 2008.

[72] Stephens PJ, Greenman CD, Fu B, Yang F, *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*, 144(1):27–40, 2011.

[73] Oshlack A, Robinson MD, Young MD. From RNA-seq reads to differential expression results. *Genome Biology*, 11(12):220, 2010.

Additional Figures



Additional Figure 2.1: This figure illustrates a histogram of nonadjusted, moderated p-values and the empirical cumulative distribution of p-values for the studies of *A*, Cleton-Jansen *et al.* (32) and *B*, Kuijjer *et al.* (28), both describing no significant difference in mRNA expression in pretreatment biopsies of patients with poor versus good response to chemotherapy. The figures were generated using Bioconductor package *SSPA* (38).

Chapter 3

mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts

This chapter is based on the publication: Kuijjer ML, Namløs HM, Hauben EI, Machado I, Kresse SH, Serra M, Llombart-Bosch A, Hogendoorn PCW, Meza-Zepeda LA, Myklebost O, Cleton-Jansen AM. *BMC Med Genomics.* 2011 Sep 20;4:66

Abstract

Background: Conventional high-grade osteosarcoma is a primary malignant bone tumor, which is most prevalent in adolescence. Survival rates of osteosarcoma patients have not improved significantly in the last 25 years. Aiming to increase this survival rate, a variety of model systems are used to study osteosarcomagenesis and to test new therapeutic agents. Such model systems are typically generated from an osteosarcoma primary tumor, but undergo many changes due to culturing or interactions with a different host species, which may result in differences in gene expression between primary tumor cells, and tumor cells from the model system. We aimed to investigate whether gene expression profiles of osteosarcoma cell lines and xenografts are still comparable to those of the primary tumor.

Methods: We performed genome-wide mRNA expression profiling on osteosarcoma biopsies ($n = 76$), cell lines ($n = 13$), and xenografts ($n = 18$). Osteosarcoma can be subdivided into several histological subtypes, of which osteoblastic, chondroblastic, and fibroblastic osteosarcoma are the most frequent ones. Using nearest shrunken centroids classification, we generated an expression signature that can predict the histological subtype of osteosarcoma biopsies.

Results: The expression signature, which consisted of 24 probes encoding for 22 genes, predicted the histological subtype of osteosarcoma biopsies with a misclassification error of 15%. Histological subtypes of the two osteosarcoma model systems, *i.e.* osteosarcoma cell lines and xenografts, were predicted with similar misclassification error rates (15% and 11%, respectively).

Conclusions: Based on the preservation of mRNA expression profiles that are characteristic for the histological subtype we propose that these model systems are representative for the primary tumor from which they are derived.

Background

Conventional high-grade osteosarcoma is the most frequent primary malignant bone tumor, with a peak occurrence in children and adolescents and a second peak in patients older than 40 years. It is a highly genetically unstable tumor, of which karyotypes often show aneuploidy, high level amplification and deletion, and translocations (1). No precursor lesion is known, although part of the osteosarcomas in patients over 40 years is secondary, and is induced by radiation, chemicals, or by an underlying history of Paget's disease of bone (2). The leading cause of death of osteosarcoma patients are distant metastases, which despite aggressive chemotherapy regimens still develop in approximately 45% of all patients (3). Overall survival of osteosarcoma patients has increased from 10–20% before the introduction of preoperative chemotherapy in the 1970s, to about 60% (4). However, survival has reached a plateau, and treating with higher doses of chemotherapy does not lead to better overall survival (5).

Osteosarcoma is a heterogeneous tumor type, which can be subdivided into various subtypes (6). Conventional high-grade osteosarcoma is the most common subtype, and can be further subdivided in different histological subtypes, of which osteoblastic (50%), chondroblastic (25%), and fibroblastic osteosarcoma (25%) are the most frequent ones. Other subtypes of conventional high-grade osteosarcoma, such as chondromyxoid fibroma-like, clear cell, epitheliod, sclerosing, and giant cell rich osteosarcoma, are extremely rare (2). Often, osteosarcoma tissue contains a mixture of morphologically differing cell types, and the classification is based on the most dominant type (7). The three main histological subtypes have different survival profiles. Patients with fibroblastic osteosarcoma have a significantly better response to preoperative chemotherapy, which is a known predictor for improved survival, than do osteoblastic osteosarcoma patients (8). Although patients with chondroblastic osteosarcoma are relatively poor responders to preoperative chemotherapy (7, 9), which is probably caused by the impermeability of the chondroid areas of the tumor, there is a trend for these patients to have better 5-year survival profiles (7), but also a higher risk for late relapse (10).

The search for new (targeted) therapies to treat osteosarcoma is ongoing (11). Because the disease is relatively rare, large efforts need to be done in order to collect a considerable amount of patient samples. Moreover, material is usually scarce due to necrosis in resections of the primary tumor, which is mostly present in tumors of patients who respond fairly well to neoadjuvant chemotherapy. No necrosis is present in prechemotherapy biopsies, but these are often very small and are not readily available for research because they are needed for diagnosis. Because of these limitations, model systems are widely used to study osteosarcomagenesis and for preclinical testing of candidate drugs. Osteosarcoma cell lines, especially SAOS-2 and U-2 OS are frequently used as model systems, remarkably not only to study osteosarcoma, but all types of *in vitro* cell biological processes, as these cell lines grow fast and are relatively easy to transfect. Recently, the EuroBoNeT (`www.eurobonet.eu`) osteosarcoma panel of 19 cell lines was characterized, which allows us to study osteosarcoma in a high-throughput manner (12). This panel of osteosarcoma cell lines has been shown to resemble osteosarcoma phenotypically and functionally (13). Other established model systems include xenografts from primary tumors or osteosarcoma cell lines in immunodeficient nude mice, which subsequently develop into tumors resembling osteosarcoma (13–15). Osteosarcomagenesis can also be induced in mice by radiation or orthotopically implanting chemical carcinogens (16). We have previously shown that DNA copy number profiles of xenografts resemble those of the corresponding primary tumor, although some significant changes for osteosarcoma were observed (15).

Established cancer cell lines are often thought not to be representative for the originating primary tumor. Since there could have been a selection for their propensity to grow in culture, they lack interaction with stroma and may have acquired additional mutations in culture (17). Xenografts do have tumor–host interactions, but can lose matrix as well after several passages. It is not clear whether such changes in matrix composition

of xenografts are caused by the tumor cells, or by changes in mouse stroma (14). Despite these biological differences, model systems are useful for studying signal transduction pathways important in tumor biology, of which mRNA expression, as measured by qPCR or using gene expression microarrays, is frequently used as a readout. It is therefore highly important to determine whether gene expression levels of these model systems are comparable to those of the corresponding primary tumors, which we aimed to do in this study. We performed gene expression analysis on a panel of 76 conventional high-grade osteosarcoma pretreatment biopsies. We set out to recapitulate representative expression profiles from primary untreated osteosarcoma biopsies in corresponding models *i.e.* cell lines and xenografts. We could demonstrate that both model systems still express genes that are characteristic for the specific histological subtype of the primary tumor. We therefore endorse that, despite biological differences, both xenografts and cell lines are representative model systems for studying mRNA expression in high-grade osteosarcoma. Specific models may be identified that would be appropriate to use for studies of specific subgroups of osteosarcoma.

Methods

Ethics statement

All biological material was handled in a coded fashion. Ethical guidelines of the individual European partners were followed and samples and clinical data were stored in the EuroBoNet biobank. For xenograft experiments, informed consent and sample collection were approved by the Ethical Committee of Southern Norway (Project S-06132) and the Institutional Ethical Committee of Valencia University.

Patients cohorts

Genome-wide expression profiling was performed on pretreatment diagnostic biopsies of 76 resectable highgrade osteosarcoma patients from the EuroBoNet consortium (`www.eurobonet.eu`). Clinicopathological details of these 76 samples can be found in Table 3.1. Samples with a main histological subtype ($n = 66$) were selected for subsequent subtype analyses. These 66 samples included 50 osteoblastic, 9 chondroblastic, and 7 fibroblastic osteosarcomas. Five additional osteosarcoma biopsies (1 chondroblastic and 4 osteoblastic osteosarcomas), 12 mesenchymal stem cell (MSC) and 3 osteoblast cultures, and 12 chondrosarcoma biopsies were used for validation.

Category	Patient characteristics	Biopsies (%)	Cell lines (%)	Xenografts (%)
Total nr of samples		76 (100)	13 (100)	18 (100)
Institution	LUMC, Netherlands	29 (38.2)	0 (0)	0 (0)
	IOR, Italy	11 (14.5)	7 (53.8)	0 (0)
	LOH, Sweden	3 (3.9)	0 (0)	0 (0)
	Radiumhospitalet, Norway	1 (1.3)	3 (23.1)	12 (66.7)
	UV, Spain	0 (0)	0 (0)	6 (33.3)
	WWUM, Germany	32 (42.1)	0 (0)	0 (0)
	Other	0 (0)	3 (23.1)	0 (0)
Origin	Biopsy	76 (100)	0 (0)	0 (0)
	Resection	0 (0)	7 (53.8)	11 (61.1)
	Metastasis	0 (0)	3 (23.1)	1 (5.6)
	Unknown	0 (0)	3 (23.1)	6 (33.3)
Location of primary tumor	Femur	36 (47.4)	0 (0)	10 (55.6)
	Tibia/fibula	26 (34.2)	0 (0)	2 (11.1)
	Humerus	10 (13.2)	0 (0)	2 (11.1)
	Axial skeleton	1 (1.3)	0 (0)	1 (5.6)
	Unknown/other	3 (3.9)	13 (100)	3 (16.7)
Histological subtype	Osteoblastic	50 (65.8)	9 (69.2)	15 (83.3)
	Chondroblastic	9 (11.8)	0 (0)	3 (16.7)
	Fibroblastic	7 (9.2)	4 (30.8)	0 (0)
	Minor	10 (13.2)	0 (0)	0 (0)
Histological response to preoperative chemotherapy in the primary tumor	Good response	33 (43.4)	0 (0)	0 (0)
	Poor response	36 (47.4)	0 (0)	0 (0)
	Unknown/NA	7 (9.2)	13 (100)	18 (100)
Sex	Male	52 (68.4)	9 (69.2)	9 (50)
	Female	24 (31.6)	4 (30.8)	3 (16.7)
	Unknown	0 (0)	0 (0)	6 (33.3)

Table 3.1: Clinicopathological details of patients with conventional high-grade osteosarcoma, including all patients from the biopsy, cell line, and xenograft datasets.

Osteosarcoma cell lines

Out of the EuroBoNeT panel of 19 cell lines, 13 cell lines were recorded to belong to a main histological subtype. This set of 13 cell lines contained 4 cell lines derived from fibroblastic, and 9 cell lines derived from osteoblastic osteosarcomas. The 13 osteosarcoma cell lines IOR/MOS, IOR/OS10, IOR/OS14, IOR/OS15, IOR/OS18, IOR/OS9, IOR/SARG, KPD, MG-63, MHM, OHS, OSA, and ZK-58 were maintained in RPMI 1640 (Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal calf serum and 1% Penicillin/Streptomycin (Invitrogen) as previously described (12). Clinical details of the tissue from which these cell lines were derived are shown in Table 3.1 and are described previously (12).

Osteosarcoma xenografts

The osteosarcoma xenograft model is described in Kresse *et al.* (15). In short, human tumors were implanted directly from patient samples and successively passaged subcutaneously in nude mice. Eighteen different xenografts were used, of which 3 were derived from chondroblastic, and 15 from osteoblastic osteosarcomas. Clinical data on primary tumor samples and xenograft passages that were used are shown in Table 3.1.

Determination of histological subtypes

Histological subtyping was performed by two pathologists (PCWH, EH) on hematoxylin and eosin (HE) stained slides of all biopsies and of all primary tumors from which the osteosarcoma cell lines and xenografts were derived. Osteoblastic, chondroblastic, and fibroblastic osteosarcoma samples were selected for further study. Other subtypes (anaplastic, chondromyxoid fibroma-like, fibroblastic MFH-like, giant cell rich, pleomorphic, and sclerosing osteosarcoma) were excluded because these subtypes are rare.

RNA isolation, cDNA synthesis, cRNA amplification, and Illumina Human-6 v2.0 Expression BeadChip hybridization

Osteosarcoma and xenograft tissue was handled as previously described (18). Osteosarcoma cell lines were prepared as in Ottaviano *et al.* (12). RNA isolation, synthesis of cDNA, cRNA amplification, and hybridization of cRNA onto the Illumina Human-6 v2.0 Expression BeadChips were performed as previously described (18).

Microarray data preprocessing

Microarray data processing and quality control were performed using the statistical language R (19) as described previously (18). MIAME-compliant data have been deposited in the GEO database (`www.ncbi.nlm.nih.gov/geo/`, accession number GSE30699). High

correlations between these microarray data and corresponding qPCR results have been demonstrated previously (18).

Detection of significantly differentially expressed genes

We performed *LIMMA* analyses (20) in order to determine differential expression for the following clinical parameters: sex (52 male *vs* 24 female), tumor location (36 femur, 10 humerus, 26 fibula/tibia), response to preoperative chemotherapy (36 poor responders, or Huvos grade 1–2, *vs* 33 good responders, or Huvos grade 3–4), and histological subtype (an analysis comparing 50 osteoblastic, 9 chondroblastic, and 7 fibroblastic osteosarcomas). Genes that play a role in metastasis-free survival are described in Buddingh *et al.* (18). Probes with Benjamini and Hochberg False discovery rate-adjusted p-values (adjP) < 0.05 were considered to be significantly differentially expressed.

Prediction analysis

The gene expression profile was generated on the dataset of biopsies using Bioconductor (21) package *pamr* (22). Internal cross-validation was performed 50 times. A threshold was selected where the error rate of the prediction profile was minimal. The minimum error rate was representative of 50 independent simulations. In order to minimize optimization bias (23), we validated the profile on an independent dataset of osteosarcoma biopsies ($n = 5$), containing 1 chondroblastic osteosarcoma and 4 osteoblastic osteosarcomas. In addition, we applied the profile on datasets containing positive controls—mesenchymal stem cells (MSC, $n = 12$), osteoblasts ($n = 3$), and chondrosarcoma biopsies ($n = 12$, previously published in (24), GEO accession number GSE12532). We subsequently applied the validated prediction profile to two independent datasets, the first consisting of gene expression data of osteosarcoma cell lines, the second of xenografts. Expression of the probes that composed the prediction profile was verified using a *LIMMA* analysis, comparing chondroblastic, fibroblastic, and osteoblastic osteosarcoma biopsy samples.

Gene set enrichment

Network analysis was performed using Ingenuity Pathways Analysis (IPA, Ingenuity Systems, `www.ingenuity.com`). For both chondroblastic-specific and fibroblastic-specific analyses, data for all reference sequences containing expression values and FDR-adjusted p-values were uploaded into the application. Each identifier was mapped to its corresponding object in Ingenuity's Knowledge Base. An adjP cut-off of 0.05 was set to select genes whose expression was significantly differentially regulated. The Network Eligible molecules were overlaid onto a global molecular network developed from information contained in Ingenuity's Knowledge Base. Networks of Network Eligible Molecules were then

algorithmically generated based on their connectivity. GO term enrichment was tested using Bioconductor package *topGO* (25). Lists of significantly affected genes were compared with all genes eligible for the analysis. GO terms with Fisher's exact p-values < 0.001, as calculated by the *weight* algorithm from *topGO*, were defined significant.

Results

Histological subtypes of osteosarcoma biopsies have different gene expression profiles

We determined differential expression for different clinical parameters. Of all comparisons of clinical parameters only histological subtypes appeared to give a sufficient number of differentially expressed genes to design a prediction profile. *LIMMA* analyses resulted in one location-specific differentially expressed gene: *HOXD4*, which was overexpressed in tumors at the humerus versus at fibula/tibia and femur. Between tumors from male and female patients, 18 genes were significantly differentially expressed, all belonging to X- and Y-chromosome-specific genes, which are not considered as representative for osteosarcoma, yet this finding validates the analysis. No significantly affected genes were returned with regards to response to preoperative chemotherapy. To determine differential expression between the three main histological subtypes, we excluded all samples with unknown or rare subtypes. This resulted in a dataset of 66 conventional high-grade osteosarcoma biopsies with a main histological subtype. Using a *LIMMA* analysis, we determined 1,338 significantly differentially expressed genes (adjP < 0.05) that were specific for a certain main histological subtype (depicted in a Venn diagram in Figure 3.1). A subtype-specific

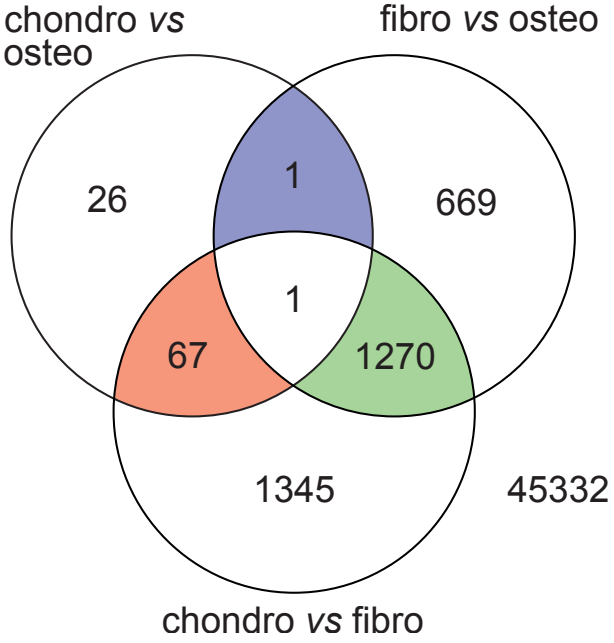


Figure 3.1: Venn diagram representing numbers of fibroblastic- (green), chondroblastic- (red), and osteoblastic (blue)-specific differentially expressed genes obtained with *LIMMA* analysis, considering chondroblastic versus osteoblastic (chondro *vs* osteo), fibroblastic versus osteoblastic (fibro *vs* osteo), and chondroblastic versus fibroblastic (chondro *vs* fibro) analyses. Subtype-specific genes are genes that are either both upregulated or both downregulated in the subtype of interest in the different comparisons.

probe was defined as a probe that had the same sign of log fold change in both analyses,

e.g. the probe was upregulated in chondroblastic samples versus osteoblastic, and in chondroblastic versus fibroblastic samples.

Gene set enrichment shows specific sets of genes are affected in fibroblastic and chondroblastic osteosarcoma

Network analysis using IPA showed that fibroblastic osteosarcoma-specific networks mostly had a role in cellular growth and proliferation, which was also the most significant biological function as detected by IPA (see Additional File 1 (*available online* (26)) for all affected networks and biological functions). The most significant network is illustrated in Additional File 2A (*available online* (26)) and shows that mRNA of various genes with a connection to the NF-κB pathway and STAT5A signaling are upregulated in fibroblastic osteosarcoma biopsies, as compared with both osteoblastic and chondroblastic osteosarcoma. The most significant network specific for the chondroblastic subtype consisted of genes important in skeletal connective tissue development and function (Additional File 2B (*available online* (26))), and shows that, also on the gene expression level, chondroblastic osteosarcoma is mainly distinguished from osteoblastic and fibroblastic osteosarcoma based on the composition of the extracellular matrix of the tumor (Additional File 1 (*available online* (26)) shows all affected networks and biological functions).

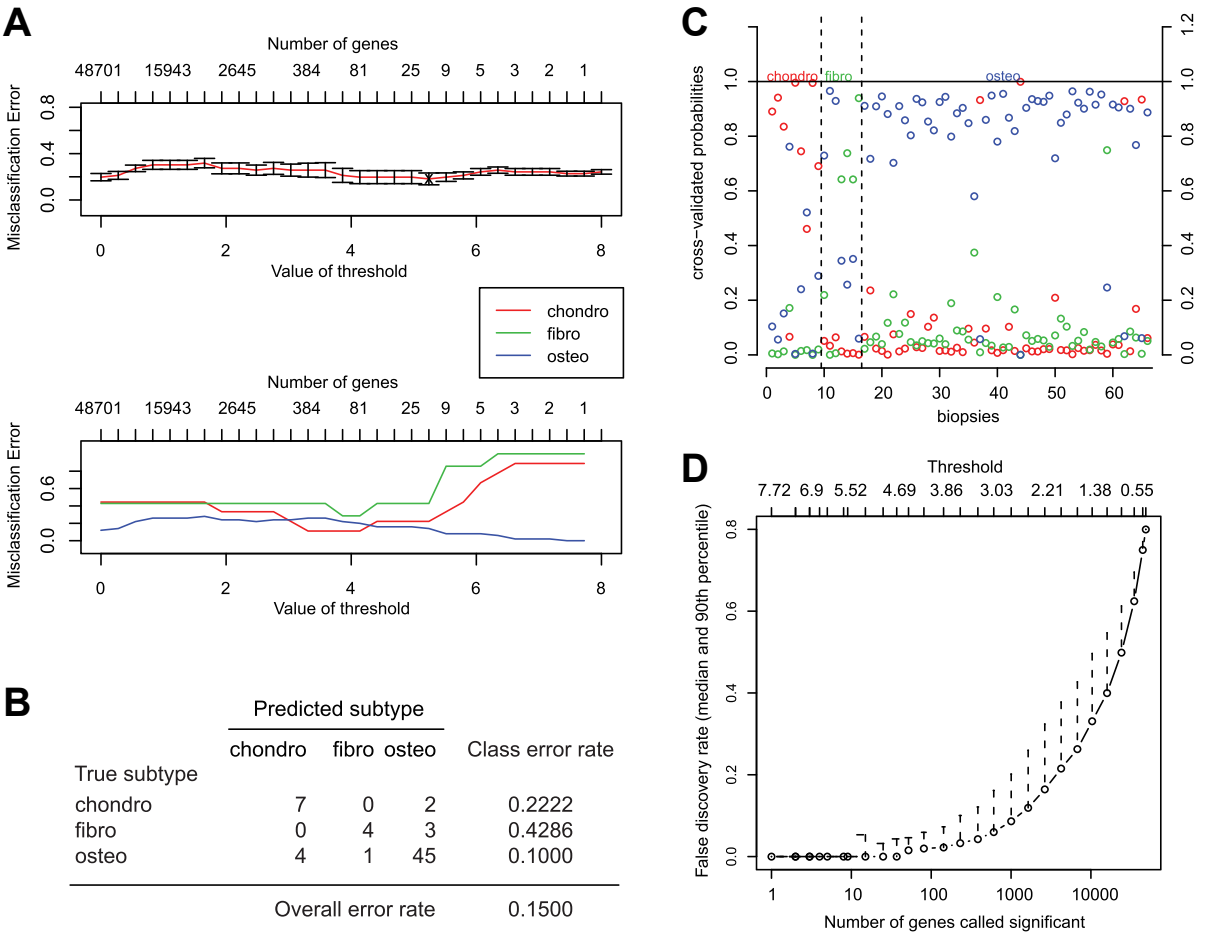
Results from network analysis were confirmed using *topGO*, a gene set enrichment approach analyzing the significance of GO terms in a specific dataset. These analyses resulted in two significant fibroblastic specific GO terms in osteosarcoma: regulation of tyrosine phosphorylation of Stat5 protein (GO:0042522, p-value $= 4.8 \cdot 10^{-4}$) and regulation of survival gene product expression (GO:0045884, p-value $= 8.2 \cdot 10^{-4}$). Significantly differentially expressed genes from both GO terms partly overlap the fibroblastic osteosarcoma-specific network detected with IPA. Two GO terms were significant in the chondroblastic-specific analysis as well: skeletal system development (GO:0001501), and extracellular matrix organization (GO:0030198), which strengthen the results found in the IPA network analyses. GO term subgraphs of the five most significant GO terms for both analyses are shown in Additional File 3 (*available online* (26)).

Gene set enrichment on genes specific for osteoblastic osteosarcoma was not performed, because only one osteoblastic osteosarcoma-specific probe was detected that distinguishes the osteoblastic subtype from fibroblastic and chondroblastic. This probe matches to *UNQ1940*, or *FAM180A*, a protein-coding gene with a yet unknown function.

Generation and validation of the prediction profile

Because we could not directly compare subtype-specific genes between our different model systems due to small sample sizes, we generated a profile that could predict the histological subtype of osteosarcoma. The prediction profile was generated on 66 high-grade

conventional osteosarcoma prechemotherapy biopsies, using nearest shrunken centroids classification. Optimal control of error rate in the prediction profile was found at delta thresholds of $4.9 - 5.1$ (Figure 3.2A), where merely 10 out of 66 samples (15%) in the training set were wrongly assigned to a specific histological subtype. This error rate was



B

True subtype	Predicted subtype: chondro	fibro	osteo	Class error rate
chondro	7	0	2	0.2222
fibro	0	4	3	0.4286
osteo	4	1	45	0.1000
			Overall error rate	0.1500

Figure 3.2: *A*, Illustration of training the *pamr* prediction profile on osteosarcoma biopsies. At thresholds of $4.9 - 5.1$, the misclassification error rate was minimal. *B*, True versus predicted values from the nearest shrunken centroid fit. *C*, Probabilities of each biopsy to belong to any of the three histological subtypes. Samples are separated (dotted lines) based on their true subtypes. Cross-validated probabilities for each histological subtype are shown on the y-axis, so that for every sample three open circles are present (blue, red, and green circles for osteo-, chondro-, and fibroblastic osteosarcoma, respectively). A sample is classified into a specific subtype if the probability to belong to that specific subtype is higher than the probabilities to belong to the other subtypes. *D*, The FDR plotted against different thresholds of the prediction profile. At a threshold of 5.0, 24 genes are included in the prediction profile. These 24 genes have a FDR $< 5\%$.

representative for a set of 50 simulations, which resulted in error rates between 13.5% and 15%. Subtype-specific error rates were 22%, 43%, and 10% for chondroblastic, fibroblastic, and osteoblastic subtypes, respectively (Figure 3.2B). Probabilities of each sample to belong to any of the three histological subtypes are shown in Figure 3.2C. At a threshold delta of 5.0, the prediction profile consisted of 24 probes encoding for 22 genes. All genes

were below a FDR threshold of 5% (Figure 3.2D). Expression of the 24 probes of the profile were verified in a *LIMMA* analysis which was corrected for multiple testing. All 24 probes were confirmed to be significantly differentially expressed in the *LIMMA* analysis as well. Results from *pamr* and *LIMMA* analyses are shown in Table 3.2. A supervised heatmap depicting expression of the 24 probes in all samples is shown in Additional File 4 (*available online* (26)). The prediction profile was validated at threshold delta of 5.0 in an independent dataset of osteosarcoma biopsies and positive controls. Histological subtypes of biopsies had a prediction error of 0% (0/5). Mesenchymal stem cells and osteoblasts all fitted in the osteoblastic group, while 11/12 chondrosarcoma samples were best corresponding to the group of chondroblastic osteosarcoma. The remaining chondrosarcoma sample was a dedifferentiated chondrosarcoma and was predicted in the fibroblastic group, probably because of the high amount of spindle cells present in the biopsy. Additional File 5 (*available online* (26)) shows prediction probabilities for each subtype of these additional datasets.

A prediction profile based on osteosarcoma biopsy data can predict histological subtypes of cell lines and xenografts

Unsupervised clustering of all biopsies, xenografts, and cell lines demonstrated that xenografts and cell lines show different overall expression profiles from most biopsies, and that there are no subtype-specific clusters based on overall expression (Additional File 6 (*available online* (26))). In order to determine whether histological subtypes of cell lines and xenografts could be predicted as well with the 24-gene prediction profile, we applied this profile to two independent datasets. In the first dataset, consisting of osteosarcoma cell line data, 2 out of 13 samples (15%, Figure 3.3A) were wrongly classified. These samples were MG63, a cell line derived from a fibroblastic osteosarcoma, which was subtyped as being osteoblastic, and IOR/OS18, derived from an osteoblastic osteosarcoma, which was subtyped by the prediction profile as a fibroblastic osteosarcoma. Interestingly, HOS, HOS-MNNG, and HOS-143B, all cell lines derived from the HOS cell line, which is derived from fibroblastic and epithelial osteosarcoma and therefore was not added to our set of 13 osteosarcoma cell lines, were all predicted as fibroblastic osteosarcoma (*data not shown*). Two out of 18 xenograft samples (11%, Figure 3.3B) were wrongly classified. One of these samples was OKx, a xenograft derived from a chondroblastic tumor, which was subtyped as an osteoblastic osteosarcoma. The other sample was KPDx, a xenograft derived from an osteoblastic tumor, which was subtyped as fibroblastic. The KPD cell line was subtyped rightly as an osteoblastic osteosarcoma.

probeID	symbol	*LIMMA* logFC C*vs*F	*LIMMA* logFC C*vs*O	*LIMMA* logFC F*vs*O	*LIMMA* adjP	*pamr* chondro-score	*pamr* fibro-score	*pamr* osteo-score
5910377	*ACAN*	2.42	2.24	-0.18	0.0000	0.9294	0	-0.0147
3390678	*NFE2L3*	-1.74	-0.02	1.71	0.0000	0	0.9184	0
1990523	*COL9A1*	3.49	3.01	-0.47	0.0000	0.6011	0	0
360553	*SCRG1*	4.55	3.51	-1.04	0.0001	0.4571	0	0
3310368	itID3	1.87	-0.29	-2.16	0.0003	0	-0.4053	0
10561	*ITIH5L*	0.68	0.65	-0.04	0.0001	0.295	0	0
5050110	*MGC34761*	0.93	0.83	-0.09	0.0002	0.2818	0	0
4780368	*ACAN*	1.34	1.19	-0.14	0.0004	0.2716	0	0
7150719	*COL2A1*	4.82	4.36	-0.47	0.0016	0.183	0	0
3830341	*LYRM1*	-1.23	-0.18	1.06	0.0007	0	0.1677	0
3990500	*MATN4*	1.96	1.69	-0.27	0.0012	0.151	0	0
4280370	*POPDC3*	-0.88	-0.08	0.80	0.0009	0	0.0909	0
6520487	*UNQ830*	4.10	2.90	-1.20	0.0016	0.0817	0	0
2850202	*COL11A2*	1.37	1.10	-0.27	0.0014	0.0735	0	0
4220452	*C11ORF41*	-0.89	-0.03	0.86	0.0011	0	0.0721	0
4560091	*COL9A3*	1.14	1.21	0.07	0.0018	0.0698	0	0
5890452	*LOC652881*	0.43	0.37	-0.06	0.0001	0.0666	0	0
3990259	*PPP2R2B*	-1.00	0.10	1.10	0.0016	0	0.0603	0
5340392	*MAN2A1*	-1.42	-0.22	1.20	0.0018	0	0.0477	0
3360139	*DLX5*	1.84	-0.20	-2.04	0.0033	0	-0.0358	0
2630762	*C14ORF78*	-1.07	1.45	2.52	0.0011	0	0	-0.0307
3460037	*UNQ1940*	0.44	1.71	1.27	0.0018	0	0	-0.0219
6110722	*COL2A1*	1.22	1.44	0.22	0.0032	0.0087	0	0
6980164	*ALPL*	2.52	-0.67	-3.19	0.0038	0	-0.0036	0

Table 3.2: Comparison of the 24 genes obtained with *pamr* prediction with a *LIMMA* analysis between the three different histological subtypes (CvsF: chondroblastic *vs* fibroblastic, CvsO: chondroblastic *vs* osteoblastic, FvsO: fibroblastic *vs* osteoblastic osteosarcoma), for which log fold changes (logFC) are shown for the different coefficients of the analysis. Note that the adjP shows the significance for the whole *LIMMA* analysis, and does not reflect the adjPs per subanalysis.

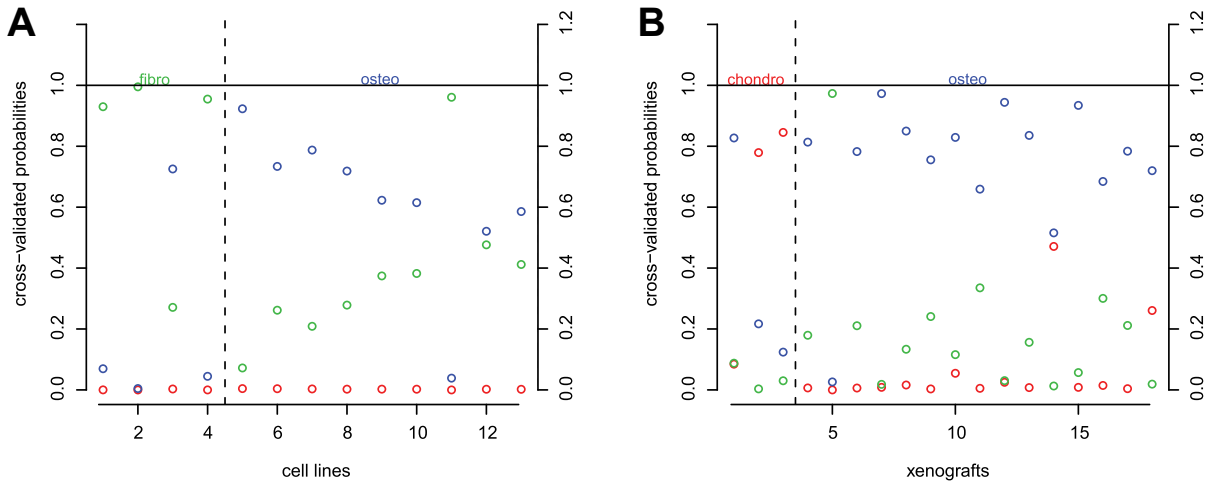


Figure 3.3: Probabilities of *A*, cell lines and *B*, xenografts to belong to any of the three histological subtypes. For an explanation of what is represented by these graphs, see Figure 3.2*C*.

Discussion

In this study, we aimed to compare gene expression profiles of osteosarcoma biopsies with cell lines and xenografts, in order to investigate whether these model systems are representative for the primary tumor. We have determined differential gene expression for different clinical parameters important in high-grade osteosarcoma on a dataset consisting of 76 conventional high-grade osteosarcoma samples. Importantly, pretreatment biopsies were used instead of resected specimens, because preoperative chemotherapy may cause tumor necrosis in responsive patients, thus altering gene expression and hampering the generation of high quality mRNA. We intended to generate a gene expression profile that could not only predict a specific clinical parameter in biopsies, but in osteosarcoma cell lines and xenografts as well. The metastasis/survival profile is described previously and may serve as a tool to predict prognosis and as a target for therapy (18). However, since most of the genes associated with osteosarcoma metastasis were macrophage associated, and no stroma or infiltrate is present in cell lines, this profile could not be applied to osteosarcoma cell lines. We therefore compared gene expression profiles of these different sample sets based on other clinical parameters. No significant differentially expressed genes were found between poor and good responders to chemotherapy. Several reports on genome-wide expression profiling in osteosarcoma have been published describing detection of differential expression between poor and good responders of preoperative chemotherapy (27–30). However, the cohorts used in these studies are all relatively small (n = 13–30), and, more importantly, the reported p-values were not corrected for multiple testing in these studies. Remarkably, only two of the genes that were found to correlate with response to chemotherapy in these studies overlap, and one of these two genes was upregulated in poor responders in one study, whereas it was upregulated in

good responders in the other study (27, 30). Another report described differential expression between a metastatic and a nonmetastatic cell line, for which metastatic capacity correlates with response to chemotherapy (31). In that particular study, four genes out of 252 were found to overlap with a patient study by Mintz *et al.* (27). However, the up- and downregulation of these four genes were not consistent between the two studies. We clearly show in a large cohort that there are no differences between these groups of patients, as the most significant probe had an adjP of 0.9998. These results are in line with our previous findings obtained by analyzing an osteosarcoma cohort on a different platform (32). The parameter 'histological subtypes' resulted in a considerable number of differentially expressed genes. Our prediction profile is not directly applicable to other platforms, but there is no real need to have a prediction profile for primary osteosarcoma histological subtype, since pathologists are very well able to assess this on an HE-section, even on a biopsy, with a concordance of 98% between histological subtype of biopsies and corresponding resections (7). Yet, we here show a quite important use of this profile, *i.e.* to determine the histological subtype of cell lines and xenografts. *In vitro* 2-dimensional growing cells lack extra cellular matrix formation, which is the characteristic feature to distinguish histological subtypes in high-grade central osteosarcoma.

The gene expression profiles as detected by analyzing osteosarcoma biopsy data show a large number of subtype-specific differentially expressed genes. In particular, fibroblastic osteosarcoma differed most from the two other main subtypes. Using gene set enrichment, we detected a network of genes upregulated in fibroblastic osteosarcoma, with a role in cellular growth and proliferation, and connection to the NF-κB pathway. This may be a readout of the high cellularity and low matrix composition of fibroblastic osteosarcoma in comparison with osteoblastic and chondroblastic osteosarcoma (2). GO term enrichment analysis confirmed these results. These pathways may explain why it is comparatively easy to culture fibroblastic osteosarcoma cells, which also may explain why the percentage of fibroblastic osteosarcoma is relatively high in our cell line dataset (31%, compared to 11% in the biopsy dataset). Next to this link to cellular growth and proliferation, the most significant network with fibroblastic-specific upregulated genes showed a connection to the immune system. GO analysis of the five most significant GO terms pointed to involvement of the immune system as well (GO term GO:0006955, p-value $= 3.9 \cdot 10^{-3}$, see Additional File 3 (*available online* (26)) for GO term subgraphs). Forty-four genes in this GO term were significant, of which 43 were upregulated in fibroblastic osteosarcoma. An elevated immune response might be the reason why patients with fibroblastic osteosarcoma tend to have better survival profiles, as a proinflammatory environment has an important role in osteosarcoma metastasis-free survival. This profile is different from the previously found macrophage-specific profile which was associated with better metastasis-free survival of osteosarcoma patients (18). The overrepresentation of pathways involved in chondrogenesis in the chondroblastic profile is in line with the high amount of chondroid matrix in this subtype. We only detected one osteoblastic-specific gene, *UNQ1940*, or *FAM180A*, with

a yet unknown function. Since 50 osteoblastic osteosarcoma samples were compared with only 9 chondroblastic and 7 fibroblastic osteosarcoma samples, we suggest that fibroblastic and chondroblastic osteosarcoma have specific characteristics that distinguishes these tumors from osteoblastic osteosarcoma, and that the latter does not have such an extra feature in comparison with chondro- and fibroblastic osteosarcoma.

Our histological subtype prediction profile consists of 24 probes encoding for 22 genes, all with a specific score which depends on the significance of each gene. The genes that make up the chondroblastic-specific part of this expression profile are mostly chondroid matrixassociated genes, such as *ACAN*, *COL2A1*, and *MATN4*, and are all upregulated in chondroblastic osteosarcoma. Fibroblastic-specific genes that make up the prediction profile are up- or downregulated. An example of a gene upregulated in fibroblastic osteosarcoma is *NFE2L3*, a transcription factor which heterodimerizes with small musculoaponeurotic fibrosarcoma factors and for which a protective role was suggested in hematopoietic malignancies (33). *DLX5*, a transcription factor involved in bone formation, is downregulated in fibroblastic osteosarcoma, and reflects the lower amounts of matrix present in fibroblastic osteosarcoma. No known function is yet available for the two osteoblastic-specific genes. The misclassification error of the prediction profile in the training set of biopsies was 15%. Cell lines and xenografts were predicted with misclassification errors of 15% and of 11%, respectively. It seems most likely that these prediction errors are inherent to the error rate of the prediction profile, which is also 15%. Thus, because these misclassification errors are in the same range, we suggest that gene expression of these model systems is highly similar to gene expression of the tumor from which they are derived. This is especially of interest for studies in cell lines, since no stroma is present on which subtyping can be performed, but repeatedly passaged xenografts often lose stroma as well. Most genes of the prediction profile are matrix-associated genes. Even though these cell lines do not secrete any matrix, and xenografts have diminished amounts of matrix, we can still detect histological subtype markers on an mRNA level, and are able to distinguish different histological subtypes of cell lines and xenografts using this profile.

Conclusions

As osteosarcoma xenografts and cell lines still express histological subtype-specific mRNAs that are characteristic of the primary tumor, these model systems are representative for the primary tumor from which they are derived, and will be useful tools for studying mRNA expression and pathways important in high-grade osteosarcoma.

References

[1] Cleton-Jansen AM, Buerger H, Hogendoorn PCW. Central high-grade osteosarcoma of bone: diagnostic and genetic considerations. *Current Diagnostic Pathology*, 11(6):390–399, 2005.

[2] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[3] Buddingh EP, Anninga JK, Versteegh MIM, Taminiau AHM, *et al.* Prognostic factors in pulmonary metastasized high-grade osteosarcoma. *Pediatric Blood & Cancer*, 54(2):216–221, 2010.

[4] Rozeman LB, Cleton-Jansen AM, Hogendoorn PCW. Pathology of primary malignant bone and cartilage tumours. *International Orthopaedics*, 30(6):437–444, 2006.

[5] Lewis IJ, Nooij MA, Whelan J, Sydes MR, *et al.* Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup. *Journal of the National Cancer Institute*, 99(2):112–128, 2007.

[6] Mohseny AB. Bone: conventional osteosarcoma. `AtlasGeneticsOncology.org/Tumors/ConvOsteoID5344.html`. Accessed: 09/30/2010.

[7] Hauben EI, Weeden S, Pringle J, van Marck EA, *et al.* Does the histological subtype of high-grade central osteosarcoma influence the response to treatment with chemotherapy and does it affect overall survival? A study on 570 patients of two consecutive trials of the European Osteosarcoma Intergroup. *European Journal of Cancer*, 38(9):1218, 2002.

[8] Huvos AG. Bone tumors: diagnosis, treatment and prognosis. WB Saunders Company, 1991.

[9] Bacci G, Ferrari S, Delepine N, Bertoni F, *et al.* Predictive factors of histologic response to primary chemotherapy in osteosarcoma of the extremity: study of 272 patients preoperatively treated with high-dose methotrexate, doxorubicin, and cisplatin. *Journal of Clinical Oncology*, 16(2):658–663, 1998.

[10] Hauben EI, Bielack S, Grimer R, Jundt G, *et al.* Clinico-histologic parameters of osteosarcoma patients with late relapse. *European Journal of Cancer*, 42(4):460–466, 2006.

[11] Hattinger CM, Pasello M, Ferrari S, Picci P, *et al.* Emerging drugs for high-grade osteosarcoma. *Expert Opinion on Emerging Drugs*, 15(4):615–634, 2010.

[12] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[13] Mohseny AB, Machado I, Cai Y, Schaefer KL, *et al.* Functional characterization of osteosarcoma cell lines provides representative models to study the human disease. *Laboratory Investigation*, 91(8):1195–1205, 2011.

[14] Mayordomo E, Machado I, Giner F, Kresse SH, *et al.* A tissue microarray study of osteosarcoma: histopathologic and immunohistochemical validation of xenotransplanted tumors as preclinical models. *Applied Immunohistochemistry & Molecular Morphology*, 18(5):453–461, 2010.

[15] Kresse SH, Meza-Zepeda LA, Machado I, Llombart-Bosch A, *et al.* Preclinical xenograft models of human sarcoma show nonrandom loss of aberrations. *Cancer*, 118(2):558–570, 2012.

[16] Jones KB. Osteosarcomagenesis: modeling cancer initiation in the mouse. *Sarcoma*, 2011:694136, 2011.

[17] Weinberg RA. The Biology of Cancer, volume 255. Garland Science, New York, 2007.

[18] Buddingh EP, Kuijjer ML, Duim RAJ, Bürger H, *et al.* Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. *Clinical Cancer Research*, 17(8):2110–2119, 2011.

[19] Team RDC. R: a language and environment for statistical computing, reference index version 2.10.0. R Foundation for Statistical Computing, 2005.

[20] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):3, 2004.

[21] Gentleman RC, Carey VJ, Bates DM, Bolstad B, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004.

[22] Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proceedings of the National Academy of Sciences*, 99(10):6567–6572, 2002.

[23] Wood IA, Visscher PM, Mengersen KL. Classification based upon gene expression data: bias and precision of error rates. *Bioinformatics*, 23(11):1363–1370, 2007.

[24] Hallor KH, Staaf J, Bovée JVMG, Hogendoorn PCW, *et al.* Genomic profiling of chondrosarcoma: chromosomal patterns in central and peripheral tumors. *Clinical Cancer Research*, 15(8):2685–2694, 2009.

[25] Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–1607, 2006.

[26] Additional files Chapter 3. `www.biomedcentral.com/1755-8794/4/66/additional`.

[27] Mintz MB, Sowers R, Brown KM, Hilmer SC, *et al.* An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Research*, 65(5):1748–1754, 2005.

[28] Ochi K, Daigo Y, Katagiri T, Nagayama S, *et al.* Prediction of response to neoadjuvant chemotherapy for osteosarcoma by gene-expression profiles. *International Journal of Oncology*, 24(3):647–656, 2004.

[29] Man TK, Chintagumpala M, Visvanathan J, Shen J, *et al.* Expression profiles of osteosarcoma that can predict response to chemotherapy. *Cancer Research*, 65(18):8142–8150, 2005.

[30] Salas S, Jézéquel P, Campion L, Deville JL, *et al.* Molecular characterization of the response to chemotherapy in conventional osteosarcomas: predictive value of HSD17B10 and IFITM2. *International Journal of Cancer*, 125(4):851–860, 2009.

[31] Walters DK, Steinmann P, Langsam B, Schmutz S, *et al.* Identification of potential chemoresistance genes in osteosarcoma. *Anticancer Research*, 28(2A):673–679, 2008.

[32] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

[33] Chevillard G, Paquet M, Blank V. Nfe2l3 (Nrf3) deficiency predisposes mice to T-cell lymphoblastic lymphoma. *Blood*, 117(6):2005–2008, 2011.

Chapter 4

Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents

This chapter is based on the publication: Buddingh EP†, Kuijjer ML†, Duim RA, Bürger H, Agelopoulos K, Myklebost O, Serra M, Mertens F, Hogendoorn PCW, Lankester AC, Cleton-Jansen AM. *Clin Cancer Res.* 2011 Apr 15;17(8):2110-9 †Shared first authorship

Abstract

Purpose: High-grade osteosarcoma is a malignant primary bone tumor with a peak incidence in adolescence. Overall survival (OS) of patients with resectable metastatic disease is approximately 20%. The exact mechanisms of development of metastases in osteosarcoma remain unclear. Most studies focus on tumor cells, but it is increasingly evident that stroma plays an important role in tumorigenesis and metastasis. We investigated the development of metastasis by studying tumor cells and their stromal context.

Experimental Design: To identify gene signatures playing a role in metastasis, we carried out genome-wide gene expression profiling on prechemotherapy biopsies of patients who did ($n = 34$) and patients who did not ($n = 19$) develop metastases within 5 years. Immunohistochemistry (IHC) was performed on pretreatment biopsies from 2 additional cohorts ($n = 63$ and $n = 16$) and corresponding postchemotherapy resections and metastases.

Results: A total of 118/132 differentially expressed genes were upregulated in patients without metastases. Remarkably, almost half of these upregulated genes had immunological functions, particularly related to macrophages. Macrophage-associated genes were expressed by infiltrating cells and not by osteosarcoma cells. Tumor-associated macrophages (TAM) were quantified with IHC and associated with significantly better overall survival (OS) in the additional patient cohorts. Osteosarcoma samples contained both M1- (CD14/HLA-DRα positive) and M2-type TAMs (CD14/CD163 positive and association with angiogenesis).

Conclusions: In contrast to most other tumor types, TAMs are associated with reduced metastasis and improved survival in high-grade osteosarcoma. This study provides a biological rationale for the adjuvant treatment of high-grade osteosarcoma patients with macrophage activating agents, such as muramyl tripeptide.

Introduction

High-grade osteosarcoma is a malignant bone tumor characterized by the production of osteoid. The highest incidence is in adolescent patients, with a second peak in patients older than 40 years (1). Despite wide-margin surgery and intensification of chemotherapeutic treatment, overall survival (OS) rates have reached a plateau at about 60% (2–4). Novel administration modalities are needed, but data on critical biological mechanisms allowing the development of novel therapeutic agents are scarce for this relatively rare tumor. In addition to conventional chemotherapeutic agents, recent trials have explored immunostimulatory strategies. The ongoing EURAMOS-1 trial randomizes for treatment with IFN-α in patients with good histological response to neoadjuvant chemotherapy (5). A recently published clinical trial has shown improved OS for osteosarcoma patients treated with the macrophage activating agent muramyl tripeptide (MTP) added to the standard

chemotherapy regimen (6). However, only limited information on macrophage infiltration and activation in osteosarcoma is available (7).

Tumor-associated macrophages (TAM) may promote tumorigenesis through immunosuppression, expression of matrix-degrading proteins and support of angiogenesis. In numerous cancer types, high numbers of M2 or 'alternatively activated' TAMs are associated with a worse prognosis (8–13). M2 macrophages have important functions in wound healing and angiogenesis, express high levels of the immunosuppressive cytokines interleukin (IL)-10 and TGF-β, and express scavenger receptors such as CD163 (14, 15). 'Classical activation' of macrophages by IFN-γ or microbial products results in polarization toward M1-type macrophages. M1 macrophages express high levels of proinflammatory cytokines such as IL-12, IL-1, and IL-6 and have potent antitumor efficacy, both by reactive oxygen species and cytokine-induced cytotoxicity and by induction of natural killer (NK) and T cell activity (16). Rarely, high numbers of TAMs are associated with better prognosis (17, 18). In these cases, TAMs are presumably polarized toward an M1 phenotype, although macrophage subtypes were not reported in these two studies. Alternatively, macrophages may directly phagocytose tumor cells, as has been shown in acute myeloid leukemia (19).

To investigate the role of stroma and stroma–tumor interactions important in metastasis of osteosarcoma, we investigated the development of metastasis by studying tumor cells and their stromal context. By using genome-wide expression analysis, we showed that high expression of macrophage-associated genes in pretreatment biopsies was associated with a lower risk of developing metastases. In addition, we quantified and characterized TAMs in two independent cohorts, including pretreatment biopsies, postchemotherapy resections, and metastatic lesions. In contrast to the tumor-supporting role for TAMs in most epithelial tumor types, higher numbers of infiltrating TAMs correlated with better survival in osteosarcoma. Our findings suggest that macrophages have direct or indirect antiosteosarcoma activity and provide a possible explanation for the beneficial effect of treatment with macrophage activating agents in osteosarcoma.

Materials and methods

Patient cohorts

Genome-wide expression profiling was performed on snap-frozen pretreatment diagnostic biopsies containing viable tumor material of 53 resectable high-grade osteosarcoma patients from the EuroBoNet consortium (`www.eurobonet.eu`; cohort 1). For immunohistochemical validation, a tissue microarray containing 145 formalin-fixed paraffin-embedded (FFPE) samples of 88 consecutive high-grade osteosarcoma patients with primary resectable disease (cohort 2) and 29 FFPE samples of a cohort of 20 consecutive high-grade osteosarcoma patients with resectable disease were used (cohort 3), including material

from pretreatment biopsies, postchemotherapy resections, and metastatic lesions (20). Clinicopathological details can be found in Supplemental Table S1 (*available online* (21)). All biological material was handled in a coded fashion. Ethical guidelines of the individual European partners were followed and samples and clinical data were stored in the EuroBoNet biobank.

Cell lines

The 19 osteosarcoma cell lines 143B, HAL, HOS, IOR/MOS, IOR/OS10, IOR/OS14, IOR/OS15, IOR/OS18, IOR/OS9, IOR/SARG, KPD, MG-63, MHM, MNNG-HOS, OHS, OSA, SAOS-2, U-2 OS, and ZK-58 were maintained in RPMI 1640 (Invitrogen) supplemented with 10% fetal calf serum and 1% penicillin/streptomycin (Invitrogen) as previously described (22).

RNA isolation, cDNA synthesis, cRNA amplification, and Illumina Human-6 v2.0 Expression BeadChip hybridization

Osteosarcoma tissue was snap-frozen in 2-methylbutane (Sigma-Aldrich) and stored at 70°C. By using a cryostat, 20mm sections from each block were cut and stained with hematoxylin and eosin to ensure at least 70% tumor content and viability. RNA was isolated with TRIzol (Invitrogen), followed by RNA cleanup using the QIAGEN Rneasy mini kit with on-column DNase treatment. RNA quality and concentration were measured using an Agilent 2100 Bioanalyzer and Nanodrop ND-1000 (Thermo Fisher Scientific), respectively. Synthesis of cDNA, cRNA amplification, and hybridization of cRNA onto the Illumina Human-6 v2.0 Expression BeadChips was carried out as per manufacturer's instructions.

Reverse transcriptase quantative PCR

Reverse transcriptase quantative PCR (qPCR) analysis of selected target genes was performed as previously described (23). Each experiment was conducted in duplicate by using an automated liquid-handling system (Tecan, Genesis RSP 100). Data were normalized by geometric mean expression levels of 3 reference genes, *i.e.* *SRPR*, *CAPNS1*, and *TBP* using geNorm (`medgen.ugent.be/~jvdesomp/genorm/`). Primer sequences can be found in Supplemental Table S2 (*available online* (21)).

Enzymatic and fluorescent immunostainings

Enzymatic and fluorescent immunostainings were performed on 4mm sections of FFPE tissue as previously described (20). Details regarding antibodies and procedures can be found

in Supplemental Table S3 (*available online* (21)). In case of double immunohistochemistry (IHC), incubation with anti-CD45 and development with DAB+ (Dako) occurred first, followed by a second antigen retrieval before incubation with either anti-CD163 or anti-HLA-DRα and development using the alkaline phosphatase substrate Vector Blue (Vector Labs). In case of double immunofluorescent (IF) stainings, primary antibodies were coincubated overnight. As a positive control, normal and formic acid decalcified tonsil was used, and as a negative control, no primary antibody was added. Tissue microarray slides were scanned using the MIRAX SCAN slide scanner and software (Zeiss, Mirax 3D Histech). Numbers of positively stained cells and vessels were counted using ImageJ (National Institutes of Health, Bethesda, MD) and averaged per 0.6mm core. IF and double IHC images were acquired using a Leica DM4000B microscope fitted with a CRI Nuance spectral analyzer (Cambridge Research and Instrumentation, Inc.) and analyzed using the supplied colocalization tool to determine the percentage of single and double positive pixels per region of interest.

Microarray data analysis

Gene expression data were exported from BeadStudio version 3.1.3.0 (Illumina) in GeneSpring probe profile format and processed and analyzed using the statistical language R (24). As Illumina identifiers are not stable and consistent between different chip versions, raw oligonucleotide sequences were converted to nuIDs (25). Data were transformed using the variance stabilizing transformation algorithm to take advantage of the large number of technical replicates available on the Illumina BeadChips (26). Transformed data were normalized using robust spline normalization, an algorithm combining features of quantile and loess normalization, specifically designed to normalize variance-stabilized data. All microarray data processing was carried out by Bioconductor package *lumi* (27, 28). Quality control was performed using Bioconductor package *arrayQualityMetrics* (29). MIAME (minimum information about a microarray experiment) compliant data have been deposited in the GEO database (`www.ncbi.nlm.nih.gov/geo/`, accession number GSE21257).

Statistical analysis

Differential expression between patients who did ($n = 34$) and did not ($n = 19$) develop metastases within 5 years from diagnosis of the primary tumor was determined using linear models for microarray data (*LIMMA* (30)), applying a Benjamini and Hochberg false discovery rate-adjusted p-value cutoff of 0.05. Other univariate statistical analyses were performed using GraphPad Prism Software (version 5.01). Multivariate survival analyses were carried out according to the Cox proportional hazards model in SPSS (version 16.0.2). Two-sided p-values < 0.05 were determined to be significant; p-values between 0.05 and 0.15 were defined to be a trend.

Results

High expression of macrophage-associated genes in osteosarcoma biopsies of patients who did not develop metastases within 5 years from diagnosis (cohort 1)

Comparison of genome-wide gene expression in tumors of patients who did and did not develop metastases within 5 years resulted in 139 significantly differentially expressed (DE) probes, of which 125 corresponded to 118 upregulated and 14 to downregulated genes in patients who did not develop metastases. A summary of DE genes and detailed descriptions of all probes can be found in Table 4.1 and Supplemental Table S4 (*available online* (21)), respectively. Two DE genes were specific for macrophages (*CD14* and *MSR1*) and 30/132 of the DE genes were associated with macrophage functions such as antigen processing and presentation (*e.g. HLA-DRA* and *CD74*) or pattern recognition (*e.g. TLR4* and *NLRP3*). Overall, approximately 20% of the upregulated probes corresponded to genes that were associated with macrophage function and development and an additional 25% of the upregulated probes corresponded to genes with other immunological functions, such as cytokine production and phagocytosis. Four genes were selected for validation of the microarray data using qPCR: *CD14*, *HLA-DRA*, *CLEC5A*, and *FCGR2A*. Expression levels as determined by qPCR correlated well with expression levels obtained by microarray analysis (Supplemental Figure S1 (*available online* (21))). Metastases-free survival curves of the same cohort, generated using median expression of the probe of interest as a cutoff determining low and high expression, are shown in Figure 4.1B and Supplemental Figure 2 (*available online* (21)). Cox proportional hazards analysis revealed expression of macrophage-associated genes *CD14* and *HLA-DRA* to be independently associated with metastasis-free survival (Supplemental Table S5 (*available online* (21))).

	Higher expression in patients without metastases			Lower expression in these patients		
	Number of probes	Number of genes	Examples	Number of probes	Number of genes	Examples
Pattern recognition receptor or signaling	18	17	*MSR1*, *CD14*, *NLRP3*, *TLR7*, *TLR8*, *TLR4*, *NAIP*, *IL1B*, *PYCARD*, *NLRC4*	0	0	
Immunological	16	15	*CD86*, *C1QA*, *LY9*, *CD37*, *LY86*	0	0	
HLA class II	12	12	*HLA-DMB*, *HLA-DRA*, *CD74*, *HLA-DQA1*	0	0	
Hematopoietic cells	11	10	*HMHA1*, *MYO1G*, *LST1*	0	0	
Cytokines and cytokine signaling	7	6	*CXCL16*, *CSF2RA*, *IFNGR1*, *IL10RA*	1	1	*MAP2K7*
Metabolism	9	9	*PFKFB2*, *SLC2A9*, *CECR1*, *ALOX5*	0	0	
Fc receptor	6	4	*FCGR2B*, *FCGR2A*, *FGL2*, *PTPN6*	0	0	
Cytoskeleton	5	5	*HCLS1*, *WAS*, *IQGAP2*	1	1	*DNAI2*
(An)ion transporters and channels	4	4	*SLCO2B1*, *SLC11A1*	1	1	*SLC24A4*
AKT pathway	3	3	*PIK3IP1*, *PKIB*	0	0	
Endocytosis	3	3	*APPL2*, *NECAP2*	0	0	
Apoptosis, cell cycle control, and proliferation	4	4	*TMBIM4*, *TNFRSF1B*, *OGFRL1*	1	1	*BCCIP*
Signaling	4	4	*RGS10*, *MFNG*, *FHL2*, *PILRA*	0	0	
Growth hormone signaling	0	0		1	1	*GHR*
Morphogenesis	0	0		1	1	*HOXC4*
Others	7	6	*CUGBP2*, *CYP2S1*, *VAV1*, *GGN*	2	2	*NSUN5*, *MRPL4*
Unknown	16	16	*VMO1*, *MICALCL*, *MS4A6A*	6	6	*NHN1*, *BRWD1*
Total	125	118		14	14	

Table 4.1: DE genes and probes by category comparing high-grade osteosarcoma patients with and without metastases within 5 years by genome-wide expression profiling (cohort 1).

Macrophage-associated genes are expressed by infiltrating hematopoietic cells and not by tumor cells

The most probable source of expression of the DE macrophage-associated genes was infiltrating immune cells and not osteosarcoma cells. To confirm this, we performed qPCR of *CD14* and *HLA-DRA* on osteosarcoma cell lines ($n = 19$) and biopsies ($n = 45$, a subset of cohort 1). *CD14* and *HLA-DRA* expression was variable in osteosarcoma biopsies, but almost undetectable in cell lines. This indicates that these macrophage-associated genes were not expressed by tumor cells but by infiltrating cells because only osteosarcoma biopsies contain macrophage infiltrate, whereas RNA from cell lines is exclusively from tumor cells (Figure 4.1A, Mann-Whitney U test p-value < 0.0001). In addition, we performed double IHC for the hematopoietic cell marker CD45, which is not expressed by osteosarcoma tumor cells, and the macrophage marker CD163 or the macrophage-associated protein HLA-DRα (Figure 4.1C). We chose this approach because no reliable osteosarcoma markers are available (1). Our results confirmed that infiltrating hematopoietic cells were the source of the macrophage-associated gene expression levels. Together, these data show that osteosarcoma tumor cells do not express macrophage-associated genes, neither *in vitro* nor *in vivo*.

Macrophage numbers in osteosarcoma biopsies correlate with *CD14* gene expression levels and are positively associated with localized disease and better outcome (cohorts 2 and 3)

To confirm the presence of TAMs in osteosarcoma, we stained a tissue microarray containing 145 samples of 88 patients for the macrophage marker CD14 and counted the number of positive cells per tissue microarray core (cohort 2; Figure 4.2A). CD14 was chosen as opposed to CD68 because the latter marker is not expressed by monocytes and often shows cross-reactivity with mesenchymal tissue (*data not shown*). Number of CD14-positive cells per tissue microarray core correlated significantly with *CD14* mRNA expression levels (14 samples overlap with gene expression analysis, Spearman correlation coefficient 0.64, p-value = 0.01). Similar to the gene expression data, there was a trend for patients with primary localized disease to have higher numbers of macrophages in pretreatment diagnostic biopsies than patients with metastatic disease at presentation (mean number of macrophages per core, 55 *vs* 27; Mann-Whitney U test p-value = 0.09). Also, patients with high macrophage counts at diagnosis tended to be less likely to develop metastases within 5 years (χ^2, p-value = 0.13).

We subdivided this cohort into four quartiles based on numbers of CD14-positive cells to determine the group with the best OS. No significant differences were found between quartiles 2 and 4, but patients belonging to this group had better OS than patients with low CD14 counts (lowest quartile, or less than 12 CD14-positive cells per tissue array core;

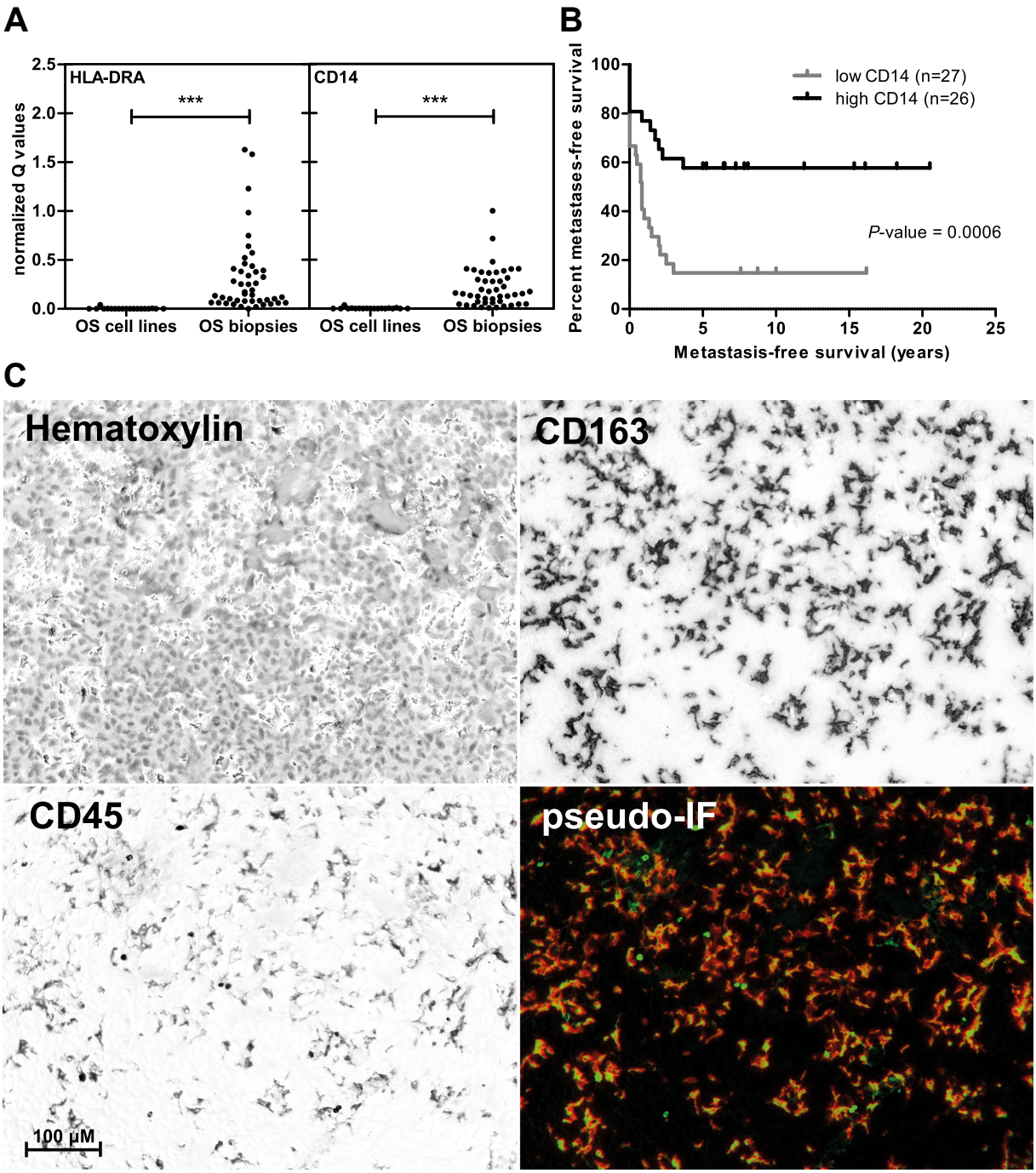


Figure 4.1: Macrophage-associated genes are not expressed by osteosarcoma tumor cells. *A*, qPCR of osteosarcoma cell lines and biopsies of *CD14* and *HLA-DRA* demonstrating lack of expression by osteosarcoma cells. Mann-Whitney U test p-value < 0.0001, $***$. *B*, High expression of macrophage associated genes was associated with a better metastasis-free survival (cohort 1, Kaplan-Meier curve, p-value obtained by Logrank test, patients with metastasis at diagnosis have an event at $t = 0$. These patients are included, because patients who develop metastases later on may as well have micrometastases at the time of diagnosis). Metastasis-free survival curves for *HLA-DRA*, *CLEC5A*, and *FCGR2A* can be found in Supplemental Figure S2 (*available online* (21)). *C*, Double immunohistochemical staining of CD163 with the hematopoietic cell marker CD45 was performed and analyzed using spectral imaging microscopy. The pseudo-IF image (pseudo-IF) shows CD163-positive cells in red, CD45-positive cells in green, and colocalization of both markers in orange. Lack of expression of CD163 and CD45 on surrounding tumor cells (dark blue) and some single positive CD45 cells can be noted.

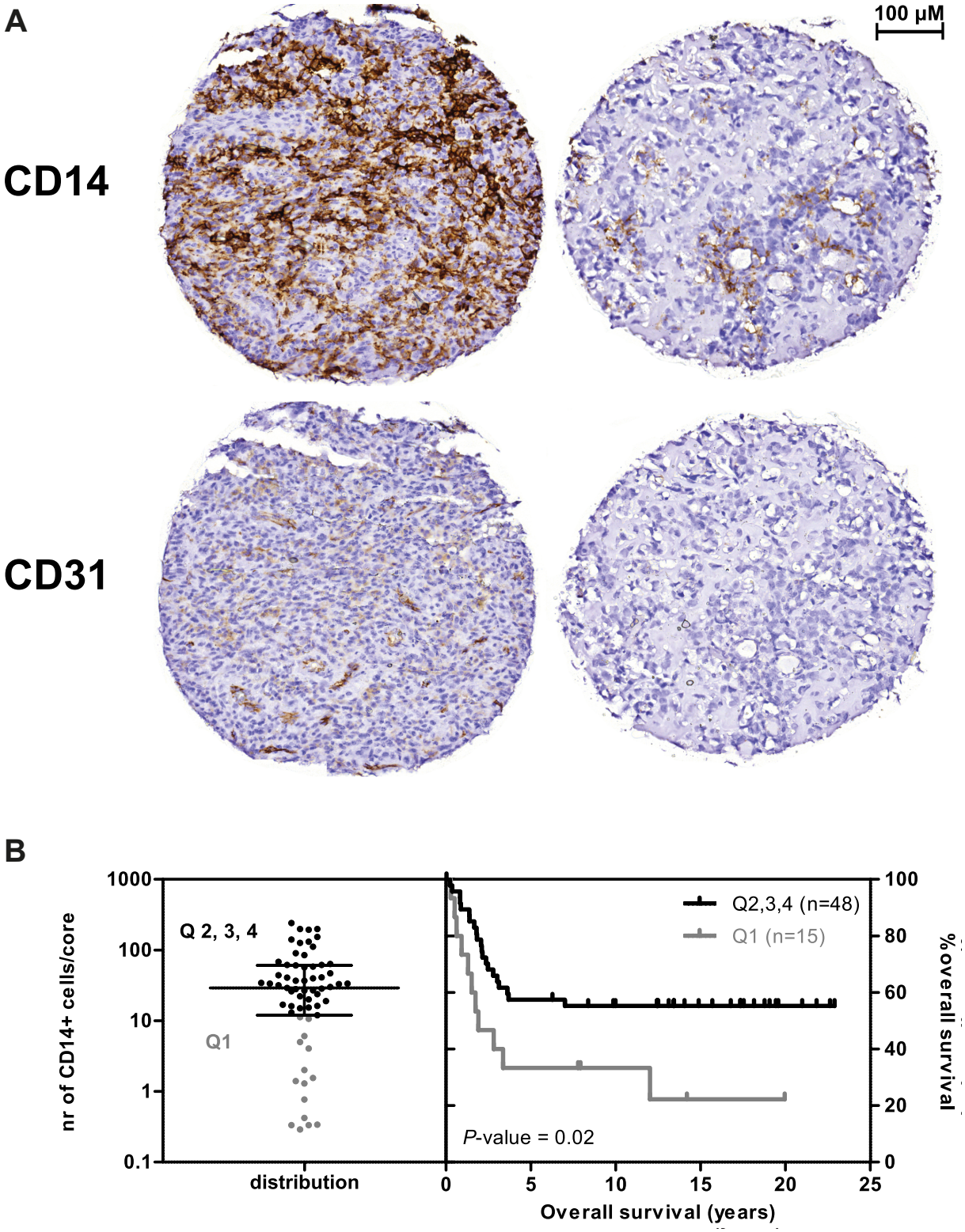


Figure 4.2: *A*, Example of representative stainings of high-grade osteosarcoma with high (left) versus low (right) levels of macrophage infiltration (CD14 staining) and vascular density (CD31 staining). *B*, High numbers of infiltrating macrophages (left, defined as the 3 upper quartiles, or more than 12 CD14-positive cells per tissue array core) are associated with better OS (right, Logrank test p-value = 0.02, cohort 2). Q1: lowest quartile, Q2, 3, 4, 3: highest quartiles.

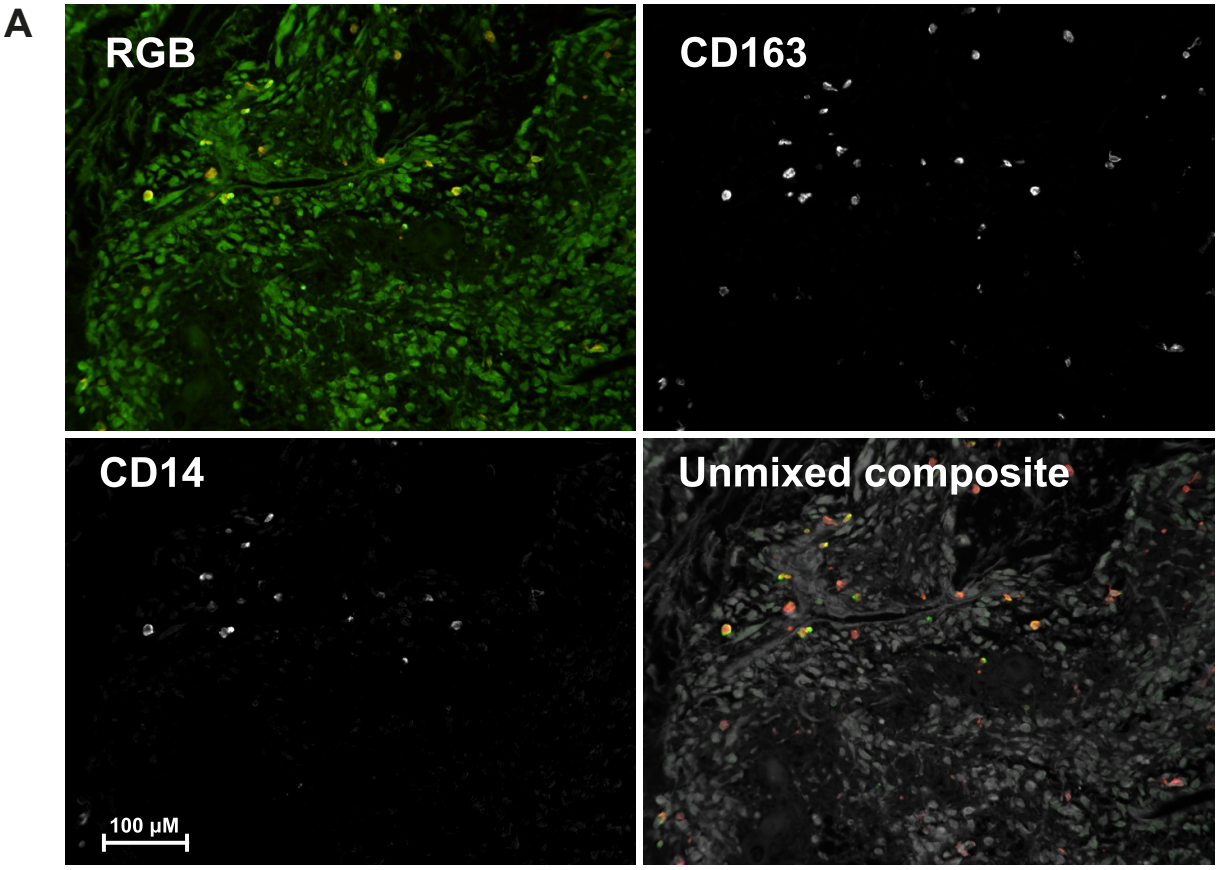
Figure 4.2B, Logrank test p-value = 0.02). In another cohort of 16 patients, IF staining of CD14, CD163, and HLA-DRα was performed, again confirming a potential prognostic value of high macrophage numbers (cohort 3, Figure 4.3, Logrank test p-value = 0.01, Supplemental Figure S3 (*available online* (21))).

Macrophages in osteosarcoma have both M1 and M2 characteristics

To determine the phenotype of macrophages present in osteosarcoma, we performed double IHC with CD14 and either the M1-associated marker HLA-DRα or the M2-associated marker CD163. Not all CD163 and HLA-DR-positive infiltrating cells expressed CD14 (Figure 4.3A and Supplementary Figure S3A). The total number of macrophages as determined by quantifying CD14-positive macrophages was associated with good survival (Figure 4.3B), but the phenotype of the macrophages (CD14/CD163 double positive versus CD14/HLA-DRα double positive) was not (Supplemental Figure S3B (*available online* (21)); *data not shown*). Another M2 characteristic is support of angiogenesis. The number of CD14-positive macrophages correlated with the number of CD31-positive vessels (Figure 4.2A and Figure 4.4), but vascularity did not correlate with prognosis (*data not shown*).

Macrophage numbers in diagnostic biopsies may predict histological response to chemotherapy and macrophage number increases following chemotherapy treatment

There was a trend for high macrophage count (highest three quartiles or more than 12 CD14-positive cells per tissue array core) in prechemotherapy diagnostic biopsies of the primary tumor to predict for good histological response to neoadjuvant chemotherapy (defined as more than 90% nonvital tumor tissue upon final resection), since 46% of patients with high macrophage numbers and 18% of patients with low macrophage numbers had a good histological response (cohort 2; $\chi^2 = 0.09$). The prognostic benefit of macrophage counts in osteosarcoma was not independent of histological response using Cox proportional hazard analysis. Macrophage numbers were higher in postchemotherapy resections of the primary tumor than in prechemotherapy biopsies (Supplemental Figure S4 (*available online* (21))). Moreover, gene expression analysis showed upregulation of macrophage-associated probes in postchemotherapy resections ($n = 4$) as compared with prechemotherapy biopsies ($n = 79$, *data not shown*).



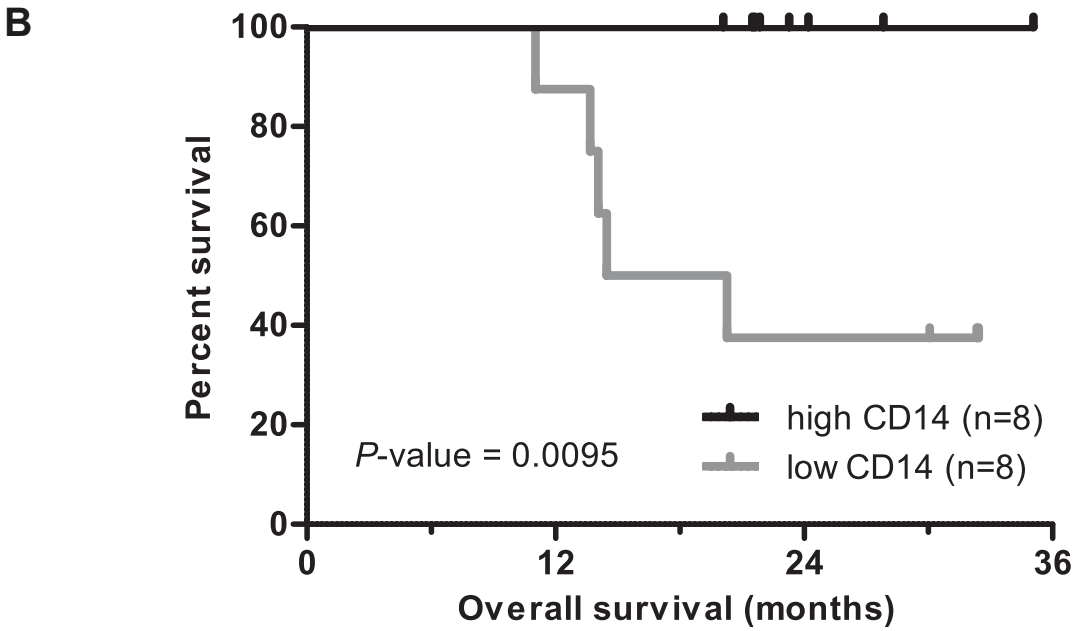


Figure 4.3: *A*, Osteosarcoma samples are infiltrated with CD14 and CD163 single and double positive macrophages. Spectral imaging was used to reduce autofluorescence of osteosarcoma cells. In the composite image, CD14-positive cells are represented in green, CD163-positive cells are represented in red, and CD14/CD163 double positive cells are represented in yellow. Background autofluorescence of tumor cells is represented in gray. *B*, In an independent cohort of 16 patients (cohort 3), high macrophage infiltration as determined by IF CD14 staining was associated with significantly improved OS. p-values obtained using Logrank test, cutoff at the median.

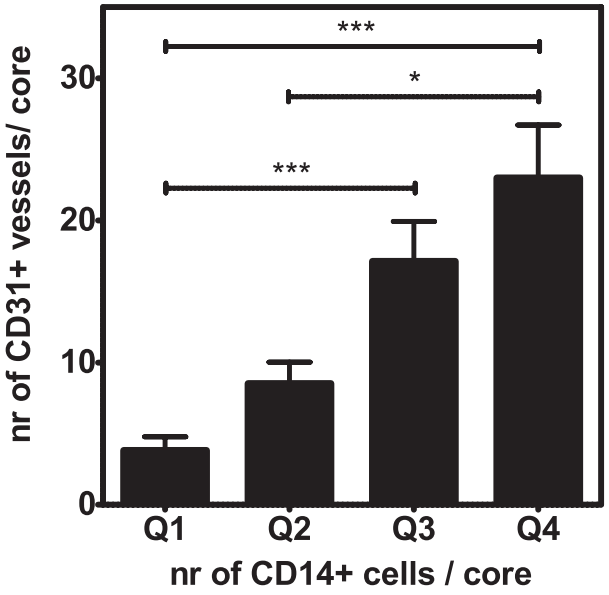


Figure 4.4: Macrophage infiltration as determined by CD14-positive cell count correlated with vascularity as determined by CD31-positive vessel count. Data of all osteosarcoma samples (pre- and posttreatment primary tumor and metastatic samples, cohort 2) are shown. Q1: lowest quartile, Q2, 3, 4: three highest quartiles. Kruskal-Wallis test p-value < 0.0001. $*$, Dunn's posttest p-value < 0.05, $***$, Dunn's posttest p-value < 0.001.

Discussion

OS of high-grade osteosarcoma patients with resectable metastatic disease is poor at about 20% (31). Mechanisms for the development of metastases in osteosarcoma are elusive. To identify genes that play a role in this process, we performed genome-wide expression profiling on prechemotherapy biopsies of osteosarcoma patients. We compared patients who developed clinically detectable metastases within 5 years with patients who did not develop metastases within this time frame (cohort 1). About 20% of genes overexpressed in patients without metastases were macrophage-associated, whereas an additional 25% of genes had other immunological functions (*e.g.* in phagocytosis, complement activation or cytokine production and response) but could still be attributed to macrophages (Table 1 and Supplemental Table S4 (*available online* (21))). Thus, in total, almost half of the DE genes belonged to one specific process, *i.e.* macrophage function. Macrophage-associated genes were expressed by infiltrating hematopoietic cells and not by osteosarcoma tumor cells (Figure 4.1), indicating a possible role for macrophages in preventing metastasis in osteosarcoma. To confirm these findings, we quantified infiltrating macrophages in two additional cohorts (cohorts 2 and 3) and found an association with better OS in both cohorts.

The antimetastatic effect of TAMs in osteosarcoma is remarkable because TAMs support tumor growth in a substantial number of other cancers, which are mostly tumors of epithelial origin. For example, macrophages are associated with the angiogenic switch in breast cancer (32). We find an association between macrophage infiltration and higher microvessel density, which suggests that the influx of macrophages may support certain aspects of tumor growth in osteosarcoma as well. However, in the case of osteosarcoma, direct or indirect antitumor activity of macrophages apparently outweighs their possible tumor-supporting effects. Macrophages can alter their phenotype from M2 to M1 and become the tumor's foe instead of its friend, given the right circumstances (33–35). The

TAMs that were identified in this study in osteosarcoma had both M1 and M2 characteristics. The expression of CD163 and the association with angiogenesis are M2 characteristics (32, 36). Some of the DE genes, such as *MSR1* and *MS4A6A* are specific for M2 macrophages *in vitro* (37). Others, such as the proinflammatory cytokine *IL1B*, are more indicative of an M1 phenotype (16). How macrophages inhibit osteosarcoma metastasis and whether a balance between M1- and M2-type functions is responsible is unknown.

In a multivariate regression model, the survival benefit of high TAM numbers was at least partly dependent on histological response to chemotherapy. Chemotherapy can cause 'immunogenic cell death' of cancer cells, resulting in the release of endogenous danger signals (38, 39). The binding of these dangerous signals to pattern recognition receptors on macrophages can skew polarization of M2- to M1-type TAMs. The interaction between dying tumor cells and resident TAMs may facilitate clearance or inhibit outgrowth of metastatic tumor cells. However, patients with localized disease at diagnosis tended to have a larger macrophage infiltrate than patients with metastatic disease at diagnosis (mean number of macrophages per core 55 *vs* 27). At this point, patients have not undergone chemotherapeutic treatment yet and an interaction between chemotherapy and macrophages can therefore not be responsible for the antimetastatic effect of macrophages. Perhaps, the antimetastatic effect of TAMs in these patients is due to the constitutive presence of macrophages with an M1 phenotype. Alternatively, the presence of macrophages might be a reflection of a microenvironment not conducive for metastasis.

Although preliminary analysis of a clinical trial investigating the effect of treatment with the macrophage activating agent MTP yielded conflicting results, subsequent analysis revealed that treatment with MTP improved 6-year OS from 70% to 78% in a cohort of patients with primary localized disease (6, 40). Similar results were obtained in canine osteosarcoma (41). MTP is a synthetic derivative of muramyl dipeptide (MDP), a common bacterial cell wall component. Muropeptides bind to intracellular pattern recognition receptors of the nucleotide binding and oligomerization domain (NOD)-like receptor (NLR) family, expressed by macrophages (42). In our study, 5 genes associated with NLR family signaling and the associated 'inflammasome' were highly expressed in pretreatment biopsies of patients who do not develop metastases. The DE genes *NLRP3*, *NAIP*, *NLRC4*, and *PYCARD* are components of the inflammasome, *LYZ* is a lysozyme that processes bacterial cell wall peptidoglycan into MDP, a ubiquitous natural analogue of MTP, and *IL1B* is the downstream effector cytokine of the inflammasome pathway. Further research is needed to clarify whether only patients with high numbers of TAMs benefit from MTP treatment, or whether MTP treatment is effective regardless of macrophage number or activation status pretreatment. Also, it is unknown whether treatment with agents promoting macrophage migration or with other macrophage activating agents like toll-like receptor ligands or IFNs has a similar beneficial effect on outcome.

Previous genome-wide expression profiling studies in osteosarcoma focused on identifying genes that predict histological response to neoadjuvant chemotherapy (43–46). As

a consequence, the importance of macrophages in controlling metastases was not recognized. However, we previously compared gene expression profiles of osteosarcoma biopsies and cultured mesenchymal stem cells and determined which genes are expressed by tumor stroma and not by tumor cells (47). There is a considerable overlap between the stromal genes identified in our previous study and the macrophage-associated genes identified in the present study (including HLA class II genes as the most prevalent DE group of genes and the macrophage-associated genes *MSR1*, *MS4A6A*, and *FCGR2A*).

In conclusion, we showed the presence and clinical significance of TAMs in pretreatment samples of high-grade osteosarcoma. TAMs in osteosarcoma are a heterogeneous cell population with both M1 antitumor and M2 protumor characteristics. Although the exact mechanism by which macrophages exert their antimetastatic functions is still unknown, this study provides an important biological rationale for the treatment of osteosarcoma patients with macrophage activating agents.

References

[1] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[2] Lewis IJ, Nooij MA, Whelan J, Sydes MR, *et al.* Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup. *Journal of the National Cancer Institute*, 99(2):112–128, 2007.

[3] Bacci G, Longhi A, Versari M, Mercuri M, *et al.* Prognostic factors for osteosarcoma of the extremity treated with neoadjuvant chemotherapy. *Cancer*, 106(5):1154–1161, 2006.

[4] Bielack SS, Kempf-Bielack B, Delling G, Exner GU, *et al.* Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *Journal of Clinical Oncology*, 20(3):776–790, 2002.

[5] Marina N, Bielack S, Whelan J, Smeland S, *et al.* International collaboration is feasible in trials for rare conditions: the EURAMOS experience. *Cancer Treatment and Research*, 152:339–353, 2009.

[6] Meyers PA, Schwartz CL, Krailo MD, Healey JH, *et al.* Osteosarcoma: the addition of muramyl tripeptide to chemotherapy improves overall survival—a report from the Children's Oncology Group. *Journal of Clinical Oncology*, 26(4):633–638, 2008.

[7] Kleinerman ES, Raymond AK, Bucana CD, Jaffe N, *et al.* Unique histological changes in lung metastases of osteosarcoma patients following therapy with liposomal muramyl tripeptide (CGP 19835A lipid). *Cancer Immunology, Immunotherapy*, 34(4):211–220, 1992.

[8] Hagemann T, Wilson J, Burke F, Kulbe H, *et al.* Ovarian cancer cells polarize macrophages toward a tumor-associated phenotype. *The Journal of Immunology*, 176(8):5023–5032, 2006.

[9] Lee CH, Espinosa I, Vrijaldenhoven S, Subramanian S, *et al.* Prognostic significance of macrophage infiltration in leiomyosarcomas. *Clinical Cancer Research*, 14(5):1423–1430, 2008.

[10] Lissbrant IF, Stattin P, Wikstrom P, Damber J, *et al.* Tumor associated macrophages in human prostate cancer: relation to clinicopathological variables and survival. *International Journal of Oncology*, 17(3):445–452, 2000.

[11] Volodko N, Reiner A, Rudas M, Jakesz R. Tumour-associated macrophages in breast cancer and their prognostic correlations. *The Breast*, 7(2):99–105, 1998.

[12] Jensen TO, Schmidt H, Møller HJ, Høyer M, *et al.* Macrophage markers in serum and tumor have prognostic impact in American Joint Committee on Cancer stage I/II melanoma. *Journal of Clinical Oncology*, 27(20):3330–3337, 2009.

[13] van Dongen M, Savage NDL, Jordanova ES, Briaire-de Bruijn IH, *et al.* Anti-inflammatory M2 type macrophages characterize metastasized and tyrosine kinase inhibitor-treated gastrointestinal stromal tumors. *International Journal of Cancer*, 127(4):899–909, 2009.

[14] Sica A, Larghi P, Mancino A, Rubino L, *et al.* Macrophage polarization in tumour progression. *Seminars in Cancer Biology*, 18(5):349–355, 2008.

[15] Qian BZ, Pollard JW. Macrophage diversity enhances tumor progression and metastasis. *Cell*, 141(1):39–51, 2010.

[16] Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. *Nature Reviews Immunology*, 8(12):958–969, 2008.

[17] Kim DW, Min HS, Lee KH, Kim YJ, *et al.* High tumour islet macrophage infiltration correlates with improved patient survival but not with EGFR mutations, gene copy number or protein expression in resected non-small cell lung cancer. *British Journal of Cancer*, 98(6):1118–1124, 2008.

[18] Forssell J, Öberg Å, Henriksson ML, Stenling R, *et al.* High macrophage infiltration along the tumor front correlates with improved survival in colon cancer. *Clinical Cancer Research*, 13(5):1472–1479, 2007.

[19] Jaiswal S, Chao MP, Majeti R, Weissman IL. Macrophages as mediators of tumor immunosurveillance. *Trends in Immunology*, 31(6):212–219, 2010.

[20] Mohseny AB, Szuhai K, Romeo S, Buddingh EP, *et al.* Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of Cdkn2. *The Journal of Pathology*, 219(3):294–305, 2009.

[21] Supplemental files Chapter 4. `clincancerres.aacrjournals.org/content/17/8/2110/suppl/DC1`.

[22] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[23] Rozeman LB, Hameetman L, Cleton-Jansen AM, Taminiau AHM, *et al.* Absence of IHH and retention of PTHrP signalling in enchondromas and central chondrosarcomas. *The Journal of Pathology*, 205(4):476–482, 2005.

[24] Team RDC. R: a language and environment for statistical computing, reference index version 2.9.0. R Foundation for Statistical Computing, 2005.

[25] Du P, Kibbe WA, Lin SM. nuID: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. *Biology Direct*, 2:16, 2007.

[26] Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Research*, 36(2):e11, 2008.

[27] Gentleman RC, Carey VJ, Bates DM, Bolstad B, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004.

[28] Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics*, 24(13):1547–1548, 2008.

[29] Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics*, 25(3):415–416, 2009.

[30] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):3, 2004.

[31] Buddingh EP, Anninga JK, Versteegh MIM, Taminiau AHM, *et al.* Prognostic factors in pulmonary metastasized high-grade osteosarcoma. *Pediatric Blood & Cancer*, 54(2):216–221, 2010.

[32] Lin EY, Li JF, Gnatovskiy L, Deng Y, *et al.* Macrophages regulate the angiogenic switch in a mouse model of breast cancer. *Cancer Research*, 66(23):11238–11246, 2006.

[33] Hagemann T, Lawrence T, McNeish I, Charles KA, *et al.* 'Re-educating' tumor-associated macrophages by targeting NF-κB. *The Journal of Experimental Medicine*, 205(6):1261, 2008.

[34] Sinha P, Clements VK, Ostrand-Rosenberg S. Reduction of myeloid-derived suppressor cells and induction of M1 macrophages facilitate the rejection of established metastatic disease. *The Journal of Immunology*, 174(2):636–645, 2005.

[35] Buhtoiarov IN, Sondel PM, Eickhoff JC, Rakhmilevich AL. Macrophages are essential for antitumour effects against weakly immunogenic murine tumours induced by class B CpG-oligodeoxynucleotides. *Immunology*, 120(3):412–423, 2006.

[36] Ojalvo LS, King W, Cox D, Pollard JW. High-density gene expression analysis of tumor-associated macrophages from mouse mammary tumors. *The American Journal of Pathology*, 174(3):1048–1064, 2009.

[37] Martinez FO, Gordon S, Locati M, Mantovani A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *The Journal of Immunology*, 177(10):7303–7311, 2006.

[38] Zitvogel L, Apetoh L, Ghiringhelli F, Kroemer G. Immunological aspects of cancer chemotherapy. *Nature Reviews Immunology*, 8(1):59–73, 2008.

[39] Kono H, Rock KL. How dying cells alert the immune system to danger. *Nature Reviews Immunology*, 8(4):279–289, 2008.

[40] Meyers PA, Schwartz CL, Krailo M, Kleinerman ES, *et al.* Osteosarcoma: a randomized, prospective trial of the addition of ifosfamide and/or muramyl tripeptide to cisplatin, doxorubicin, and high-dose methotrexate. *Journal of Clinical Oncology*, 23(9):2004–2011, 2005.

[41] Kurzman ID, MacEwen EG, Rosenthal RC, Fox LE, *et al.* Adjuvant therapy for osteosarcoma in dogs: results of randomized clinical trials using combined liposome-encapsulated muramyl tripeptide and cisplatin. *Clinical Cancer Research*, 1(12):1595–1601, 1995.

[42] Geddes K, Magalhães JG, Girardin SE. Unleashing the therapeutic potential of NOD-like receptors. *Nature Reviews Drug Discovery*, 8(6):465–479, 2009.

[43] Ochi K, Daigo Y, Katagiri T, Nagayama S, *et al.* Prediction of response to neoadjuvant chemotherapy for osteosarcoma by gene-expression profiles. *International Journal of Oncology*, 24(3):647–656, 2004.

[44] Salas S, Jézéquel P, Campion L, Deville JL, *et al.* Molecular characterization of the response to chemotherapy in conventional osteosarcomas: predictive value of HSD17B10 and IFITM2. *International Journal of Cancer*, 125(4):851–860, 2009.

[45] Man TK, Chintagumpala M, Visvanathan J, Shen J, *et al.* Expression profiles of osteosarcoma that can predict response to chemotherapy. *Cancer Research*, 65(18):8142–8150, 2005.

[46] Mintz MB, Sowers R, Brown KM, Hilmer SC, *et al.* An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Research*, 65(5):1748–1754, 2005.

[47] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

Chapter 5

IR/IGF1R signaling as potential target for treatment of high-grade osteosarcoma

This chapter is based on the manuscript: Kuijjer ML, Peterse EFP, van den Akker BEWM, Briaire-de Bruijn IH, Serra M, Meza-Zepeda LA, Myklebost O, Hassan AB, Hogendoorn PCW, Cleton-Jansen AM. Accepted for publication in *BMC Cancer*

Abstract

Background: High-grade osteosarcoma is an aggressive tumor most often developing in the long bones of adolescents, with a second peak in the 5th decade of life. Better knowledge on cellular signaling in this tumor may identify new possibilities for targeted treatment.

Methods: We performed gene set analysis on previously published genome-wide gene expression data of osteosarcoma cell lines ($n = 19$) and pretreatment biopsies ($n = 84$). We characterized overexpression of the insulin-like growth factor receptor (IGF1R) signaling pathways in human osteosarcoma as compared with osteoblasts and with the hypothesized progenitor cells of osteosarcoma—mesenchymal stem cells. This pathway plays a key role in the growth and development of bone. Since most profound differences in mRNA expression were found at and upstream of the receptor of this pathway, we set out to inhibit IR/IGF1R using OSI-906, a dual inhibitor for IR/IGF1R, on four osteosarcoma cell lines. Inhibitory effects of this drug were measured by Western blotting and cell proliferation assays.

Results: OSI-906 had a strong inhibitory effect on proliferation of 3 of 4 osteosarcoma cell lines, with IC_{50}s below 100nM at 72hrs of treatment. Phosphorylation of IRS-1, a direct downstream target of IGF1R signaling, was inhibited in the responsive osteosarcoma cell lines.

Conclusions: This study provides an *in vitro* rationale for using IR/IGF1R inhibitors in preclinical studies of osteosarcoma.

Background

High-grade osteosarcoma is the most prevalent primary malignant bone tumor. The disease occurs most frequently in children and adolescents at the site where proliferation is most active, *i.e.* the metaphysis adjacent to the epiphyseal plate (1). The 5-year overall survival of osteosarcoma patients has raised from 10–20% to about 60% after the introduction of preoperative chemotherapy in the 1970s. However, about 45% of all patients still die because of distant metastasis. No additional treatments have been found that can increase survival significantly, and administering higher doses of preoperative chemotherapy does not result in improved outcomes (2, 3). Better knowledge on cellular signaling in high-grade osteosarcoma may identify new possibilities for targeted treatment of this highly aggressive tumor.

We have previously described the roles of bone developmental pathways Wnt, TGFβ-BMP, and Hedgehog signaling in osteosarcoma, but unfortunately so far could not identify suitable targets for treatment (4, 5). In addition to these signal transduction pathways, insulin-like growth factor 1 receptor (IGF1R) signaling plays a key role in the growth and development of bone. Aberrant signaling of this pathway has been implicated in various

cancer types, among others sarcomas (6, 7). Key players of insulin-like growth factor (IGF) signaling are the ligands IGF1, IGF2, which are circulating polypeptides that can be expressed in endocrine, paracrine, and autocrine manners, and the tyrosine kinase receptor IGF1R, which forms homodimers, or hybrid receptors with the insulin receptor (IR) (8). IGF1R and IR/IGF1R hybrids are activated by both IGF1 and -2, which trigger autophosphorylation of IGF1R and subsequent downstream signal transduction. A second IGF receptor, IGF2R, can bind IGF2, but does not confer intracellular signaling, thereby diminishing the bioavailability of IGF2 to IGF1R (9). Autophosphorylation of IR/IGF1R receptors recruits the signaling proteins insulin receptor substrate (IRS) and Src homology 2 domain containing transforming protein (Shc) to the cell membrane, which get phosphorylated and subsequently activate the downstream PI3K/Akt and Ras/Raf/ERK signaling pathways, both of which are known to be important in cancer. These pathways ultimately act on several biological processes, such as transcription, proliferation, growth, and survival (9–11). Interestingly, treatment targeted against IGF1R signaling has shown to be effective in a subset of Ewing sarcoma, another bone tumor that manifests at young age (12).

The role of the IGF1R pathway in growth has been illustrated in studies of knockout mice. It was shown that IGF1 null mice are 40% smaller than littermates, while IGF1R null mice are approximately 55% smaller (13). In dogs, the size of different breeds was demonstrated to be dependent on IGF1 plasma levels (7). Additionally, a specific IGF1 SNP haplotype was described to be common in small breed dogs and nearly absent in giant breeds (14). Interestingly, large and giant dog breeds are more prone to develop osteosarcoma (15), which in dogs is biologically very similar to the human disease (16). Two recent studies on human osteosarcoma suggest a positive correlation between patient birth weight and height at diagnosis and the development of the disease (17, 18). Involvement of some members of IGF1R signaling in osteosarcoma has been described (as has been reviewed in Kolb *et al.* (19)), but the activity of this pathway remains to be determined.

We have analyzed genome-wide gene expression in high-grade osteosarcoma cell lines and pretreatment biopsies, and observed significantly altered activity of genes involved in IGF1R signaling when compared to profiles of mesenchymal stem cells and osteoblasts. Specifically, upstream inhibitors of IGF1R signaling were found to be downregulated in osteosarcoma, and low expression of these genes correlated with worse event-free survival. We inhibited IR/IGF1R signaling with the dual IR/IGF1R inhibitor OSI-906. This showed inhibition of phosphorylation of IRS-1 and of strong inhibition of proliferation in 3/4 osteosarcoma cell lines. Interestingly, the cell line which could not be inhibited with OSI-906, 143B, has a *KRAS* oncogenic transformation, which is a component of the Ras/Raf/ERK pathway, one of downstream effectors of IGF1R signaling. These results suggest that IR/IGF1R signaling may be an effective targeted for treatment of high-grade osteosarcoma patients.

Methods

Cell culture

The 19 high-grade osteosarcoma cell lines that were used in this study were characterized and are described by Ottaviano *et al.* (20). The 12 mesenchymal stem cell and 3 osteoblast cultures were previously described (21). MSCs have been previously (22) characterized through FACS analysis and have been tested for their ability to be committed under proper conditions towards adipogenesis, chondrogenesis and osteogenesis as described in Bernardo *et al.* (23). Osteoblast cultures were derived from MSCs which were treated to undergo osteogenic differentiation. Cell line DNA was short tandem repeat profiled to confirm cell line identity with use of the Cell ID system of Promega (Madison, WI). For Western blotting experiments, cells were maintained in RPMI 1640 (Invitrogen, Carlsbad, CA), supplemented with 10% fetal bovine serum (F7524, Sigma-Aldrich, St. Louis, MO) and 1% glutamax (Gibco 35050, Invitrogen, Carlsbad, CA).

Microarray experiments, preprocessing, and data analysis

For genome-wide gene expression analysis, we used Illumina Human-6 v2.0 BeadChips. Microarray experiments and data preprocessing are described in Kuijjer *et al.* (21). Previously deposited genome-wide gene expression data of mesenchymal stem cells (MSCs) and osteoblasts can be found in the Gene Expression Ombinus (GEO accession numbers GSE28974 and GSE33382, respectively). Data from osteosarcoma cell lines have been published before (24), but since we normalized and processed all raw data together, we deposited normalized values in the Gene Expression Omnibus (GEO, accession number GSE42351, superseries accession GSE42352). Data from the 84 high-grade osteosarcoma pretreatment biopsies have been previously published (GEO accession number GSE33382) (21). Ethical guidelines of the individual European partner institutions were followed and samples and clinical data were handled in a coded fashion and stored in the EuroBoNeT biobank. We determined significant differential expression between osteosarcoma cell lines ($n = 19$) and mesenchymal stem cells ($n = 12$), and between osteosarcoma cell lines and osteoblasts ($n = 3$) using Bioconductor (25) package *LIMMA* (26) in statistical language R (27). Probes with Benjamini and Hochberg false discovery rate-adjusted p-values < 0.05 were considered to be significant. Gene set analysis was performed on KEGG pathways (28) (Release 63.0, July 1, 2012) using R-package globaltest (29). For each analysis, the top 15 significant KEGG pathways were returned. All returned pathways had a Benjamini and Hochberg false-discovery rate-corrected p-value $< 1 \cdot 10^{-5}$. To visualize differential expression in the IGF1R pathway, we performed Core analyses using Ingenuity Pathways Analysis (IPA, Ingenuity Systems, `www.ingenuity.com`).

Antibodies and reagents

Rabbit monoclonal and polyclonal antibodies against IGF1R and IRS-1, respectively (both $1:1,000$) were obtained from Cell Signaling (Danvers, MA). Rabbit polyclonal antibody against phospho-IRS-1 (Y612, $1:1,000$) was purchased from Biosource, Invitrogen (Carlsbad, CA). A mouse monoclonal antibody against α-tubulin from Abcam (Cambridge, UK) was used as a loading control ($1:3,000$). Secondary antibodies (both $1:10,000$, BD Transduction Laboratories, Lexington, KY) were horseradish peroxidase (HRP) conjugated polyclonal goat-anti-rabbit IgG for components of the IR/IGF1R pathway, and HRP conjugated polyclonal goat-anti-mouse for α-tubulin. OSI-906 was purchased from Selleck Chemicals LLC (Houston, TX).

Western blotting

Osteosarcoma cell lines OHS, KPD, SAOS-2, and 143B were treated with 0.5% DMSO or with 1μM OSI-906 for 3hrs, and were subsequently lysed using Mammalian Protein Extraction Reagent (Thermo Scientific 78503), to which Halt Phosphatase and Protease Inhibitor Cocktails (Thermo Scientific 78420 and 78418, respectively) were added according to the manufacturer's protocol. Concentrations of cell lysates were determined using the BioRad DC Protein Assay Kit (Biorad, Hercules, CA). Per sample, 20μg of protein was loaded on SDS-PAGE gels. Lysate of HepG2-A16 cells transfected with IR and stimulated with insulin, containing 10μg of protein, was taken along as a positive control. Western blotting was performed as described by Schrage *et al.* (30).

Proliferation assays

OSI-906 was diluted in DMSO and stored at -20°C. OHS, SAOS-2, KPD, and 143B cells were plated in 96 wells plates, using $4,000$, $2,000$, $12,000$, and $2,000$ cells per well, respectively. After 24hrs, OSI-906 was added in triplicate at different concentrations—0nM, 0.01nM, 0.1nM, 1nM, 10nM, 100nM, 1μM, and 10μM. The inhibitor was incubated for 72hrs and 96hrs, in different experiments. The results shown are representative results from at least three independent experiments. Cell proliferation reagent WST-1 (Roche) was incubated for 2hrs and subsequently measured using a Wallac 1420 VICTOR2 (Perkin Elmer, Waltham, MA). Data were analyzed in Graphpad Prism 5.0 (`www.graphpad.com`). Relative IC_{50}s were calculated using results from the different concentrations up to the highest dose where toxicity was not yet present.

Results

Enrichment of IGF1R signaling in high-grade osteosarcoma

Genome-wide gene expression data were of good quality for all cell lines. *LIMMA* analysis resulted in 7,891 probes encoding for differentially expressed (DE) genes between osteosarcoma cell lines and MSCs, and 2,222 probes encoding for DE genes between osteosarcoma cells and osteoblasts. We tested the global expression patterns of KEGG pathways using globaltest (29) and determined the intersection of the pathways most significantly different in osteosarcoma cell lines as compared with MSCs, and of osteosarcoma cell lines as compared with osteoblasts. This approach resulted in five significantly affected pathways—insulin signaling pathway, oocyte meiosis, ubiquitin mediated proteolysis, progesterone-mediated oocyte maturation, and glycerophospholipid metabolism. Details of the globaltest are shown in Table 5.1.

KEGG pathway	Analysis	adjP	Statistic	Expected	Std.dev
Insulin signaling	OScellvsOB	$1.01 \cdot 10^{-7}$	26.34	4.76	1.92
pathway	OScellvsMSC	$3.07 \cdot 10^{-15}$	35.12	3.33	1.78
Oocyte meiosis	OScellvsOB	$2.70 \cdot 10^{-7}$	37.45	4.76	2.90
	OScellvsMSC	$5.04 \cdot 10^{-16}$	53.70	3.33	2.84
Ubiquitin mediated	OScellvsOB	$3.21 \cdot 10^{-7}$	22.88	4.76	1.75
proteolysis	OScellvsMSC	$5.04 \cdot 10^{-16}$	37.99	3.33	1.89
Progesterone-mediated	OScellvsOB	$7.16 \cdot 10^{-7}$	34.26	4.76	2.71
oocyte maturation	OScellvsMSC	$1.34 \cdot 10^{-15}$	55.35	3.33	2.77
Glycerophospholipid	OScellvsOB	$1.40 \cdot 10^{-6}$	27.13	4.76	2.25
metabolism	OScellvsMSC	$2.25 \cdot 10^{-15}$	55.86	3.33	2.82

Table 5.1: The top five significant pathways with aberrant expression in both osteosarcoma cell lines versus osteoblasts (OScellvsOB) and osteosarcoma cell lines versus mesenchymal stem cells (OScellvsMSC). adjP: FDR-adjusted p-value, Statistic: test statistic of the globaltest, Expected: expected test statistic of the globaltest, Std.dev: standard deviation under the null hypothesis.

IGF1R signaling is involved in three out of the five detected KEGG pathways (insulin signaling pathway, oocyte meiosis, and progesterone-mediated oocyte maturation). Interestingly, a globaltest on mRNA expression of previously published pretreatment biopsies (21) compared with normal bones (31) also returned insulin signaling as the most significantly affected pathway (*data not shown*). Notably, there is no specific IGF1R signaling pathway in the KEGG database (28). Because of the overrepresentation of IGF1R signaling, and because of its known role in cancer, we decided to study expression of members of this pathway in detail.

Differentially expressed genes of the IGF1R pathway

To determine which genes have the most specific up- or downregulation in osteosarcoma, we combined lists of significantly differentially expressed genes of osteosarcoma cell lines ($n = 19$) and a previously published set of osteosarcoma pretreatment biopsies ($n = 84$,

GEO accession GSE33382) in comparison with mesenchymal stem cells ($n = 12$) and osteoblasts ($n = 3$) by four-way Venn analysis of all significantly affected probes with the same direction of fold change (upregulated or downregulated in all four analyses). We identified *IGFBP4* and *GAS6* as the most downregulated genes in osteosarcoma (average log fold changes of -4.43 and -4.29, respectively). *IGFBP2* was also present in the top 20 results from this four-way analysis. In addition, *IGFBP3* and -*7* were significantly downregulated, and *IGF2BP3* was significantly upregulated in three out of the four analyses. Both *IGFBP4* and *GAS6* show high variability in expression in osteosarcoma cell lines and biopsies (Figure 5.1A). Patients whose biopsies had very low expression of these genes

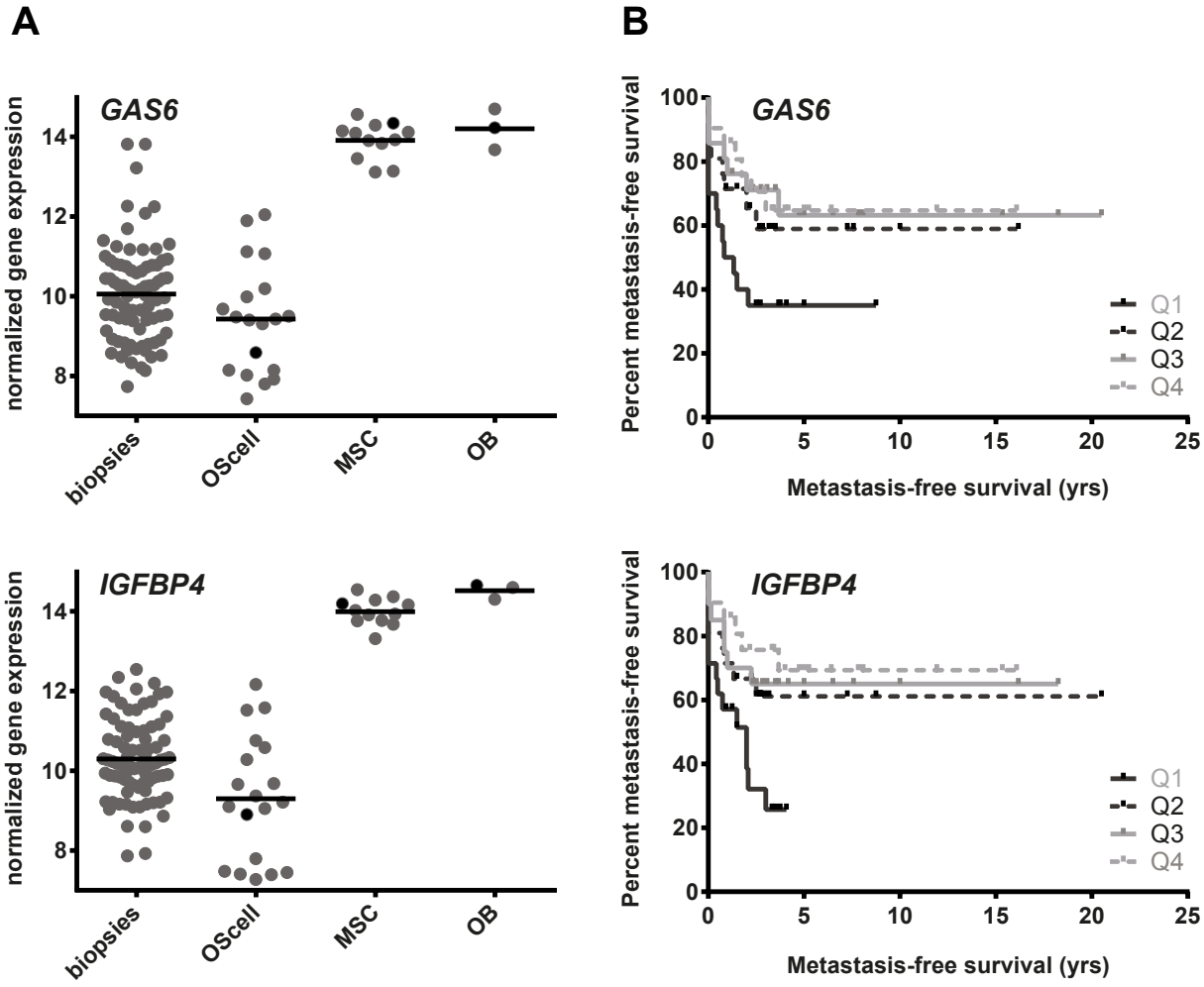


Figure 5.1: *A*, Normalized gene expression levels of *GAS6* and *IGFBP4* in osteosarcoma biopsies, cell lines, mesenchymal stem cells (MSCs), and osteoblasts (OB). Expression of both proteins is considerably higher in the controls (FDR-adjusted p-value < 0.001 for both genes in all four analyses). *B*, Kaplan-Meier curves depicting metastasis-free survival in years for 83 high-grade osteosarcoma patients (for 1/84 patients, we did not have follow-up data available), based on quartiles of mRNA expression of the genes of interest.

had poor event-free survival profiles (Logrank test for trend, p-value $= 0.01$ for *IGFBP4* and p-value $= 0.04$ for *GAS6*, Figure 5.1B). To visualize mRNA expression of the IGF1R signaling pathway members, we used Ingenuity Pathways Analysis on *LIMMA* toptables

from osteosarcoma cells as compared with mesenchymal stem cells and from osteosarcoma cells as compared with osteoblasts (Figure 5.2). As can be seen in this figure, overlap of differentially expressed genes between these analyses was detected upstream of IGF1R.

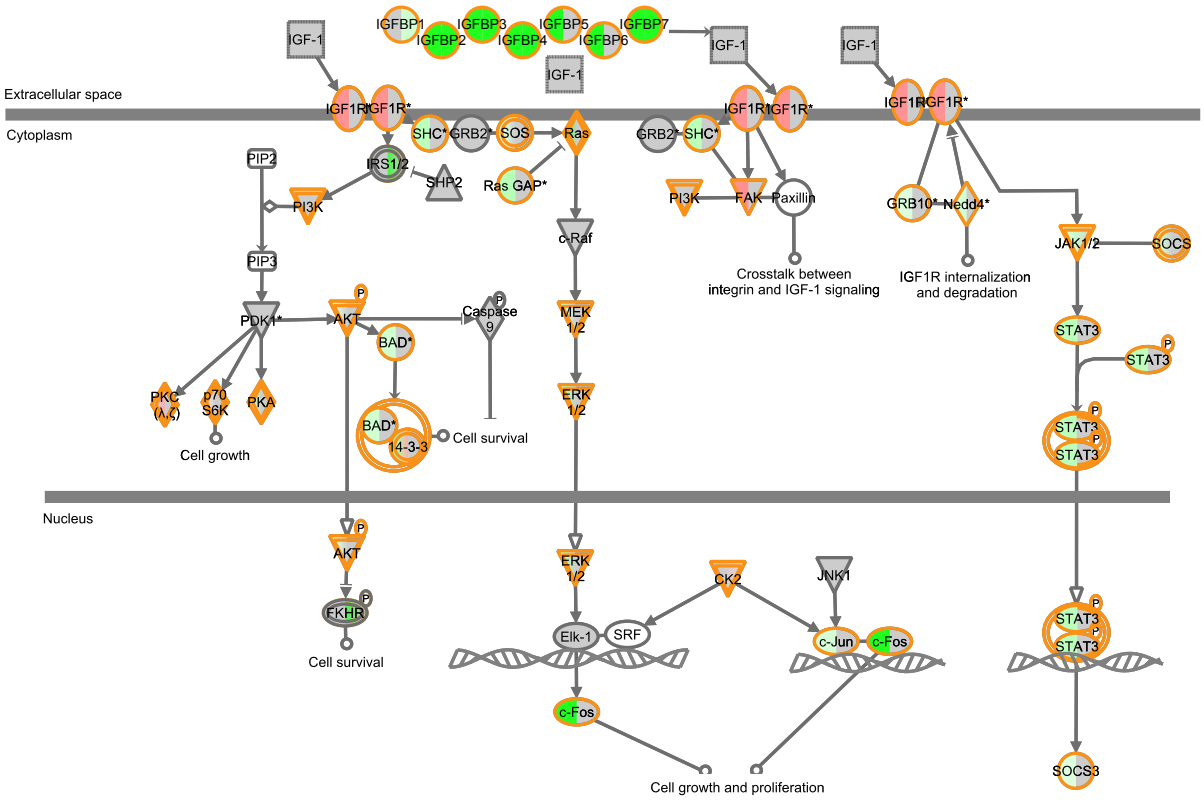


Figure 5.2: This figure shows the IGF1R signaling pathway, with significantly upregulated genes in red, downregulated genes in green, and genes that did not meet our criteria for significance in gray. The left part of the symbols shows the analysis of osteosarcoma cell lines as compared with mesenchymal stem cells, the right part as compared with osteoblasts. Most consensus in gene expression is found upstream IGF1R signaling, in the expression of the IGF binding proteins.

OSI-906 inhibits phosphorylation of IRS-1

Gene expression levels of IGF1R and IRS-1 were validated at the protein level by Western blot analysis (*data not shown*). We used phosphorylated IRS-1 as a readout for IR/IGF1R signal transduction activity, as IRS-1 is a direct downstream target of these receptors. We performed Western blot analysis on cell lysates of OHS, KPD, SAOS-2, and 143B, using antibodies against IRS-1 and phosphorylated IRS-1, before and after treatment with OSI-906—a selective small molecule dual kinase inhibitor of both IR and IGF1R. An inhibition of intrinsic IRS-1 phosphorylation at Y612 was detected after treatment with OSI-906 in all cell lines (Figure 5.3), indicating that this inhibitor could affect signaling downstream IGF1R in osteosarcoma cells.

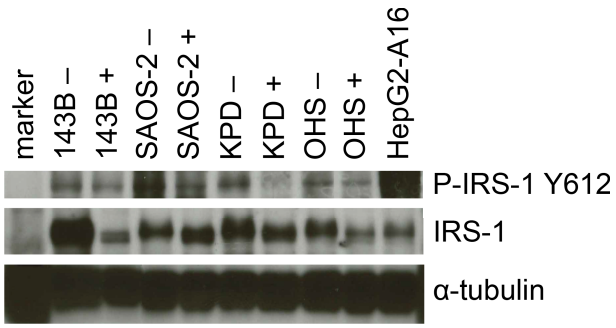


Figure 5.3: Western blot of IRS-1 and p-IRS-1 of lysates of untreated (–) osteosarcoma cell lines OHS, KPD, SAOS-2, and 143B, and of these cells treated for 3hrs with 1μM of OSI-906 (+).

OSI-906 inhibits proliferation of 3 of 4 osteosarcoma cell lines

In 3 of 4 osteosarcoma cell lines tested, inhibition with OSI-906 was dose-dependent (Figure 5.4). Except for a toxic response at the maximum dose of 10μM (*data not shown*), there was no effect on 143B. Because of this toxicity, relative IC_{50}s were determined using measurements until 1μM. OHS, SAOS-2, and KPD had an IC_{50} of 25nM, 92nM, and 90nM at 72hrs, respectively, and of 37nM, 57nM, and 23nM at 96hrs of inhibition, respectively. At 1μM OSI-906, approximately 60% of proliferation of OHS, SAOS-2, and KPD cells was inhibited, while 143B proliferation was not inhibited (Figure 5.4).

Discussion

Genome-wide gene expression and subsequent gene set analysis on osteosarcoma cell lines and biopsies indicated increased insulin-like growth factor signaling in high-grade osteosarcoma as compared with the hypothesized osteosarcoma progenitors, which is currently the best control, since there is no benign precursor and no certainty of the normal counterpart of osteosarcoma. Because IGF1R signaling can be exploited as a therapeutic target, and osteosarcoma patients are in severe need of new therapies, we examined mRNA expression of members of this signaling pathway in detail. *IGFBP4* and *GAS6*, which code for proteins that inhibit IGF1R signaling, showed the highest significant downregulation (log fold changes < -4) in a four-way analysis, in which osteosarcoma pretreatment biopsies or cell lines were compared with osteoblastic cultures ($n = 3$) or MSCs ($n = 12$). Insulin-like growth factor binding proteins (IGFBPs) generally inhibit IGF1R signaling by competitively binding IGFs, but can under certain circumstances also stimulate IGF1R signaling (32). IGFBP4 is a negative regulator of IGF signaling in various tissues, including bone (33). GAS6, or growth arrest-specific 6, was shown to inhibit the growth promoting effects of IGF signaling and to stimulate differentiation in the chondrogenic cell line ATDC5 (34). Both *IGFBP4* and *GAS6* expression have previously been shown to be downregulated in osteosarcoma cell lines (*IGFBP4* in MG-63 (35), *GAS6* in MG-63 and SAOS-2 cells (36)). Next to *GAS6* and *IGFBP4*, *IGFBP2* was also significantly downregulated in all four analyses, with log fold changes of approximately −3. IGFBP2 generally inhibits IGF action and may play a role in IGF2-induced osteoblast differentia-

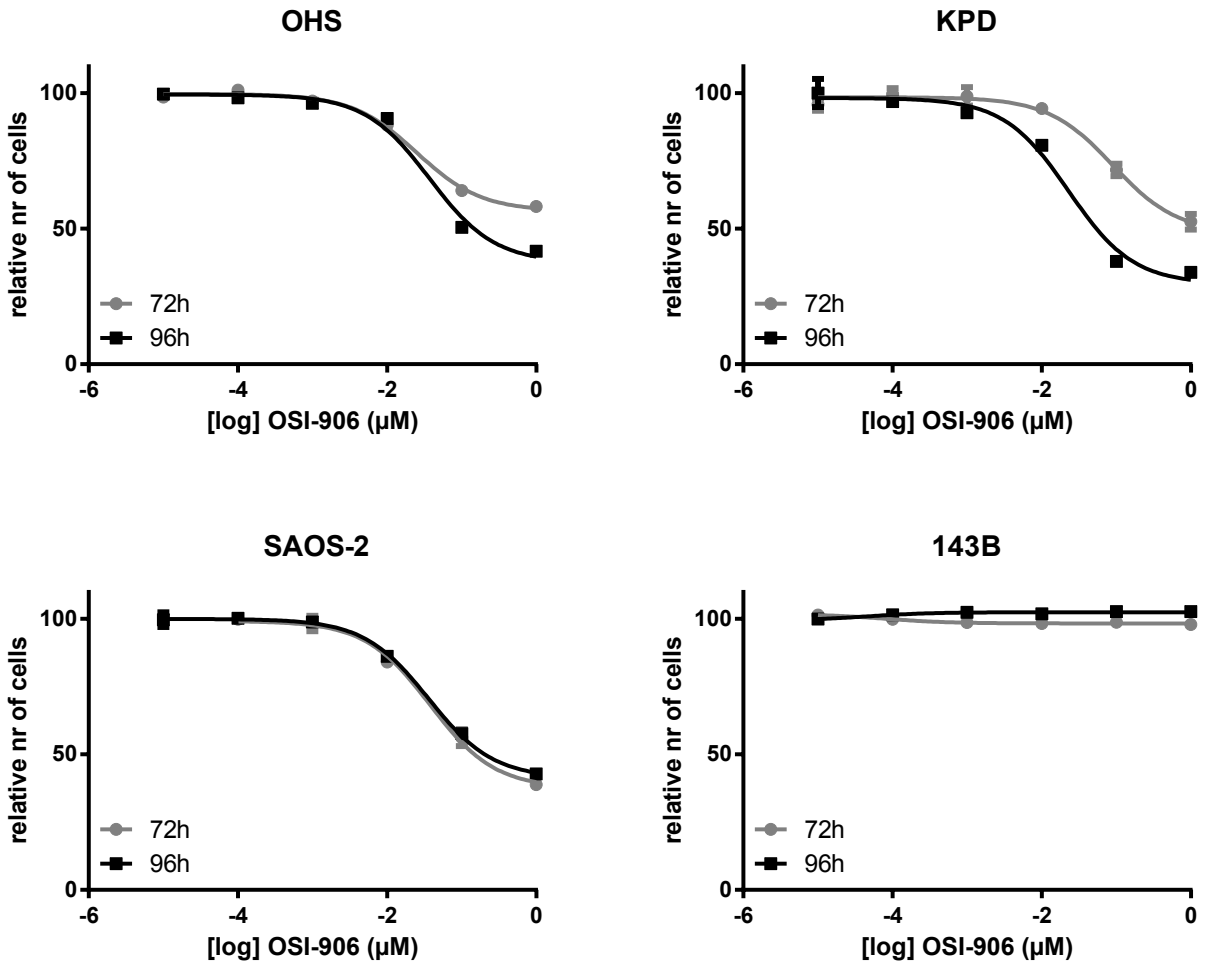


Figure 5.4: Osteosarcoma cell lines were inhibited with different concentrations of OSI-906, for 72 (gray line) or 96 (black line) hours. OHS (*A*), KPD (*B*), and SAOS-2 (*C*) showed a dose-dependent inhibition, while 143B (*D*) did not respond to OSI-906.

tion (33). *IGFBP3* was highly downregulated in three out of four analyses, and has been shown to elicit anticancer effects by inhibiting IGF1R signaling in Ewing sarcoma (37). IGFBP7 activity has not yet been reported in sarcoma, but has been associated with *e.g.* hepatocellular carcinoma (38). Interestingly, *IGF2BP3* was highly overexpressed in 3 of 4 analyses. This binding protein can bind IGF2 mRNA, thereby probably activating the translation of IGF2 (39). Overexpression of *IGF2BP3* has been reported in several cancer types (40, 41). Figure 5.2 shows that differential expression is most pronounced in upstream regulators of IGF1R, while downstream components, such as SHC and FOS, are slightly downregulated, although for most genes this only holds when compared with mesenchymal stem cells, and not with osteoblasts. This may be caused by negative feedback loops, triggered by the active IGF1R signaling pathway. These results suggest that, in osteosarcoma, the IGF1R signaling pathway can be inhibited at the level of the receptor. We therefore validated protein levels of IGF1R and of IRS-1, a direct downstream component of IGF1R and IR signaling using Western blotting. IGF1R and IRS-1 protein levels correlated fairly well with mRNA expression levels. Most importantly, phosphorylated IRS-1, which is a measure for pathway activity, was detected in all four osteosarcoma cell lines, indicating that IGF1R signaling is active in osteosarcoma, and is possibly regulated upstream of IGF1R. Accordingly, targeting this receptor may be an effective way to inhibit this pathway.

OSI-906 is a selective small molecule dual kinase inhibitor of both IR and IGF1R (42). We specifically chose to treat osteosarcoma cells with a dual inhibitor, because the insulin receptor can activate the same downstream signaling pathways as IGF1R, therefore providing cells a way to circumvent single inhibition of IGF1R. This has formerly been demonstrated in osteoblasts (43) and in Ewing sarcoma cells (44). In fact, this dual inhibitor has been shown to cause enhanced inhibition of the Akt signaling pathway when compared with a selective monoclonal antibody against IGF1R, which could inhibit IR/IGF1R hybrids, but not IR homodimers (45). OSI-906 is currently being tested by OSI Pharmaceuticals in a Phase III trial in adrenocortical carcinoma and in a Phase I/II clinical trial in ovarian cancer. Treatment of osteosarcoma cells with OSI-906 at physiological levels leads to decreased phosphorylation of IRS-1 at Y612. Inhibition of IRS-1 at Y612 after treatment with OSI-906 was previously reported by Buck *et al.* in direct complementation breast cancer cells for IGF1R-IGF2 and IR(A)-IGF2 (45). Interestingly, we also detected a small shift in the size of p-IRS-1 on the Western Blot, indicating that multiple phosphorylation groups are removed after treatment with OSI-906. Surprisingly, total IRS-1 levels were highest in 143B, and were downregulated after treatment with OSI-906 in this cell line, although this had no effect on cell growth in this line, as opposed to the three others, which showed low IC_{50}s. Proliferation of 143B was only inhibited most likely unspecifically at high and toxic levels of the drug. The 143B cell line is a derivative of the osteosarcoma cell line HOS, transformed by a *KRAS* oncogene. Constitutive activation of the Ras/Raf/ERK pathway can explain why proliferation of this cell

line cannot be inhibited by OSI-906. Of the cell lines that were responsive to OSI-906, KPD and OHS showed that treatment of 96hrs was most effective, while SAOS-2 already reached maximum inhibition at 72hrs.

IGF1R signaling has been previously modulated in sarcoma in preclinical and clinical models. Several phase I and II clinical trials including treatment with IGF1R monoclonal antibodies are currently being conducted in sarcoma, especially in Ewing sarcoma (an overview of these trials is given in Olmos *et al.* (46)). Monoclonal antibodies against IGF1R have modest activity against Ewing sarcoma, as was observed in a phase I/II study of figitumumab (partial response in 14.2% of all subjects) (47) and in a phase II study using R1507 (complete/partial response rate of 10%) (48). Results of a phase II study of ganitumab in subjects with Ewing sarcoma and desmoplastic small round cell tumors were published very recently, and reported clinical benefit in 17% of all patients (49). Preclinically, treatment with different monoclonal antibodies against IGFR1 has been performed in osteosarcoma xenograft models, in which a response was detected in at least 60% of all cases studied (50–52). However, no objective responses were observed in phase I trials testing monoclonal antibodies in osteosarcoma (47, 53, 54), although 2 of 3 patients treated with R1507 had prolonged stable disease (53). Clinical data using dual IGF1R/IR inhibitors osteosarcoma is still very limited (55). Because resistance to highly specific IGF1R inhibitors may develop through IR (44), blocking both IGF1R and IR with a dual kinase inhibitor will most likely lead to better inhibition of downstream IRS-1 signaling. We thus expect clinical outcomes to improve for osteosarcoma patients treated with dual IGF1R/IR inhibitor OSI-906. The effects of combination of OSI-906 with chemotherapeutics in osteosarcoma still need to be assessed before such a treatment can be clinically tested.

Phosphorylated IRS could be used as a biomarker in order to determine whether patients would respond to IGF1R inhibition. Patients with tumors exhibiting an activating mutation in downstream pathways will most likely not respond to IGF1R inhibition. Further research needs to be performed in order to assess these candidate biomarkers for response to treatment. The IGF1R pathway acts on several biological mechanisms that promote tumor progression—mitogenesis, protection from apoptosis, malignant transformation, and metastasis (6). It is therefore possible that inhibiting these pathways with a dual IR/IGF1R kinase inhibitor, such as OSI-906, may reduce tumor sizes, as well as osteosarcoma metastasis, the leading cause of death in these patients.

Conclusions

Using gene set analysis of genome-wide gene expression data of high-grade osteosarcoma biopsies and cell lines, we detected an overrepresentation of IGF1R signaling. Specifically, different upstream inhibitors of IGF1R signaling, *e.g.* several IGF binding proteins, were downregulated. As this indicated the IGF1R receptor as a potential target for treatment

of osteosarcoma, we set out to inhibit this receptor in four osteosarcoma cell lines. We used OSI-906, a selective small molecule dual kinase inhibitor of both IR and IGF1R, since the insulin receptor can activate the same downstream signaling pathways as IGF1R, thereby providing a way to circumvent single inhibition of IGF1R. Treatment with OSI-906 resulted in inhibition of phosphorylation of IRS-1 Y612, a direct downstream target of IGF1R, and in strong inhibition of proliferation in 3 of 4 osteosarcoma cell lines. The nonresponsive cell line, 143B, has a *KRAS* oncogenic transformation, and may therefore not respond to this treatment. In conclusion, we have shown that IGF1R signaling is active in osteosarcoma, and that dual inhibition of IR/IGF1R inhibits downstream signaling and proliferation of these cells. Responsiveness to this treatment may be evaluated by Western blotting against phosphorylated IRS. This study provides an *in vitro* rationale for using dual IR/IGF1R inhibitors in preclinical studies of osteosarcoma.

References

[1] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[2] Lewis IJ, Nooij MA, Whelan J, Sydes MR, *et al.* Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup. *Journal of the National Cancer Institute*, 99(2):112–128, 2007.

[3] Eselgrim M, Grunert H, Kühne T, Zoubek A, *et al.* Dose intensity of chemotherapy for osteosarcoma and outcome in the Cooperative Osteosarcoma Study Group (COSS) trials. *Pediatric Blood & Cancer*, 47(1):42–50, 2006.

[4] Cai Y, Mohseny AB, Karperien M, Hogendoorn PCW, *et al.* Inactive Wnt/β-catenin pathway in conventional high-grade osteosarcoma. *The Journal of Pathology*, 220(1):24–33, 2010.

[5] Mohseny AB, Cai Y, Kuijjer ML, Xiao W, *et al.* The activities of Smad and Gli mediated signalling pathways in high-grade conventional osteosarcoma. *European Journal of Cancer*, 48(18):3429–3438, 2012.

[6] Rikhof B, de Jong S, Suurmeijer AJH, Meijer C, *et al.* The insulin-like growth factor system and sarcomas. *The Journal of Pathology*, 217(4):469–482, 2009.

[7] Maki RG. Small is beautiful: insulin-like growth factors and their role in growth, development, and cancer. *Journal of Clinical Oncology*, 28(33):4985–4995, 2010.

[8] Pollak M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nature Reviews Cancer*, 12(3):159–169, 2012.

[9] Siddle K. Molecular basis of signaling specificity of insulin and IGF receptors: neglected corners and recent advances. *Frontiers in Endocrinology*, 3:34, 2012.

[10] Foulstone E, Prince S, Zaccheo O, Burns JL, *et al.* Insulin-like growth factor ligands, receptors, and binding proteins in cancer. *The Journal of Pathology*, 205(2):145–153, 2005.

[11] Siddle K. Signalling by insulin and IGF receptors: supporting acts and new players. *Journal of Molecular Endocrinology*, 47(1):R1–R10, 2011.

[12] Subbiah V, Anderson P. Targeted therapy of Ewing's sarcoma. *Sarcoma*, 2011:686985, 2010.

[13] Liu JP, Baker J, Perkins AS, Robertson EJ, *et al.* Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). *Cell*, 75(1):59–72, 1993.

[14] Sutter NB, Bustamante CD, Chase K, Gray MM, *et al.* A single IGF1 allele is a major determinant of small size in dogs. *Science Signalling*, 316(5821):112, 2007.

[15] Selvarajah GT, Kirpensteijn J. Prognostic and predictive biomarkers of canine osteosarcoma. *The Veterinary Journal*, 185(1):28–35, 2010.

[16] Kirpensteijn J, Kik M, Teske E, Rutteman GR. TP53 gene mutations in canine osteosarcoma. *Veterinary Surgery*, 37(5):454–460, 2008.

[17] Arora RS, Kontopantelis E, Alston RD, Eden TO, *et al.* Relationship between height at diagnosis and bone tumours in young people: a meta-analysis. *Cancer Causes and Control*, 22(5):681–688, 2011.

[18] Mirabello L, Pfeiffer R, Murphy G, Daw NC, *et al.* Height at diagnosis and birth-weight as risk factors for osteosarcoma. *Cancer Causes and Control*, 22(6):899–908, 2011.

[19] Kolb EA, Gorlick R. Development of IGF-IR inhibitors in pediatric sarcomas. *Current Oncology Reports*, 11(4):307–313, 2009.

[20] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[21] Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, *et al.* Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes, Chromosomes and Cancer*, 51(7):696–706, 2012.

[22] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

[23] Bernardo ME, Emons JAM, Karperien M, Nauta AJ, *et al.* Human mesenchymal stem cells derived from bone marrow display a better chondrogenic differentiation compared with other sources. *Connective Tissue Research*, 48(3):132–140, 2007.

[24] Namløs HM, Meza-Zepeda LA, Barøy T, Østensen IHG, *et al.* Modulation of the Osteosarcoma Expression Phenotype by MicroRNAs. *PloS One*, 7(10):e48086, 2012.

[25] Gentleman RC, Carey VJ, Bates DM, Bolstad B, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004.

[26] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):3, 2004.

[27] Team RDC. R: a language and environment for statistical computing, reference index version 2.15.0. R Foundation for Statistical Computing, 2011.

[28] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27–30, 2000.

[29] Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics*, 20(1):93–99, 2004.

[30] Schrage YM, Briaire-de Bruijn IH, de Miranda NFCC, van Oosterwijk JG, *et al.* Kinome profiling of chondrosarcoma reveals SRC-pathway activity and dasatinib as option for treatment. *Cancer Research*, 69(15):6216–6222, 2009.

[31] Namløs HM, Kresse SH, Müller CR, Henriksen J, *et al.* Global gene expression profiling of human osteosarcomas reveals metastasis-associated chemokine pattern. *Sarcoma*, 2012:639038, 2012.

[32] Grimberg A, Cohen P. Role of insulin-like growth factors and their binding proteins in growth control and carcinogenesis. *Journal of Cellular Physiology*, 183(1):1–9, 2000.

[33] Conover CA. Insulin-like growth factor-binding proteins and bone metabolism. *American Journal of Physiology-Endocrinology and Metabolism*, 294(1):E10–E14, 2008.

[34] Hutchison MR, Bassett MH, White PC. SCF, BDNF, and Gas6 are regulators of growth plate chondrocyte proliferation and differentiation. *Molecular Endocrinology*, 24(1):193–203, 2010.

[35] Scharla SH, Strong DD, Rosen C, Mohan S, *et al.* 1,25-Dihydroxyvitamin D3 increases secretion of insulin-like growth factor binding protein-4 (IGFBP-4) by human osteoblast-like cells in vitro and elevates IGFBP-4 serum levels in vivo. *Journal of Clinical Endocrinology & Metabolism*, 77(5):1190–1197, 1993.

[36] Shiozawa Y, Pedersen EA, Patel LR, Ziegler AM, *et al.* GAS6/AXL axis regulates prostate cancer invasion, proliferation, and survival in the bone marrow niche. *Neoplasia*, 12(2):116–127, 2010.

[37] Benini S, Zuntini M, Manara MC, Cohen P, *et al.* Insulin-like growth factor binding protein 3 as an anticancer molecule in Ewing's sarcoma. *International Journal of Cancer*, 119(5):1039–1046, 2006.

[38] Chen D, Yoo BK, Santhekadur PK, Gredler R, *et al.* Insulin-like growth factor–binding protein-7 functions as a potential tumor suppressor in hepatocellular carcinoma. *Clinical Cancer Research*, 17(21):6693–6701, 2011.

[39] Liao B, Hu Y, Herrick DJ, Brewer G. The RNA-binding protein IMP-3 is a translational activator of insulin-like growth factor II leader-3 mRNA during proliferation of human K562 leukemia cells. *Journal of Biological Chemistry*, 280(18):18517–18524, 2005.

[40] Schaeffer DF, Owen DR, Lim HJ, Buczkowski AK, *et al.* Insulin-like growth factor 2 mRNA binding protein 3 (IGF2BP3) overexpression in pancreatic ductal adenocarcinoma correlates with poor survival. *BMC Cancer*, 10(1):59, 2010.

[41] Suvasini R, Shruti B, Thota B, Shinde SV, *et al.* Insulin growth factor-2 binding protein 3 (IGF2BP3) is a glioblastoma-specific marker that activates phosphatidylinositol 3-kinase/mitogen-activated protein kinase (PI3K/MAPK) pathways by modulating IGF-2. *Journal of Biological Chemistry*, 286(29):25882–25890, 2011.

[42] Mulvihill MJ, Cooke A, Buck E, Foreman K, *et al.* Discovery of OSI-906: a selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor. *Future Medicinal Chemistry*, 1(6):1153–1171, 2009.

[43] Fulzele K, DiGirolamo DJ, Liu Z, Xu J, *et al.* Disruption of the insulin-like growth factor type 1 receptor in osteoblasts enhances insulin signaling and action. *Journal of Biological Chemistry*, 282(35):25649–25658, 2007.

[44] Garofalo C, Manara MC, Nicoletti G, Marino MT, *et al.* Efficacy of and resistance to anti-IGF-1R therapies in Ewing's sarcoma is dependent on insulin receptor signaling. *Oncogene*, 30(24):2730–2740, 2011.

[45] Buck E, Gokhale PC, Koujak S, Brown E, *et al.* Compensatory insulin receptor (IR) activation on inhibition of insulin-like growth factor-1 receptor (IGF-1R): rationale for cotargeting IGF-1R and IR in cancer. *Molecular Cancer Therapeutics*, 9(10):2652–2664, 2010.

[46] Olmos D, Tan DSW, Jones RL, Judson IR. Biological rationale and current clinical experience with anti-insulin-like growth factor 1 receptor monoclonal antibodies in treating sarcoma: twenty years from the bench to the bedside. *The Cancer Journal*, 16(3):183–194, 2010.

[47] Juergens H, Daw NC, Geoerger B, Ferrari S, *et al.* Preliminary efficacy of the anti-insulin-like growth factor type 1 receptor antibody figitumumab in patients with refractory Ewing sarcoma. *Journal of Clinical Oncology*, 29(34):4534–4540, 2011.

[48] Pappo AS, Patel SR, Crowley J, Reinke DK, *et al.* R1507, a monoclonal antibody to the insulin-like growth factor 1 receptor, in patients with recurrent or refractory Ewing sarcoma family of tumors: results of a phase II Sarcoma Alliance for Research through Collaboration study. *Journal of Clinical Oncology*, 29(34):4541–4547, 2011.

[49] Tap WD, Demetri G, Barnette P, Desai J, *et al.* Phase II study of ganitumab, a fully human anti–type-1 insulin-like growth factor receptor antibody, in patients with metastatic ewing family tumors or desmoplastic small round cell tumors. *Journal of Clinical Oncology*, 30(15):1849–1856, 2012.

[50] Kolb EA, Kamara D, Zhang W, Lin J, *et al.* R1507, a fully human monoclonal antibody targeting IGF-1R, is effective alone and in combination with rapamycin in inhibiting growth of osteosarcoma xenografts. *Pediatric Blood & Cancer*, 55(1):67–75, 2010.

[51] Kolb EA, Gorlick R, Houghton PJ, Morton CL, *et al.* Initial testing (stage 1) of a monoclonal antibody (SCH 717454) against the IGF-1 receptor by the pediatric preclinical testing program. *Pediatric Blood & Cancer*, 50(6):1190–1197, 2008.

[52] Houghton PJ, Morton CL, Gorlick R, Kolb EA, *et al.* Initial testing of a monoclonal antibody (IMC-A12) against IGF-1R by the pediatric preclinical testing program. *Pediatric Blood & Cancer*, 54(7):921–926, 2010.

[53] Bagatell R, Herzog CE, Trippett TM, Grippo JF, *et al.* Pharmacokinetically guided phase 1 trial of the IGF-1 receptor antagonist RG1507 in children with recurrent or refractory solid tumors. *Clinical Cancer Research*, 17(3):611–619, 2011.

[54] Quek RH, Morgan JA, Shapiro G, Butrynski JE, *et al.* Combination mTOR+IGF-IR inhibition: phase I trial of everolimus and CP-751871 in patients (pts) with advanced sarcomas and other solid tumors. ASCO Annual Meeting 2010. Abstract 10002.

[55] Desai J, Solomon B, Davis ID, Lipton LR, *et al.* Phase I dose-escalation study of daily BMS-754807, an oral, dual IGF-1R/insulin receptor (IR) inhibitor in subjects with solid tumors. ASCO Annual Meeting 2010. Abstract 3104.

Chapter 6

Kinome and mRNA expression profiling of osteosarcoma identifies genomic instability, and reveals Akt as potential target for treatment

This chapter is based on the manuscript: Kuijjer ML, van den Akker BEWM, Hilhorst R, Mommersteeg M, Buddingh EP, Serra M, Bürger H, Hogendoorn PCW, Cleton-Jansen AM. *Submitted*

Abstract

Background: High-grade osteosarcoma is a primary malignant bone tumor mostly occurring in adolescents and young adults, with a second peak at middle age. Overall survival is approximately 60%, and has not significantly increased since the introduction of neoadjuvant chemotherapy in the 1970s. The genomic profile of high-grade osteosarcoma is complex and heterogeneous. Integration of different types of genome-wide data may be advantageous in extracting relevant information from the large number of aberrations detected in this tumor.

Methods: We analyzed genome-wide gene expression data of osteosarcoma cell lines, and integrated these data with a kinome screen. Data were analyzed in statistical language R, using *LIMMA* for detection of differential expression/phosphorylation. We subsequently used Ingenuity Pathways Analysis to determine deregulated pathways in both data types.

Results: Gene set enrichment indicated that pathways important in genomic stability are highly deregulated in these tumors, with many genes showing upregulation, which could be used as a prognostic marker, and with kinases phosphorylating peptides in these pathways. Akt and AMPK were identified as active and inactive, respectively. As these pathways have an opposite role on mTORC1 signaling, we set out to inhibit Akt with the allosteric Akt inhibitor MK-2206. This resulted in inhibition of proliferation of osteosarcoma cell lines U-2 OS and HOS, but not of 143B, which harbors a *KRAS* oncogenic transformation.

Conclusions: We identified both overexpression and hyperphosphorylation in pathways playing a role in genomic stability. Kinome profiling identified active Akt signaling, which could inhibit proliferation in 2/3 osteosarcoma cell lines. This study provides a rationale for further testing inhibitors of the PI3K/Akt/mTORC1 pathway in preclinical studies of osteosarcoma.

Background

High-grade osteosarcoma is the most prevalent primary malignant bone tumor. Most frequently, the long bones of adolescents and young adults are affected, with a yearly incidence of approximately 5 cases per million per year (1). Patients are generally treated with high doses of neoadjuvant chemotherapy to prevent the outgrowth of micrometastases. In 15–25% of all patients, however, metastatic disease is clinically detectable at diagnosis and despite the intensive treatment, 45% of all patients develop distant metastases, the leading cause of death of osteosarcoma patients (2, 3). The introduction of neoadjuvant chemotherapy in the 1970s has increased survival from 10–20% to approximately 60%. However, survival has reached a plateau, and new treatments are urgently needed (4–6). Osteosarcoma is an extremely genomically unstable tumor, with karyotypes

harboring numerous numerical and structural changes (7, 8). In addition, osteosarcoma genotypes show a considerable degree of heterogeneity, both intra- and intertumoral. Both the complex genotype and its heterogeneity render it difficult to determine which genomic alterations are important in osteosarcomagenesis, as not all alterations may lead to a difference in mRNA, protein levels, or enzyme activity in the tumor tissue. Integration of different data types is therefore of particular relevance for studying a heterogeneous tumor with a complex genomic profile such as osteosarcoma. Previously, genomic and expression data of osteosarcoma pretreatment biopsies have been integrated, in order to detect highly recurrent osteosarcoma driver genes. The list of driver genes obtained with this study was enriched in genes playing a role in genomic stability (9). Yet, even though recurrent driver genes may provide knowledge on what pathways are affected that help tumor cells survive, such driver genes may not always be accessible as targets for treatment. This especially holds for pathways involved in genetic stability, since the damage is already done.

Oncogenic kinases are often active in tumor cells, and a number of kinases can be pharmacologically inhibited. Therapies targeting oncogenic kinases have provided promising results in inhibiting proliferation of cancer cells, and some kinases have been targeted in preclinical and clinical studies in childhood sarcomas (as reviewed in Wachtel *et al.* (10)), *e.g.* IGF1R and mTOR (11, 12). An unbiased approach to identify active kinases in cancer is to perform kinome-wide screens. Such screens have previously been effectively used in other types of sarcoma and have led to the detection of specific targets for treatment (13, 14). As combining the analysis of different data types using systems biology approaches can give a more complete impression of the state of a tumor cell, we set out to integrate genome-wide gene expression data of osteosarcoma cell lines with kinome profiling data. Osteosarcoma cell lines are widely available and have been shown to be representative for the tumor of origin, both on a genome-wide as on a functional level, and are therefore a good model to study osteosarcoma preclinically (15, 16). We previously have performed genome-wide expression analysis on a panel of 19 osteosarcoma cell lines (17). In the present study, we compared expression profiles with the different putative progenitor cells of osteosarcoma—mesenchymal stem cells (MSCs) and osteoblasts—in order to define the common denominator pathways that are deregulated in osteosarcoma. Pathways with a role in genomic stability appeared to be enriched in overexpressed genes. By integrating expression data with a serine/threonine (Ser/Thr) kinome screen, we show that these pathways are enriched in hyperphosphorylation as well, confirming that genomic stability is highly deregulated in osteosarcoma, both on a transcriptional level and on phosphorylation activity.

In order to detect overactive kinases in osteosarcoma, which may be potential targets for treatment, we identified the most significant pathways in the kinome profiling data, which indicated active PI3K/Akt and inactive AMPK signaling. These pathways play an opposite role in mTORC1 signaling, with Akt promoting and AMPK inhibiting signal

transduction (18). We pharmacologically inhibited Akt in osteosarcoma cell lines, which reduced proliferation of 2/3 cell lines. In summary, this study describes integration of mRNA and phosphorylation data, and gives a rationale for treatment of osteosarcoma with inhibitors of the PI3K/Akt pathway.

Methods

Cell culture

Osteosarcoma cell lines were previously characterized and described (17). Human bone-marrow-derived MSCs were obtained from two osteosarcoma patients, and were characterized and handled as described (19). For kinome profiling of osteosarcoma versus MSCs, cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Invitrogen, Carlsbad, CA), supplemented with 10% fetal bovine serum (Greiner Bio-one, Frickenhausen, Germany), in order to eliminate differences in kinase activity caused by culture conditions. For inhibition experiments and kinome profiling of inhibition experiments, osteosarcoma cell lines 143B, U-2 OS, and HOS were maintained in RPMI 1640 supplemented with 10% fetal calf serum (both from Invitrogen, Carlsbad, CA). The human pre-B acute lymphoblastic leukemia cell line NALM-6 cell line was kindly provided by Mw. N. Duinkerken (Department of Hematology, Leiden University Medical Center, the Netherlands), and was maintained in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with GlutaMAX-1 (Life Technologies, Carlsbad, CA) and 10% fetal bovine serum (Greiner Bio-one, Frickenhausen, Germany). All cells were regularly tested for mycoplasm and were genotyped using the Powerplex 1.2 system (Promega, Leiden, the Netherlands), as described previously (16).

Cell lysates

Kinome profiling was performed on osteosarcoma cell lines 143B and U-2 OS and on two MSCs—MSC001 and MSC006. Cells at 80% confluence were washed twice with Phosphate buffered Saline and lysed with M-PER Mammalian Extraction Buffer, supplemented with Halt Phosphatase Inhibitor Cocktail and EDTA free Halt Protease Inhibitor Cocktail (Pierce Biotechnology, Rockford, IL), according to the manufacture's protocol. Cells were incubated on ice for at least 30 minutes before collecting the lysates and centrifuging these for 15 minutes at 4°C at $> 10,000 \cdot g$. Protein concentration was measured using a detergent-compatible Protein Assay (Bio-Rad Laboratories, Hercules, CA) according to the manufacturer's protocol. Samples were snap-frozen and stored at -70°C.

Proliferation assays

MK-2206 was dissolved in DMSO at a concentration of 10mM and stored at -20°C. For 143B, U-2 OS, and HOS, 2,000, 4,000, and 2,000 cells/well respectively, were plated in a 96-wells plate. NALM-6, a human pre-B acute lymphoblastic leukemia (ALL) cell line, was included as a positive control, as ALL cell lines have been shown to be highly sensitive to MK-2206 (20). This cell line grows in suspension and was plated at 50,000 cells/well. After 24hrs, MK-2206 was added in triplicate in different concentrations—0nM, 0.5nM, 1nM, 5nM, 10nM, 50nM, 100nM, 500nM, 1μM, 5μM, and 10μM. For 143B and HOS, the effect of concentrations of 2, 3, 4, and 5nM was assessed as well. Cells were grown in the presence of inhibitor for 120hrs. Cell proliferation was determined by incubating the cells with reagent WST-1 (Roche, Basel, Switzerland) for 2hrs and subsequently measured using a Wallac 1420 VICTOR2 (Perkin Elmer, Waltham, MA). Data were analyzed in Graphpad Prism 5.01 (`www.graphpad.com`). Relative IC_{50}s were calculated using results from the different concentrations up to the highest dose where toxicity was not yet present. The results shown are representative results from at least three independent experiments.

Genome-wide gene expression profiling

We analyzed our previously published data of osteosarcoma cell lines ($n = 19$), MSCs ($n = 12$), and osteoblasts ($n = 3$) (GEO superseries, accession number GSE42352) (9). Microarray data processing and quality control were performed in the statistical language R version 2.15 (21) as described previously (22).

Kinome profiling

Kinome profiling was performed on 1μg of cell lysate on the serine/threonine (Ser/Thr) Kinase PamChip®peptide microarrays (PamGene, 's-Hertogenbosch, the Netherlands) according to the manufacturer's protocol, essentially as described in Hilhorst *et al.* (23). This peptide microarray comprises 142 peptide sequences derived from human phosphorylation sites. Peptide phosphorylation is detected in time with a mixture of fluorescently labeled antiphosphoserine/threonine antibodies. We used at least three technical replicates for each MSC line, and four technical replicates for the osteosarcoma cell lines. Images were taken every 5 minutes, over the course of 60 minutes. Signal quantification on phosphorylated peptides was performed in BioNavigator software (PamGene International, 's Hertogenbosch, the Netherlands). Subsequently, data were normalized in R (24) using the *vsn* package (25). Median signals at 60 minutes of incubation with the cell lysates were analyzed in Bioconductor (26) package *arrayQualityMetrics* (27) to identify poor quality samples, which were removed from further analysis. Technical replicates of good quality were averaged. To determine whether these data were reproducible, we analyzed data from different cycles (0, 10, 20, 30, 40, 50, and 60 minutes incubation with

cell lysates).

In the second kinome profiling experiment we compared lysates of untreated cells with lysates of cells treated with MK-2206. Different treatment durations and concentrations were used—no treatment, treatment for 5, 30, 180, and 960 minutes with 1μM MK-2206, and treatment for 180 minutes with 10μM of the drug. Kinome profiling was performed as described above, with the difference that we used 1–5 technical replicates per condition. Of this experiment, we analyzed signals at 30 minutes of incubation with the lysates.

Statistical analyses of microarray data

We performed *LIMMA* analysis (24) in order to determine differential mRNA expression between osteosarcoma cell lines ($n = 19$) and control cell lines—MSCs ($n = 12$) and osteoblasts ($n = 3$) and to determine differential phosphorylation of peptides on the PamChip®microarray between osteosarcoma cell lines ($n = 2$) and MSCs ($n = 2$). We used a Benjamini and Hochberg False Discovery Rate (FDR) of 0.05 as cut-off for significance. Kinome profiling signals obtained for the different treatment conditions were analyzed in a paired approach, in which signals from untreated cells were subtracted from the signals from treated cells. For both kinome profiling experiments, we used a cut-off of 0.1 for the absolute log fold change (logFC). Heatmaps were generated using the function heatmap.2 of R package gplots.

Pathway analysis

In order to reveal pathways which were significantly affected on mRNA levels in osteosarcoma cell lines, we intersected the toptables obtained by *LIMMA* analysis of osteosarcoma cell lines versus MSCs and of osteosarcoma cell lines versus osteoblasts. Gene symbols for all probes were imported into the software Ingenuity Pathways Analysis (IPA, Ingenuity Systems, `www.ingenuity.com`), together with FDR adjusted p-values (adjP) and average logFCs. Only the gene symbols of probes that were both significantly upregulated or both significantly downregulated in osteosarcoma cell lines as compared with MSCs and with OBs (adjP < 0.05) were selected to be considered as significantly differentially expressed in the IPA analysis. For differential phosphorylation, we imported the results from the *LIMMA* analysis on kinome profiling data, with a cut-off of 0.05 for adjusted p-value and a cut-off of 0.1 for logFC. The significance of the association between the data set and the canonical pathways was measured as described previously (28). Pathways with adjP < 0.05 were considered to be significantly affected. In addition, transcription factor analyses were performed on gene expression data in IPA in order to predict activated or inhibited transcription factors based on expression of target genes, returning p-values (with a cut-off of 0.05 for significance) and regulation z-scores.

Results

Genome-wide gene expression profiling of high-grade osteosarcoma cell lines

We started by comparing gene expression signatures of 19 osteosarcoma cell lines, 12 MSC, and 3 osteoblast cultures using unsupervised hierarchical clustering. Two separate clusters were detected—one containing all tumor cell samples and one containing control samples. Within the control sample cluster, osteoblasts clustered separately from MSCs (*data not shown*). *LIMMA* analysis resulted in 7,891 probes encoding for differentially expressed (DE) genes between osteosarcoma cell lines and MSCs, and 2,222 probes encoding for DE genes between osteosarcoma cells and osteoblasts. Intersecting of these gene lists showed 1,410 probes that were significant in both analyses, of which 1,390 were upregulated in both analyses, or downregulated in both analyses (Figure 6.1). These probes, encoding for 1,312 genes, were selected for subsequent pathways analysis, in order to determine commonly affected pathways in osteosarcoma tumor cells.

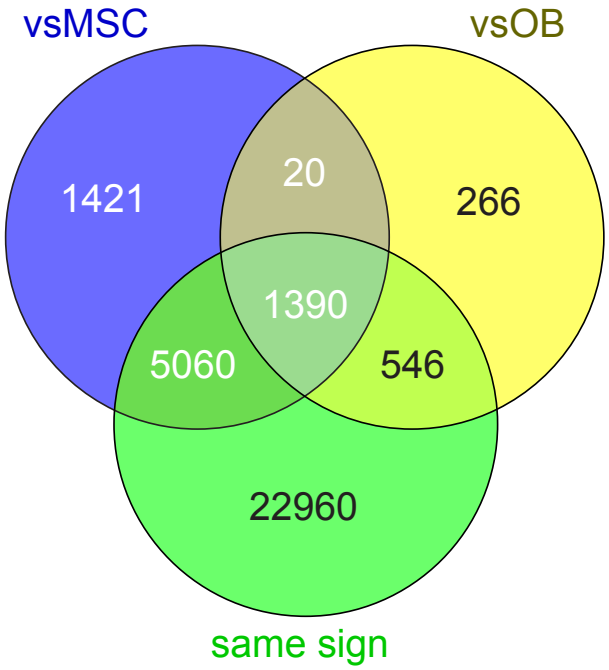


Figure 6.1: Venn diagram showing the significant probes in the analysis of osteosarcoma cell lines *vs* MSC (vsMSC) and *vs* osteoblasts (vsOB), and the intersection of these significant probes with the subset of all probes (both significant and nonsignificant) that shows both up- or both downregulation in these two analyses (same sign). In total, 1,410 probes are significant in both analyses, of which 1,390 have the same sign of logFC.

Gene expression is altered in pathways regulating genomic stability

Pathway analyses on the 1,312 differentially expressed genes resulted in 17 significantly affected pathways (Figure 6.2). Fourteen out of these 17 pathways play a direct or indirect role in genomic stability. Unsupervised hierarchical clustering of all cell line data and data from 84 osteosarcoma biopsies (GEO accession number GSE33382 (9)) was performed on all DE genes present in these 17 significantly affected pathways, which resulted in a cluster of control cells and biopsies, and larger cluster of osteosarcoma cell lines and biopsies

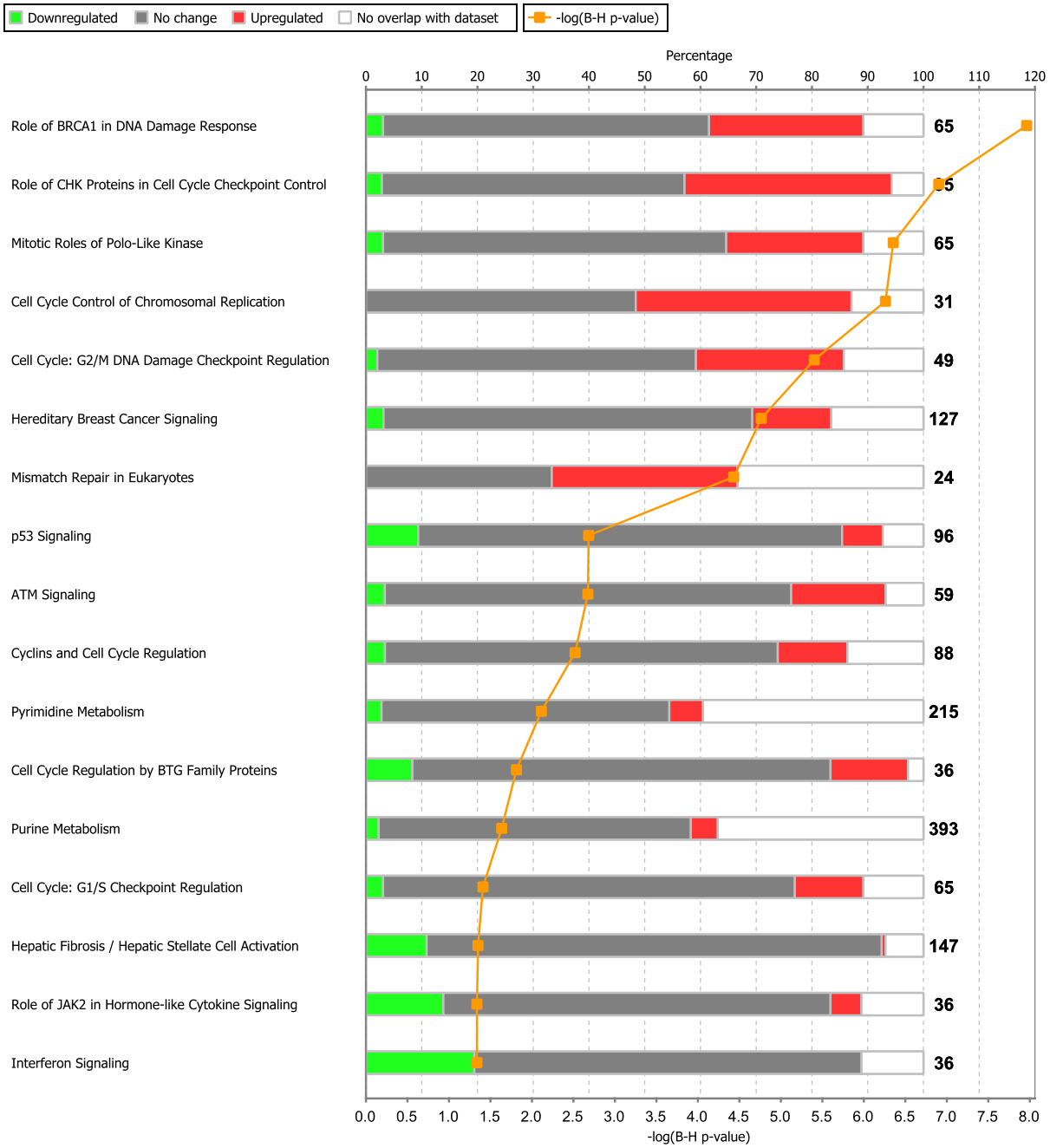


Figure 6.2: Stacked bar chart depicting all significantly affected pathways as identified by gene expression profiling of osteosarcoma cell lines, showing percentages of downregulated (green), not significantly altered (gray), and upregulated (red) genes, and genes which were not present on the microarray (white). The -log(adjP) (-log(B-H) p-value) is plotted in orange, and is above 1.3 for adjP < 0.05.

(Additional Figure 6.1). Patients whose biopsies had expression profiles of these pathways similar to osteosarcoma cell lines showed worse metastasis-free survival than patients with intermediate expression profiles, and than patients whose biopsies had expression profiles more similar to the control cultures, *i.e.* nontransformed primary mesenchymal cell cultures and osteoblast cultures (Logrank test for trend, p-value = 0.049, Figure 6.3). Transcription factors were predicted to be activated or inhibited based on expression of

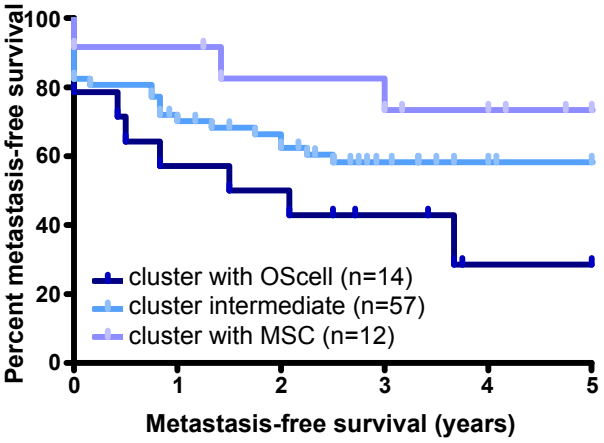


Figure 6.3: Kaplan-Meier metastasis-free survival analysis on data obtained from patient biopsies which clustered with osteosarcoma cell lines, biopsies clustering with control cell lines, and an intermediate group, based on gene expression of genes all present in the 17 significantly affected pathways (as in Additional Figure 6.1). Logrank test for trend, p-value = 0.049.

target genes are shown in IPA. The most activated transcription factor was *MYC*, while the most inactivated transcription factor was *TP53*.

Kinome profiling of osteosarcoma cell lines

To obtain more information on the activity of the pathways which showed aberrant mRNA expression, we integrated mRNA expression data with data obtained with kinase PamChip®peptide microarrays. These peptide microarrays were incubated with lysates of the osteosarcoma cell lines 143B and U-2 OS, and with lysates of two MSC cultures. Kinases present in the cell lysates can, in the presence of ATP, phosphorylate the peptides present on the microarray, which is detected by fluorescently labeled antibodies. We compared kinome profiling data at different incubation times by intersecting lists of differentially phosphorylated peptides between osteosarcoma cells and MSCs, obtained by *LIMMA* analyses, as shown in Figure 6.4. This data analysis demonstrated a large overlap in the detected differentially phosphorylated peptides, and a build-up of differentially phosphorylated peptides over time. Most peptides showed differential phosphorylation after 20 minutes of incubation with cell lysates. After 60 minutes of incubation on the peptide microarray, 49 peptides were detected to be significantly differentially phosphorylated between osteosarcoma cell lines and mesenchymal stem cells. These peptides are represented in Figure 6.5. As a reference, we performed an unsupervised hierarchical clustering including all technical replicates (*data not shown*), which showed that phosphorylation of peptides by cell lysates of most technical replicates was comparable.

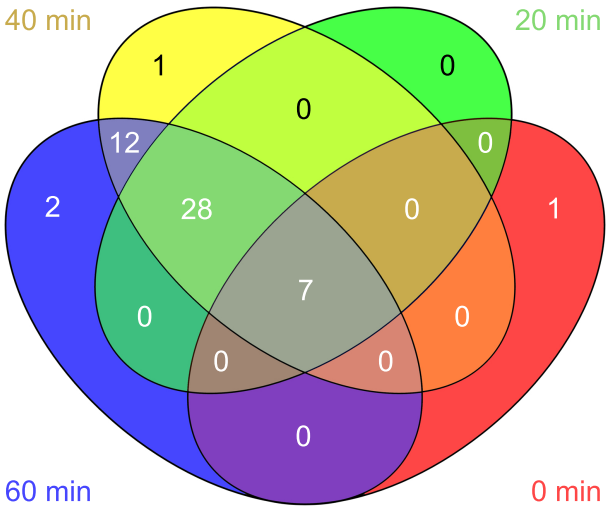


Figure 6.4: Comparison of peptide phosphorylation at different time points. *LIMMA* analyses were performed on different time points, ranging from 0 to 60 minutes of incubation with cell lysates. Venn diagrams show overlap of significantly differentially phosphorylated peptides between the consecutive time points.

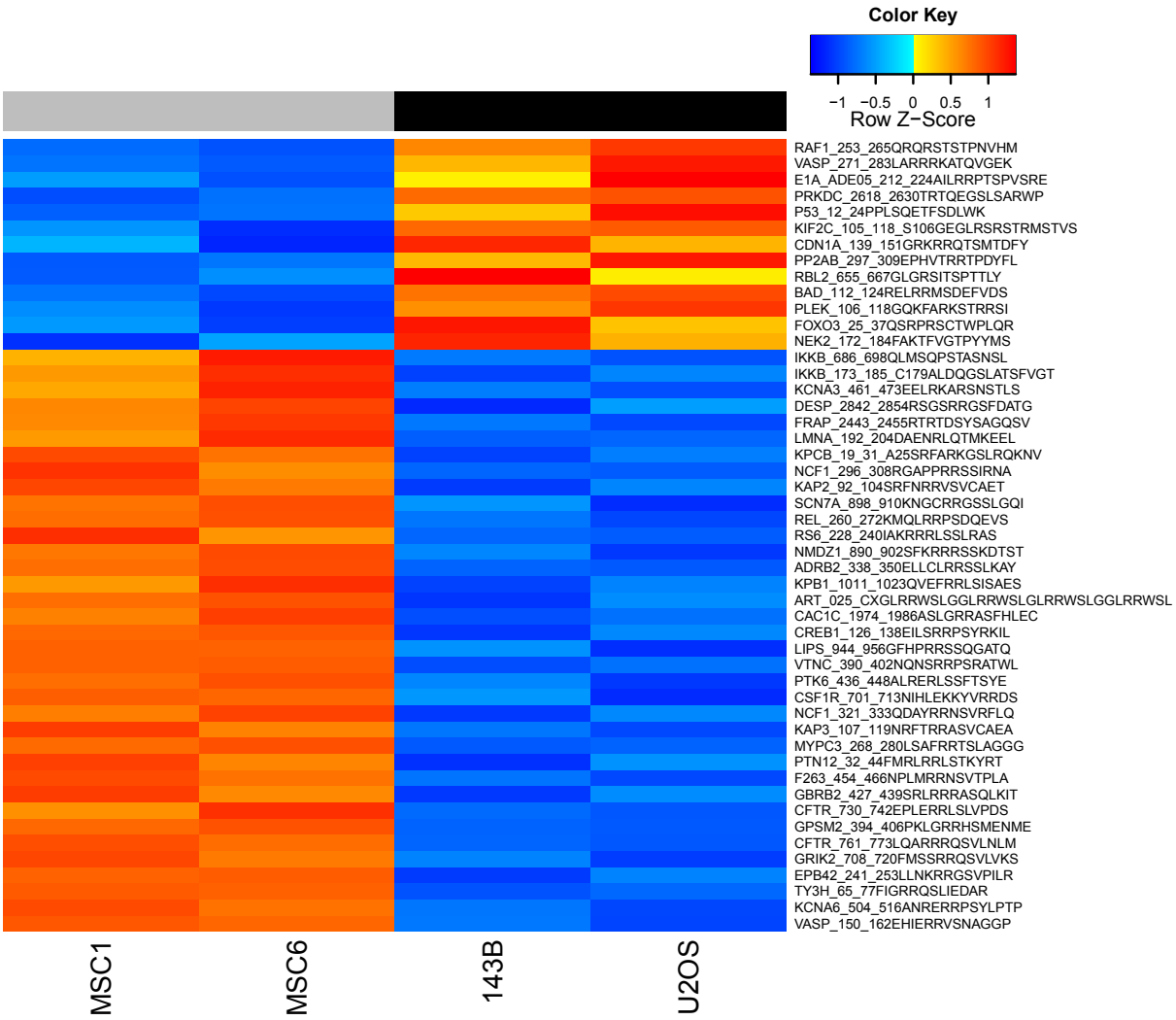


Figure 6.5: Supervised clustering of all 49 significantly differentially phosphorylated peptides identified by the comparison of two osteosarcoma cell lines with two MSC cultures. Peptides are sorted on logFC, from lower phosphorylation to higher phosphorylation in osteosarcoma cell lines. Orange: higher phosphorylation levels, blue: lower phosphorylation levels.

Altered phosphorylation in genomic stability pathways

The significance of the 17 pathways that were returned from the pathway analysis on mRNA expression data was tested on kinome profiling results in IPA. In total, 7/17 pathways were significant in kinome profiling as well. These seven pathways were a subset of the 14 pathways with a known role in genomic stability. Most significantly differentially phosphorylated peptides in these seven pathways showed higher phosphorylation levels in osteosarcoma cell lines (Figure 6.6), indicating that kinases affect phosphorylation of molecules playing a role in genomic stability.

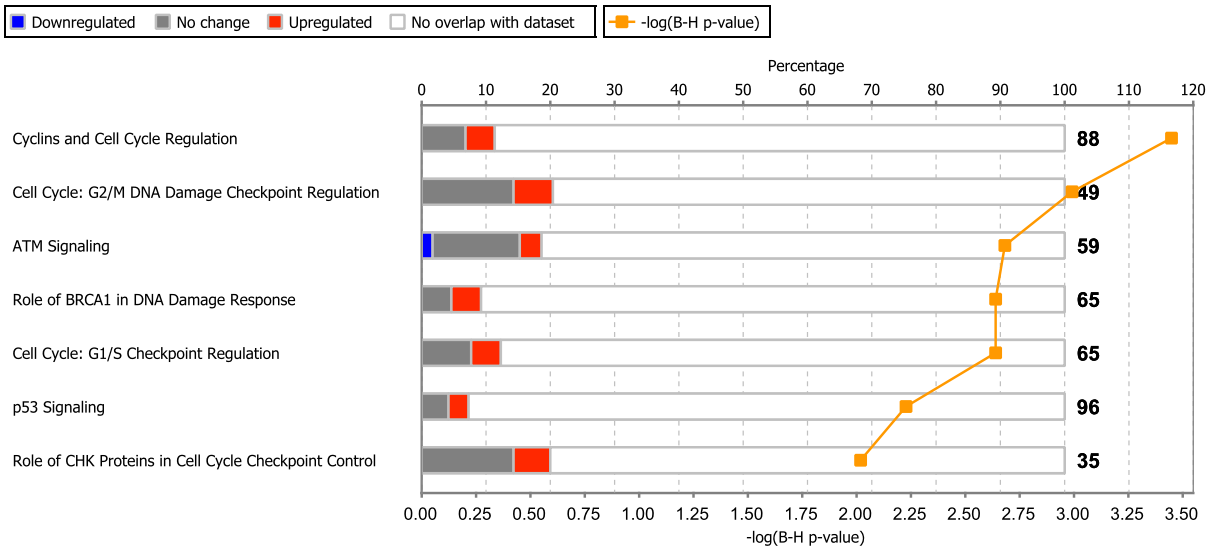


Figure 6.6: Stacked bar chart showing kinome profiling pathway analysis on the subset of pathways which were significant on gene expression profiling. Percentages of downregulated (blue), not significantly altered (gray), and upregulated (orange) genes, and genes which were not present on the microarray (white) are shown. The -log(adjP) (-log(B-H) p-value) is plotted in orange, and is above 1.3 for adjP < 0.05.

PI3K/Akt and AMPK signaling in osteosarcoma

Unsupervised pathway analysis on the kinome profiling results returned the IPA pathway PI3K/Akt signaling as the most significantly affected pathway in osteosarcoma cells (Figure 6.7) and the AMPK pathway as second most significantly affected pathway. Specifically, molecules directly downstream of Akt showed higher phosphorylation in osteosarcoma than in MSCs, while molecules downstream of AMPK showed lower phosphorylation levels. As these results indicate that Akt signaling is active in osteosarcoma and might be driving its high proliferative capacity, we set out to pharmacologically inhibit Akt using the compound MK-2206.

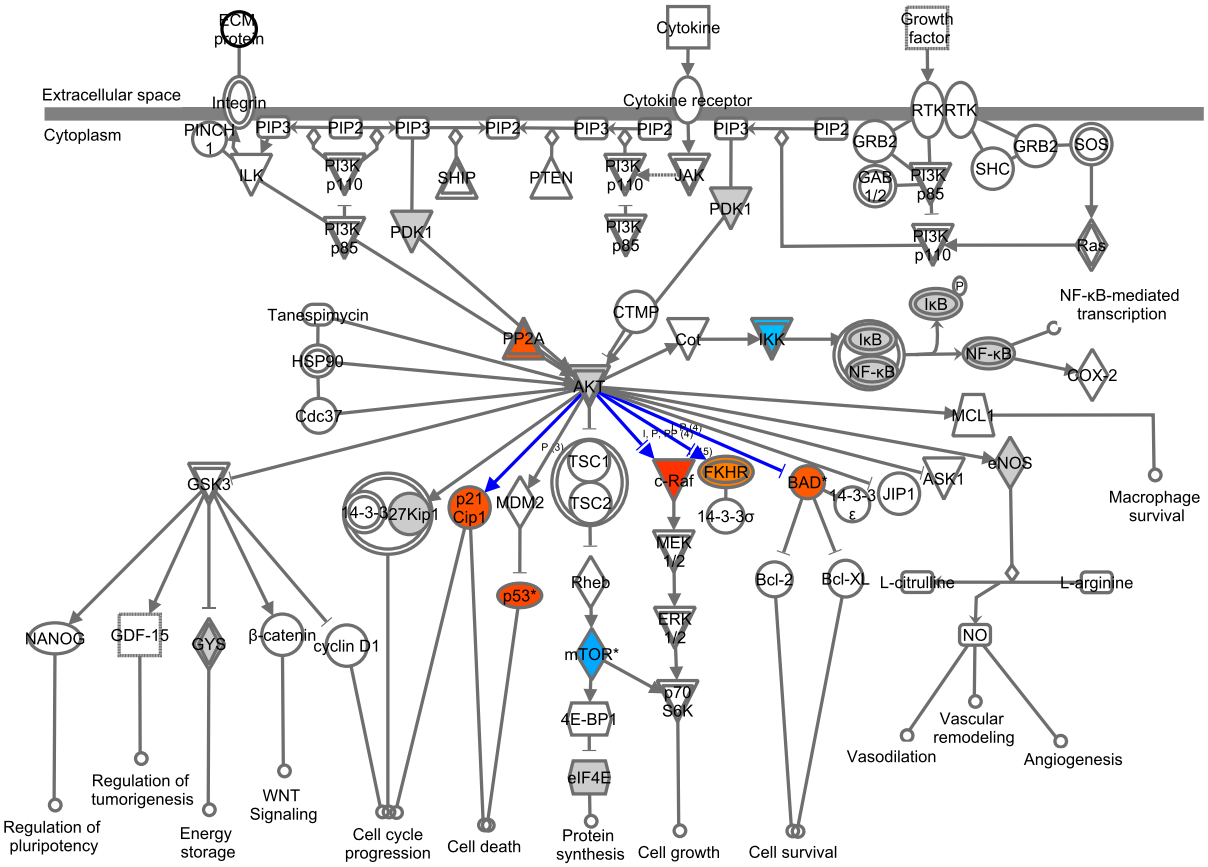


Figure 6.7: The Akt signaling pathway in IPA. Blue: significantly lower, orange: significantly higher phosphorylation in osteosarcoma cell lines, gray, no significant difference in phosphorylation, white: no phosphorylation sites of the particular protein on the PamGene Ser/Thr chip. Blue lines indicate known downstream phosphorylation by the upstream kinase.

MK-2206 inhibits proliferation of U-2 OS and HOS, but not of 143B

We inhibited osteosarcoma and control cells for 120hrs with allosteric inhibitor MK-2206. Inhibition of the positive control leukemia cell line NALM-6, and of osteosarcoma cell line U-2 OS with MK-2206 was dose-dependent, with IC_{50}s of 0.38μM and 2.5μM, and maximal responses of 94% and 71%, respectively (Figure 6.8). 143B did not show any

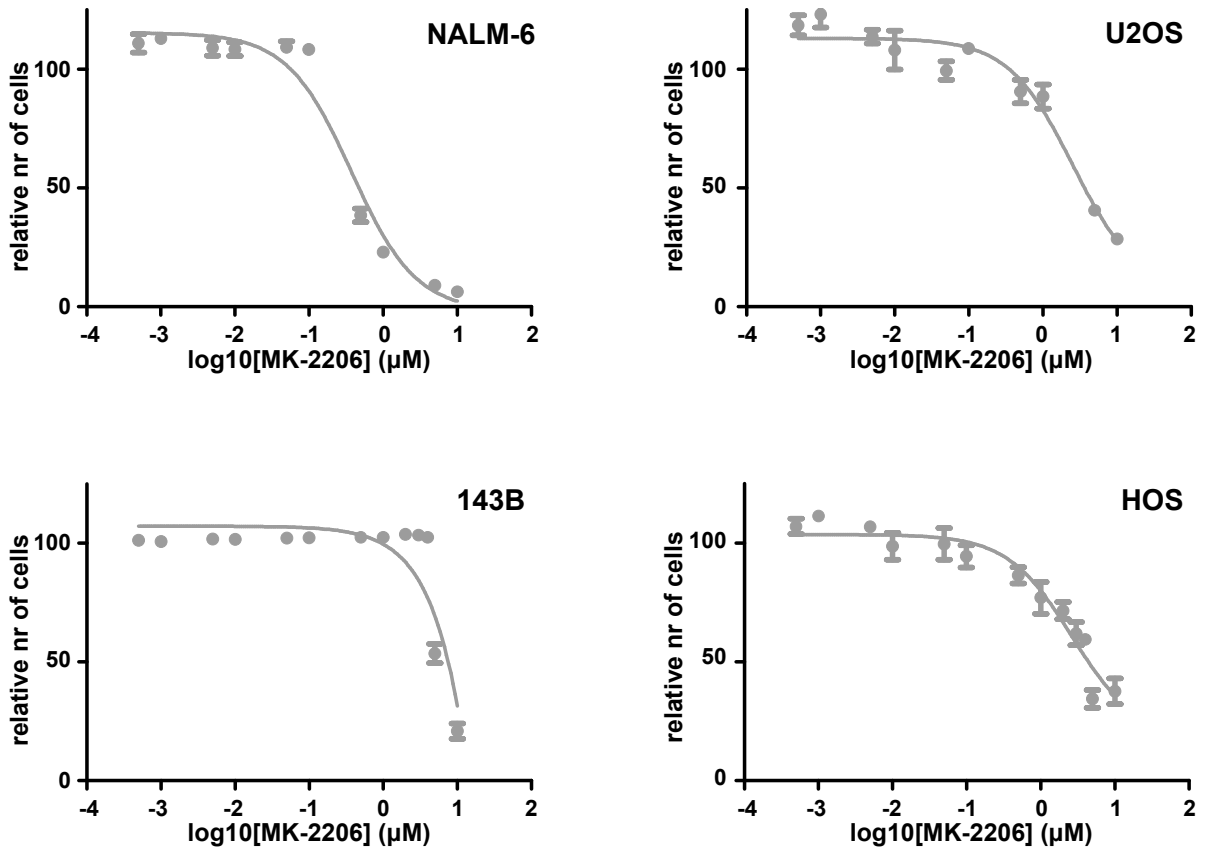


Figure 6.8: Proliferation of osteosarcoma cell lines was inhibited with different concentrations of MK-2206, for 120hrs. NALM-6, U-2 OS, and HOS showed a dose-dependent inhibition, while 143B did not respond.

response at concentrations below 5μM. Because 143B exhibits an oncogenic *KRAS* transformation, we assessed MK-2206 specificity on the parental cell line of 143B, HOS, which does not exhibit this transformation. HOS indeed responded similar to U-2 OS, with an IC_{50} of 2.6μM and maximal response of 62%.

Different phosphorylation patterns upon treatment with MK-2206

As 143B and U-2 OS showed different sensitivities to MK-2206, we performed a paired analysis between kinome profiling data obtained from lysates of cells, which were treated with different concentrations of MK-2206, and for different treatment lengths. Overall, the phosphorylation patterns differed between both cell lines, and distances between treat-

ment options within each cell line were smaller than between the cell lines (Figure 6.9). We generated a heatmap of differential phosphorylation in the paired analysis of treated

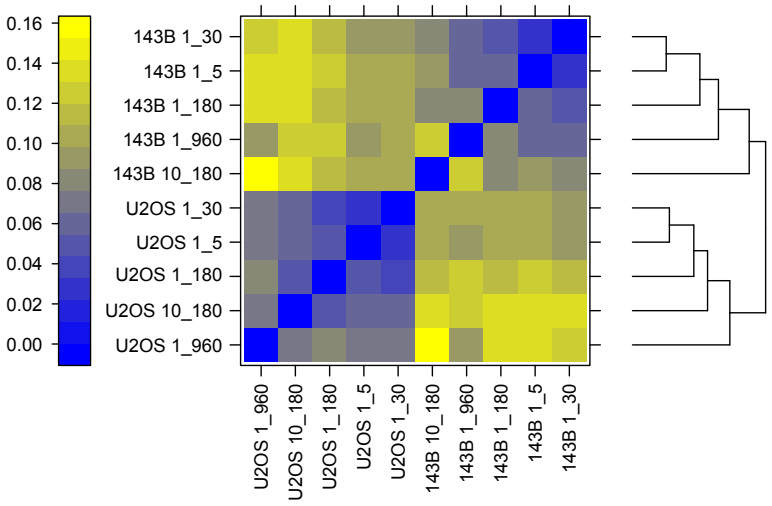


Figure 6.9: Unsupervised hierarchical clustering depicting the distances between data obtained from kinome profiling of cells treated with different concentrations of MK-2206 and for different time intervals. 1_30: treatment of 30min with 1µM of MK-2206, etc.

and untreated cells, depicting all peptides of the PamGene chip which are downstream of PI3K/Akt (Figure 6.10). This figure shows that the inhibition pattern of MK-2206 is different in the two osteosarcoma cell lines, suggesting that other upstream kinases may be affected by inhibition of Akt with MK-2206 as well.

Discussion

Osteosarcoma is a highly genomically unstable tumor. The identification of specific molecular targets that drive oncogenesis and that might be targets for therapy may thereby be hampered. Genome-wide gene expression profiling of high-grade osteosarcoma cell lines, in fact, showed an enrichment in differential expression in pathways important in genomic stability (Figure 6.2), with a role in cell cycle and checkpoint regulation (*e.g.* p53 signaling, G1/S and G2/M checkpoint regulation), DNA damage response (*e.g.* ATM signaling, role of BRCA1 in DNA damage response), and purine/pyrimidine metabolism. Most significantly differentially expressed genes in these pathways were upregulated, for example *DNA-PK*, *BRCA1*, and *CDC25A*. Some downregulated genes were detected as well, such as *CDKN1A*, which has an inhibitory role on cell cycle progression, and genes downstream of *TP53* (*e.g.* *THBS1* and *SERPINE1*, encoding TSP1 and PAI-1, respectively). Interestingly, as shown by unsupervised clustering on expression levels of genes of these pathways, osteosarcoma pretreatment biopsies can have profiles more similar to those of osteosarcoma cell lines, or more similar to profiles of the control cells. The first is associated with poor, while the latter is associated with good metastasis-free survival. Expression profiles can also be of an intermediate type, with intermediate metastasis-free survival (Additional Figure 6.1, Figure 6.3). This suggests that deregulated genomic stability is a key driver of osteosarcomagenesis, as was already previously reported (9). IPA

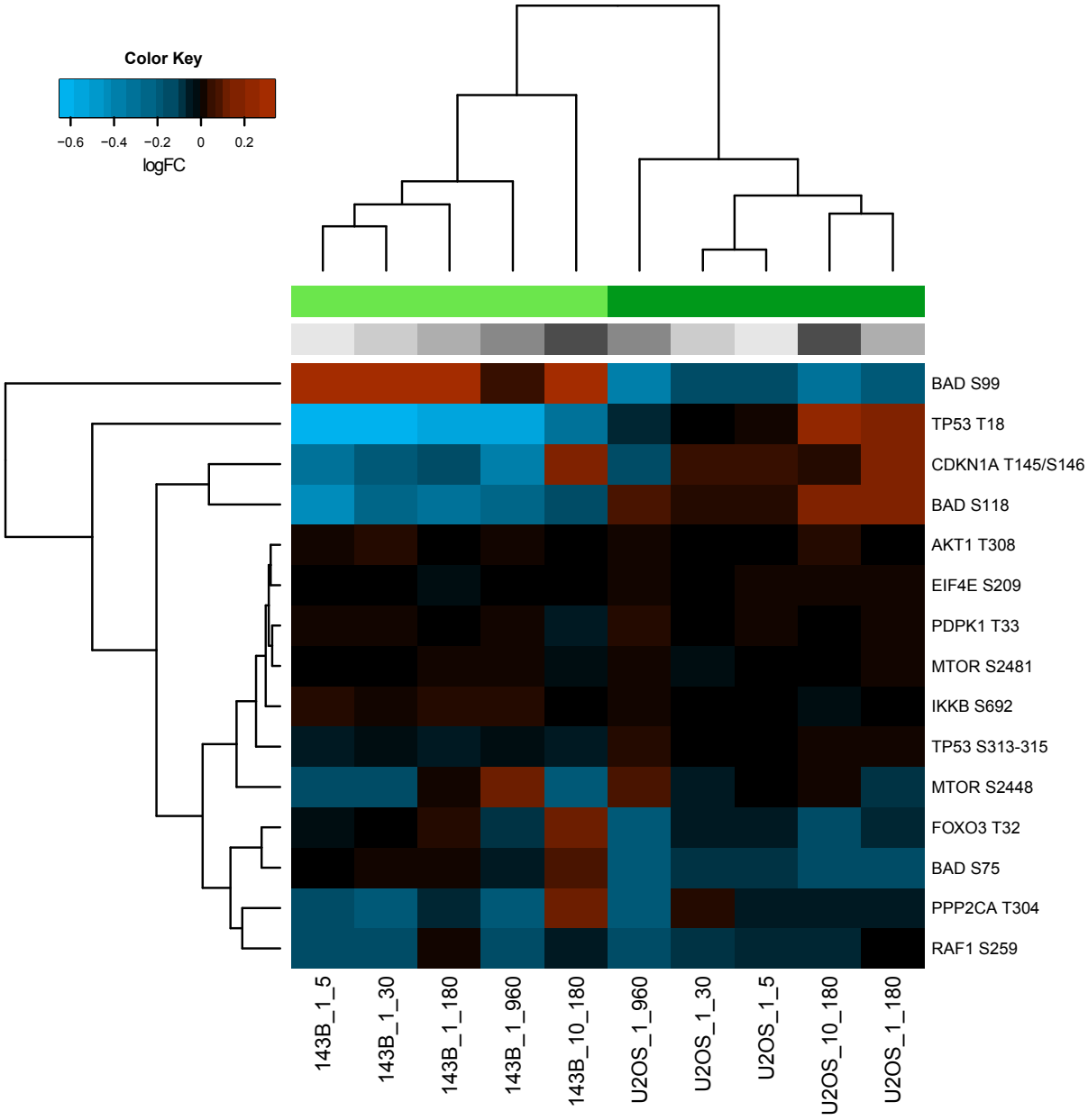


Figure 6.10: Unsupervised clustering depicting differential phosphorylation of peptides of the PI3K/Akt pathway by cell lysates treated with different concentrations of MK-2206 for different time intervals, as compared with untreated cells. Blue: logFC < 0, orange: logFC > 0. Different treatment options are shown in different shades of gray (from light to dark gray: 1µM 5, 30, 180, and 960 minutes, and 10µM 180 minutes of treatment with MK-2206. Light green: 143B, dark green: U-2 OS.

transcription factor analysis showed that *MYC* is the most significantly activated (z-score of 6.294), and *TP53* the most significantly inactivated (z-score of -7.660) transcription factor. Other highly predicted activated transcription factors are *e.g. E2F1/2/3*, whereas *CDKN2A* and *RB1* were detected as inactivated. These different genes are known to be affected in osteosarcoma (7, 9, 29). The role of these genes in cell cycle progression further confirms the importance of these pathways in osteosarcoma.

As kinome-wide screens have previously led to the detection of specific targets for treatment in other sarcoma types (13, 14), we performed kinome profiling of osteosarcoma cell lysates. Since the pathways that were shown to be significantly affected on mRNA expression mostly contained Ser/Thr kinases, we selected a Ser/Thr peptide microarray—the Ser/Thr PamChip®. Pathway analysis on kinome profiling data showed that 50% of the pathways that were significant on gene expression data were also significantly enriched in differential phosphorylation signals (Figure 6.6). All significant peptides were higher phosphorylated in osteosarcoma cell lines, except for a peptide present in the gene *CREB1*. Since most of these peptides showed higher phosphorylation, we expect these pathways to be highly active, demonstrating higher cell cycling of the tumor cells, and deregulated responses to DNA damage.

We next determined the most significantly affected pathways in the kinome data from the entire IPA canonical pathways database, and detected deregulation of the PI3K/Akt and AMPK signaling pathways. Molecules downstream of Akt showed higher phosphorylation (Figure 6.7), while downstream of AMPK, lower levels of phosphorylation were detected. Akt and AMPK act antagonistically to regulate mTOR signaling through inhibitory and activating phosphorylation of TSC2, respectively (18). The Akt pathway is one of the most commonly affected pathways in cancer, with active PI3K/Akt signaling leading to excessive cell growth and proliferation (30, 31). Inhibition of this pathway by targeting mTOR with agents such as rapamycin is effective in some cancer types (32). In a recent phase II trial in bone and soft tissue sarcomas, inhibition of mTOR with ridaforolimus resulted in better progression-free survival (12). Inhibiting mTOR can, however, also activate a strong negative feedback loop from S6K1 to enhance Akt signaling (30, 32). It may, therefore, be more effective to inhibit Akt itself. Inhibition of Akt was recently tested in a panel of xenografts of different pediatric cancers, and was most effective in osteosarcoma, with significant differences in event-free survival in 6/6 xenografts (20). In addition, AMPK activators suppress growth of cell lines of various tumor types (33).

We treated osteosarcoma cell lines with the allosteric Akt inhibitor MK-2206 (Selleck Chemicals LLC, Houston, TX). Inhibition of proliferation was dose-dependent in U-2 OS (IC_{50} = 2.5μM), but not in 143B (Figure 6.8). Important to note is that active Akt signaling can be detected by kinome profiling in this cell line, but this does not necessarily imply that this pathway can also be fully inhibited, for example in the case that downstream actors in the same pathway cause a survival benefit for the cell line. As 143B

is derived from the HOS cell line with a *KRAS* oncogenic transformation, we determined inhibitory effects of MK-2206 on HOS as well. HOS responded to MK-2206 in a similar manner as U-2 OS ($IC_{50} = 2.6\mu M$). This suggests that constitutive Ras/Raf/ERK signaling causes insensitivity to inhibition of the Akt pathway to MK-2206. Kinome profiling of cells treated with MK-2206 resulted in different phosphorylation patterns in 143B and U-2 OS of peptides of molecules in the PI3K/Akt pathway (Figure 6.10). Differences between these cell lines were found in BAD Ser-99, of which phosphorylation was inhibited after treatment with MK-2206 in the responsive cell line U-2 OS, but stimulated in 143B, and in BAD Ser-118, where an opposite pattern was detected. BAD Ser-99 is the major site of Akt phosphorylation, while Ser-118 is the major site of PKA phosphorylation (34). Opposite patterns were also detected for TP53 Thr-18 and CDKN1A Thr-145/Ser-146, of which CDKN1A Thr-145 can also be directly phosphorylated by Akt. These results suggest that activity of other kinases may be affected by inhibition of Akt using MK-2206, or by MK-2206 itself. This depends on the cellular context, as we otherwise would not have expected to detect any differences in a paired analysis for the different conditions in each cell type.

An important finding of our studies is that the PI3K/Akt and AMPK signaling pathways were detected with kinome profiling, while mRNA expression profiling did not result in the identification of these pathways. This suggests that in osteosarcoma, these pathways are regulated by phosphorylation rather than by transcriptional activity. Gene expression and protein synthesis imply a long time commitment of a cell to potential activation of its synthesized proteins. Phosphorylation, on the other hand, provides a very rapid way to mobilize extra catalytic power for a short time, and allows fine-tuning of the activation of a pathway to the needs of a cell. This difference in time scale emphasizes the importance of applying different platforms for the analysis of a complex tumor as high-grade osteosarcoma.

Conclusions

In summary, this study shows that genomic stability pathways are deregulated on both mRNA and kinome levels, with most significantly affected genes being upregulated and/or phosphorylated. Akt was detected as most probably overactive in osteosarcoma, as downstream peptides were hyperphosphorylated as compared with MSCs. Akt inhibitor MK-2206 could inhibit 2/3 osteosarcoma cell lines. Based on these results, we conclude that Akt inhibitors and other drugs inhibiting the PI3K/Akt/mTOR pathway could have an effect on survival of osteosarcoma tumor cells.

References

[1] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[2] Bacci G, Longhi A, Versari M, Mercuri M, *et al.* Prognostic factors for osteosarcoma of the extremity treated with neoadjuvant chemotherapy. *Cancer*, 106(5):1154–1161, 2006.

[3] Buddingh EP, Anninga JK, Versteegh MIM, Taminiau AHM, *et al.* Prognostic factors in pulmonary metastasized high-grade osteosarcoma. *Pediatric Blood & Cancer*, 54(2):216–221, 2010.

[4] Allison DC, Carney SC, Ahlmann ER, Hendifar A, *et al.* A meta-analysis of osteosarcoma outcomes in the modern medical era. *Sarcoma*, 2012:704872, 2012.

[5] Anninga JK, Gelderblom H, Fiocco M, Kroep JR, *et al.* Chemotherapeutic adjuvant treatment for osteosarcoma: where do we stand? *European Journal of Cancer*, 47(16):2431–2445, 2011.

[6] Hattinger CM, Pasello M, Ferrari S, Picci P, *et al.* Emerging drugs for high-grade osteosarcoma. *Expert Opinion on Emerging Drugs*, 15(4):615–634, 2010.

[7] Cleton-Jansen AM, Buerger H, Hogendoorn PCW. Central high-grade osteosarcoma of bone: diagnostic and genetic considerations. *Current Diagnostic Pathology*, 11(6):390–399, 2005.

[8] Szuhai K, Cleton-Jansen AM, Hogendoorn PCW, Bovée JVMG. Molecular pathology and its diagnostic use in bone tumors. *Cancer Genetics*, 205(5):193–204, 2012.

[9] Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, *et al.* Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes, Chromosomes and Cancer*, 51(7):696–706, 2012.

[10] Wachtel M, Schäfer BW. Targets for cancer therapy in childhood sarcomas. *Cancer Treatment Reviews*, 36(4):318–327, 2010.

[11] Kolb EA, Kamara D, Zhang W, Lin J, *et al.* R1507, a fully human monoclonal antibody targeting IGF-1R, is effective alone and in combination with rapamycin in inhibiting growth of osteosarcoma xenografts. *Pediatric Blood & Cancer*, 55(1):67–75, 2010.

[12] Chawla SP, Staddon AP, Baker LH, Schuetze SM, *et al.* Phase II study of the mammalian target of rapamycin inhibitor ridaforolimus in patients with advanced bone and soft tissue sarcomas. *Journal of Clinical Oncology*, 30(1):78–84, 2012.

[13] Schrage YM, Briaire-de Bruijn IH, de Miranda NFCC, van Oosterwijk JG, *et al.* Kinome profiling of chondrosarcoma reveals SRC-pathway activity and dasatinib as option for treatment. *Cancer Research*, 69(15):6216–6222, 2009.

[14] Willems SM, Schrage YM, Briaire-de Bruijn IH, Szuhai K, *et al.* Kinome profiling of myxoid liposarcoma reveals NF-kappaB-pathway kinase activity and Casein Kinase II inhibition as a potential treatment option. *Molecular Cancer*, 9(1):257, 2010.

[15] Kuijjer ML, Namløs HM, Hauben EI, Machado I, *et al.* mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts. *BMC Medical Genomics*, 4(1):66, 2011.

[16] Mohseny AB, Machado I, Cai Y, Schaefer KL, *et al.* Functional characterization of osteosarcoma cell lines provides representative models to study the human disease. *Laboratory Investigation*, 91(8):1195–1205, 2011.

[17] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[18] Memmott RM, Dennis PA. Akt-dependent and-independent mechanisms of mTOR regulation in cancer. *Cellular Signalling*, 21(5):656–664, 2009.

[19] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

[20] Gorlick R, Maris JM, Houghton PJ, Lock R, *et al.* Testing of the Akt/PKB inhibitor MK-2206 by the pediatric preclinical testing program. *Pediatric Blood & Cancer*, 59(3):518–524, 2012.

[21] Team RDC. R: a language and environment for statistical computing, reference index version 2.15.0. R Foundation for Statistical Computing, 2011.

[22] Buddingh EP, Kuijjer ML, Duim RAJ, Bürger H, *et al.* Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. *Clinical Cancer Research*, 17(8):2110–2119, 2011.

[23] Hilhorst R, Houkes L, Mommersteeg M, Musch J, *et al.* Peptide microarrays for profiling of Serine/Threonine kinase activity of recombinant kinases and lysates of cells and tissue samples. In Bina M, editor, Gene regulation: methods and protocols, methods in molecular biology, 259–271. Springer, 2013.

[24] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):3, 2004.

[25] Huber W, Von Heydebreck A, Sültmann H, Poustka A, *et al.* Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, 18(suppl 1):S96–S104, 2002.

[26] Gentleman RC, Carey VJ, Bates DM, Bolstad B, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004.

[27] Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics*, 25(3):415–416, 2009.

[28] Mohseny AB, Cai Y, Kuijjer ML, Xiao W, *et al.* The activities of Smad and Gli mediated signalling pathways in high-grade conventional osteosarcoma. *European Journal of Cancer*, 48(18):3429–3438, 2012.

[29] Mohseny AB, Tieken C, van der Velden PA, Szuhai K, *et al.* Small deletions but not methylation underlie CDKN2A/p16 loss of expression in conventional osteosarcoma. *Genes, Chromosomes and Cancer*, 49(12):1095–1103, 2010.

[30] Engelman JA, Luo J, Cantley LC. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nature Reviews Genetics*, 7(8):606–619, 2006.

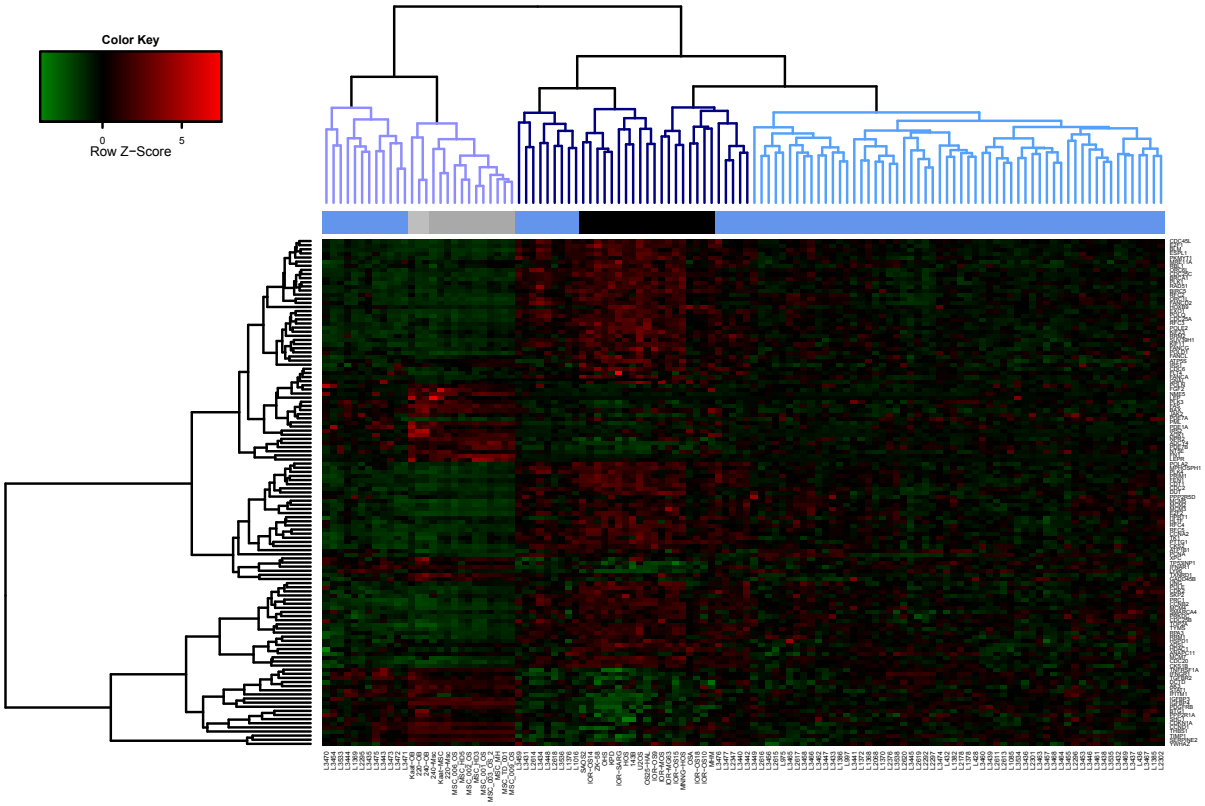
[31] Manning BD, Cantley LC. AKT/PKB signaling: navigating downstream. *Cell*, 129(7):1261–1274, 2007.

[32] Guertin DA, Sabatini DM. Defining the role of mTOR in cancer. *Cancer Cell*, 12(1):9–22, 2007.

[33] Vakana E, Altman JK, Platanias LC. Targeting AMPK in the treatment of malignancies. *Journal of Cellular Biochemistry*, 113(2):404–409, 2012.

[34] Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, *et al.* PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics*, 4(6):1551–1561, 2004.

Additional Figures



Additional Figure 6.1: Unsupervised hierarchical clustering of osteosarcoma cell line data (black bars), control cultures (MSC: dark gray bars, osteoblast: light gray bars), and data from osteosarcoma biopsies (blue bars) on mRNA expression levels of all DE genes present in the 17 significantly affected pathways as determined by IPA. The different clusters selected for Kaplan-Meier analysis are shown in the upper dendrogram in different shades of blue, corresponding to the legend of Figure 6.3. Red: upregulation, green: downregulation.

Chapter 7

Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data

This chapter is based on the publication: Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, Lid AB, Roelofs H, Bürger H, Myklebost O, Hogendoorn PCW, Meza-Zepeda LA, Cleton-Jansen AM. *Genes Chromosomes Cancer.* 2012 Jul;51(7):696-706

Abstract

High-grade osteosarcoma is a tumor with a complex genomic profile, occurring primarily in adolescents with a second peak at middle age. The extensive genomic alterations obscure the identification of genes driving tumorigenesis during osteosarcoma development. To identify such driver genes, we integrated DNA copy number profiles (Affymetrix SNP 6.0) of 32 diagnostic biopsies with 84 expression profiles (Illumina Human-6 v2.0) of high-grade osteosarcoma as compared with its putative progenitor cells, *i.e.* mesenchymal stem cells ($n = 12$) or osteoblasts ($n = 3$). In addition, we performed paired analyses between copy number and expression profiles of a subset of 29 patients for which both DNA and mRNA profiles were available. Integrative analyses were performed in Nexus Copy Number software and statistical language R. Paired analyses were performed on all probes detecting significantly differentially expressed genes in corresponding *LIMMA* analyses. For both nonpaired and paired analyses, copy number aberration frequency was set to $> 35\%$. Nonpaired and paired integrative analyses resulted in 45 and 101 genes, respectively, which were present in both analyses using different control sets. Paired analyses detected $> 90\%$ of all genes found with the corresponding nonpaired analyses. Remarkably, approximately twice as many genes as found in the corresponding nonpaired analyses were detected. Affected genes were intersected with differentially expressed genes in osteosarcoma cell lines, resulting in 31 new osteosarcoma driver genes. Cell division related genes, such as *MCM4* and *LATS2*, were overrepresented and genomic instability was predictive for metastasis-free survival, suggesting that deregulation of the cell cycle is a driver of osteosarcomagenesis.

Introduction

High-grade osteosarcoma is an aggressive primary bone tumor, which mostly occurs during adolescence, with a second peak at middle age, at the metaphysis of long bones. The tumor is characterized by aberrant production of osteoid matrix and by very complex karyotypes (1, 2). Since the introduction of DNA microarray technology, recurrent DNA copy number changes in human osteosarcoma tumor tissues have been identified by comparative genomic hybridization (CGH) and high-density single nucleotide polymorphisms (SNP) microarray analysis. There is a general consensus about gain of chromosome arms 6p, 8q, and 17p, but many additional regions are reported as well (3–7). The effects of copy number alterations may be reflected by changes in expression of genes in the affected chromosomal regions. There are various publications on human osteosarcoma gene expression, but few show robust bioinformatics (as described by Kuijjer *et al.* (8)). Often, small sample sizes and heterogeneity within groups result in only a small number of significant genes, on which usually no correction for multiple testing is applied. Another problem when studying osteosarcoma gene expression data is the lack of an osteosarcoma

benign precursor lesion and its debated cell of origin—although it becomes clearer that the mesenchymal stem cell or its derivative is the progenitor of osteosarcoma (9, 10). The disease seems to develop suddenly as a full-blown tumor, rendering it difficult to detect early drivers of osteosarcomagenesis. We have previously determined differential expression related to specific clinical parameters (8, 11). In addition, we have compared osteosarcoma with osteoblastoma—a benign tumor which develops at the same site as osteosarcoma, but which does not progress into the latter. This comparison of human osteosarcoma with a control tissue showed that cell cycle regulation is the most significantly altered pathway in osteosarcoma (12).

There are advantages of integrating copy number and expression data when aiming to identify driver genes. First, copy number data analysis of tumors with complex genomic profiles may return numerous bystander or hitch-hiker genes, as copy number alterations may occur not only because they are advantageous for the tumor but also as a result of general genomic instability. Regions of copy number alteration may therefore encompass no driver gene at all, or may include additional genes. Also, some genes with altered copy numbers may not be expressed in the tumor due to tissue-specific expression. These aspects hamper the identification of drivers of tumorigenesis, especially when the number of recurrent genes in such tumors is high. Second, at the mRNA level, drivers affect downstream genes and switch on feedback mechanisms, again rendering it difficult to determine the real osteosarcoma drivers in a pool of differentially expressed genes (13). Integration of DNA copy number and gene expression data filters out at least part of such bystanders and of genes that act downstream of drivers of tumorigenesis, because most of these genes have altered copy numbers, but no change in expression, or vice versa, while drivers are both amplified and upregulated, or deleted and downregulated. Particularly osteosarcoma is genetically extremely unstable and therefore genomic data analysis of this tumor type would benefit from an approach that distinguishes driver genes from the numerous more random genetic events.

Nonpaired integrative analysis may be performed by determining recurrent regions of copy number alterations which have higher than expected numbers of differentially expressed genes. Paired integrative analysis is a more powerful method, because the relationship between copy number alterations and gene expression can be inferred in each specific sample, instead of being based on averaged quantities. A statistically correct method for paired integrative analysis of these different data types has not yet been defined. Paired integrative analysis is usually performed by selecting genes based on the correlation between gene expression and copy number levels, such as is performed by the recently published methods DR-Integrator (14) and Regularized dual Canonical Correlation Analysis (15). However, gains and losses may not necessarily directly translate to the same quantity of change in expression levels (13), and important genes may be overlooked this way. A method where paired integrative analysis is detected for specific chromosomal regions with altered genomic and transcriptional status does exist (16), but this method

is not optimal for tumors such as osteosarcoma with highly unstable genomes, since copy number values are normalized to the mean copy number over each array, and this mean value may be altered in such tumors. Two methods, PARADIGM and CNAmet, combine different types of data on a gene-based level. In PARADIGM, integration of different data types is used to detect patient-specific pathway activities (17). CNAmet returns genes that show differential expression between samples with and without methylation and/or copy number alteration (18). This elegant approach may however hamper the identification of genes that are regulated by other frequently altered genes, such as *TP53* and *MDM2* in osteosarcoma.

Aiming to identify osteosarcoma driver genes, we performed both nonpaired and paired integrative analyses on high-grade osteosarcoma prechemotherapy biopsy data. We combined results from analyses as compared with different control sets—mesenchymal stem cells (MSCs) and osteoblasts, so that we did not exclude one of these proposed progenitors as the cell of origin of osteosarcoma. We show that the paired integrative analysis returns more affected genes than the nonpaired integrative analysis. There is an overrepresentation of genes involved in genomic stability in osteosarcoma samples. The identified genes may be important drivers in osteosarcomagenesis.

Materials and methods

Ethics statement

All biological material was handled in a coded fashion. Ethical guidelines of the individual European partner institutions were followed and samples and clinical data were handled in a coded fashion and stored in the EuroBoNeT biobank.

Patient material and cell lines

Genome-wide expression profiling was performed on pretreatment diagnostic biopsies of 84 resectable high-grade osteosarcoma patients from the EuroBoNeT consortium (`www.eurobonet.eu`). Clinicopathological details of these samples can be found in Table 7.1. Human bone-marrow-derived MSCs were obtained from five osteosarcoma patients and seven healthy donors. Osteoblasts ($n = 3$) were derived from MSCs on osteogenic differentiation. MSCs and osteoblasts were characterized and handled as described (12). Copy number analysis was performed on 32 pretreatment diagnostic biopsies, of which 29 overlapped with the 84 samples used for expression analysis.

Copy number microarray data analysis

Affymetrix Genome-Wide Human SNP 6.0 arrays (Affymetrix, Santa Clara, CA) were used for SNP data analysis. Genomic DNA preparation, labeling, hybridization, and

Category	Patient characteristics	Number of biopsies (%)
Institution	LUMC, Netherlands	36 (42.9)
	IOR, Italy	12 (14.3)
	LOH, Sweden	3 (3.6)
	Radiumhospitalet, Norway	1 (1.2)
	WWUM, Germany	32 (38.1)
Location primary tumor	Femur	40 (47.6)
	Tibia/Fibula	28 (33.3)
	Humerus	11 (13.1)
	Axial skeleton	1 (1.2)
	Unknown/other	4 (4.8)
Histological subtype	Osteoblastic	52 (61.9)
	Chondroblastic	9 (10.7)
	Fibroblastic	7 (8.3)
	Telangiectatic	4 (4.8)
	Minor subtype	11 (13.1)
	Unknown	1 (1.2)
Huvos grade	1 or 2	38 (45.2)
	3 or 4	33 (39.3)
	Unknown/NA	14 (16.7)
Metastasis at diagnosis	Yes	14 (16.7)
	No	69 (82.1)
	Unknown	1 (1.2)
Sex	Male	54 (64.3)
	Female	29 (34.5)
	Unknown	1 (1.2)
Age	$<$ 20 years	64 (76.2)
	$>=$ 20 years	19 (22.6)
	Unknown	1 (1.2)

Table 7.1: Clinicopathological details.

scanning were performed as described by Kresse *et al.* (7). Microarray data preprocessing was performed as described previously (19). Hybridization quality was estimated by the genotype call rate using the Birdseed genotype calling algorithm in Genotyping Console (version 4.0, Affymetrix). Samples of poor quality were excluded from further analyses. We performed copy number analysis in Nexus software version 5 (Biodiscovery, El Segundo, CA) using CNCHP log ratio files generated by Genotyping Console using 27 controls as a baseline, which is a subset of the reference set of 29 samples which was used by Pansuriya *et al.* (19). We rejected two samples based on results from the quality control analysis in Genotyping Console. Circular Binary Segmentation (CBS)-based SNPRank Segmentation was used to identify aberrant genomic regions. To be included as frequently aberrant, a copy number alteration was called when detected in at least 35% of all cases. Correlation of copy number alterations with clinical data was performed in Nexus software, with correction for multiple testing.

Genome-wide gene expression microarray data analysis

Osteosarcoma tissue handling, RNA isolation, synthesis of cDNA, cRNA amplification, hybridization of cRNA onto the Illumina Human-6 v2.0 Expression BeadChips (Illumina, San Diego, CA), and microarray data processing and quality control in the statistical language R version 2.10 (20) were performed as described previously (11). High correlations between these microarray data and corresponding qPCR results have been demonstrated previously (11). Unsupervised hierarchical cluster analysis was performed using R package pvclust with default settings (21).

Data deposition

MIAME-compliant copy number and gene expression data have been deposited in the GEO database (`www.ncbi.nlm.nih.gov/geo/`, superseries accession number GSE33383).

Detection of significantly differentially expressed genes

We performed a *LIMMA* analysis (22) in order to determine differential expression between high-grade osteosarcoma samples ($n = 84$) and control tissues—MSCs ($n = 12$) and osteoblasts ($n = 3$). Also, gene expression differences between MSCs and osteoblasts were determined. We used a Benjamini and Hochberg False Discovery Rate (FDR) of 0.05 as cut-off for significance.

Nonpaired integrative analysis

Nonpaired integrative analysis was performed by importing lists of differentially expressed genes into the Copy Number module of Nexus software version 5. Based on the length of

the gene list, Nexus software performs a Fisher's exact test in order to determine whether the number of differentially expressed genes in a specific region with a significant copy number alteration is larger than expected by chance. Genes present in such regions of copy number alteration with FDR-adjusted p-values (Q-bounds in Nexus software) < 0.05 were returned from this integrative analysis. We did not apply any restrictions on the size of copy number aberrations. A few small altered regions that did not encompass an entire gene were detected, but these regions did not return genes upon integration with expression data. Nexus software only reports genes which are both gained and overexpressed, or both deleted and downregulated.

Paired integrative analysis

For the paired integrative analysis, copy number data of all autosomal overlapping genes between the copy number and gene expression data were exported from Nexus software, and converted into a binary matrix containing all genes with a gain (1) and no gain (0), and a similar binary matrix for losses. As in the nonpaired integrative analysis, we did not apply any restrictions on the size of copy number alterations. Gene expression data of each probe for each sample were normalized against average gene expression values of the corresponding probes over all control samples (either expression data from 12 MSCs or from three osteoblasts)—this was performed by subtracting the average expression of the control samples from the expression levels of the sample of interest, since these are log-transformed expression values. For both analyses, only genes that were significantly differentially expressed between the 84 osteosarcoma samples and the specific control set were analyzed, in order to make sure that all genes returned from the integrative analysis were significantly differentially expressed. Subsequently, genes that overlapped between the copy number binary matrices and that matched the fold change of expression (upregulation for genes with gains, and downregulation for genes with losses) were returned as two-way contingency tables using scripts in R. Genes that were altered in two types of data were further filtered by applying the sample recurrence criterion of 35%. This generated lists of recurrent two-way altered genes. The odds ratios for having both copy number and expression changes were calculated for different combinations, for instance gain and upregulation. We used Bonferroni corrected Chi-square or Fisher's exact p-values < 0.05 to determine significance.

Gene set enrichment

GO term enrichment was tested using Bioconductor package *topGO* (23). Lists of significantly affected genes were compared with all genes eligible for the analysis. GO terms with Fisher's exact p-values < 0.0001, as calculated by the *weight01* algorithm from *topGO*, were defined significant.

Genomic instability scores and survival analysis

We calculated genomic instability scores for 83 (out of the 84) osteosarcoma biopsies (for one sample no follow-up data were available) and all controls, as well as for two normal bone samples (obtained from cancer patients at the Norwegian Radium Hospital), 20 osteosarcoma xenografts (Kresse *et al.*, *unpublished data*, and Kuijjer *et al.* (8)), 19 osteosarcoma cell lines (24), and the HeLa cervical cancer cell line. For calculation of the genomic instability scores, we refer to the article by Carter *et al.* (25). In short, this method calculates per sample per probe the expression of that particular probe minus the mean expression of that probe over all samples. For each sample, the sum of these values for all probes present in the genomic instability signature is calculated. This value is then compared between all samples and thus gives a relative measure of genomic instability. We used 24 genes of the CIN25 signature, because for one gene no probe was present on the Illumina v2.0 BeadChip. For genes with multiple probes, we used the probe that showed the highest variation in expression levels. We determined metastasis-free survival using the Kaplan-Meier method and performed a Logrank test for trend using GraphPad Software (La Jolla, CA, www.graphpad.com).

Results

Recurrent chromosomal regions with copy number aberrations in high-grade osteosarcoma

Thirty two high-grade osteosarcoma prechemotherapy biopsies were hybridized to Affymetrix SNP 6.0 arrays in order to determine recurrent copy number alterations. In total, 67 regions with recurrent alterations were detected, of which 35 regions had copy number gain, and 32 copy number loss (see Supporting Information Table 1 (*available online* (26))). Recurrent gains were present on chromosome arms 1p, 1q, 4p, 5p, 6p, and 8q, and losses on chromosome arms 1p, 1q, 2q, 3q, 6q, 7q, 8p, 10p, 10q, 12p, 13q, 15q, 16p, and 16q. A genome-wide frequency plot of copy number alterations is shown in Figure 7.1. No significant correlation was detected for specific regions with copy number alterations and clinical information (tested clinical parameters are shown in Table 7.1).

Comparison of osteoblasts and MSCs

Unsupervised hierarchical cluster analysis resulted in separate clusters for biopsies and cell lines. Within the cell line cluster, osteosarcoma cell lines formed one subcluster, whereas MSCs and osteoblasts formed a second subcluster (Supporting Information Figure 1 (*available online* (26))). This indicates that the control cell lines are more similar to one another than to osteosarcoma cells. On the basis of hierarchical clustering of gene expression data, we cannot determine the cell of origin of osteosarcoma. A total

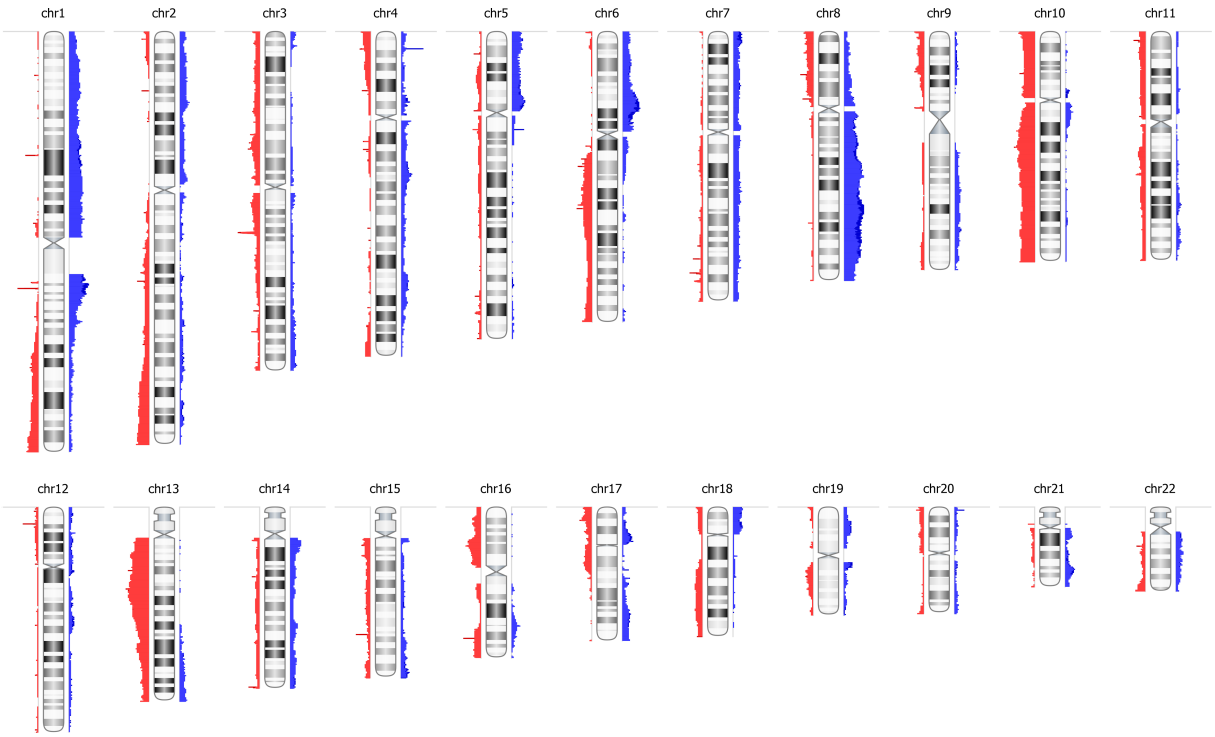


Figure 7.1: Genome-wide frequency plot of copy number alterations on chromosomes 1–22 in 32 high-grade osteosarcoma prechemotherapy biopsies. Left of the chromosomes: loss, right: gain.

of 1,382 genes were differentially expressed between osteoblasts and MSCs. GO term enrichment resulted in seven significant GO terms, which are represented in Supporting Information Figure 2 (*available online* (26)). In summary, GO term enrichment showed differences in cellular structure, proliferation, and apoptosis. Genes showing significant differences between both control cell types, however, can nonetheless be differentially expressed between osteosarcoma samples and both control cell types, thus can still be important drivers of osteosarcomagenesis. We therefore set out to select genes that showed differential expression in osteosarcoma as compared with both MSCs and osteoblasts.

Gene expression signature of high-grade osteosarcoma

We detected 12,542 and 2,939 probes encoding for genes that were significantly differentially expressed between the 84 osteosarcoma biopsies and MSCs and osteoblasts, respectively. MA plots, showing log-intensity ratios and log-intensity averages for both analyses, are depicted in Supporting Information Figure 3 (*available online* (26)). A total of 1,679 probes overlapped between both analyses, of which 1,639 were either up- or downregulated in both. GO term analysis on the genes represented by these 1,639 probes showed an enrichment of apoptosis and signal transduction genes. Antigen processing and presentation, as well as angiogenesis were also overrepresented (Supporting Information Figure 4 (*available online* (26))).

Paired integrative analysis is more sensitive than nonpaired integrative analysis

Nonpaired integrative analysis was performed on data from 32 samples hybridized on SNP arrays and from 84 samples hybridized on gene expression arrays, whereas paired analysis was performed on a subset of 29 samples for which both SNP and expression data were available. In total, 16,105 autosomal genes were represented both on SNP and on gene expression arrays. Nonpaired integrative analysis resulted in 253 significantly affected genes in osteosarcoma biopsies versus mesenchymal stem cells, whereas 71 genes were detected when osteoblasts were used as a control. A total of 45 genes were identified in both analyses versus MSCs and versus osteoblasts (Figure 7.2). Of these 45 genes, 23 were also

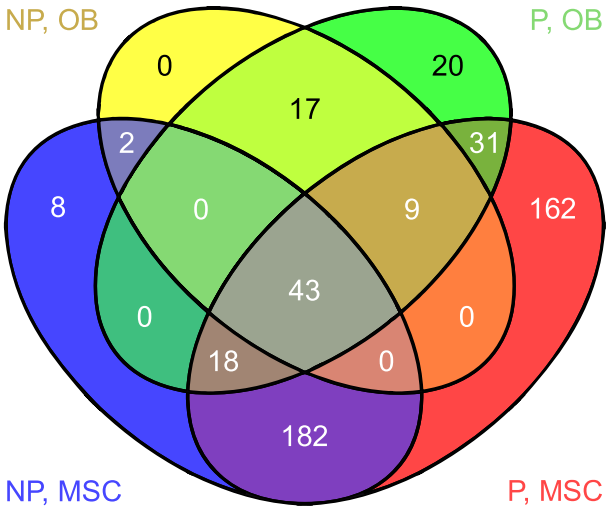


Figure 7.2: Venn diagram with numbers of affected genes in both nonpaired and paired analyses, and in osteosarcoma biopsies versus MSCs and versus osteoblasts. NP: nonpaired integrative analysis, P: paired integrative analysis, OB: analysis of osteosarcoma biopsies versus osteoblasts, MSC: analysis of osteosarcoma samples versus mesenchymal stem cells.

detected in expression analyses of a panel of 19 osteosarcoma cell lines (24) versus MSCs and osteoblasts (Supporting Information Figure 5A (*available online* (26))). For the paired integrative analyses, we determined whether the number of genes with gain combined with overexpression and with loss combined with downregulation was higher than expected per sample, based on the numbers of copy number alterations and gene expression changes in the whole genome. This was true for most samples, as depicted in Figure 7.3, in which the odds ratios and significance of data dependencies are shown. Paired integrative analysis resulted in 445 and 138 genes when compared with MSCs and osteoblasts, respectively. A total of 101 genes overlapped between these different analyses (Figure 7.2), and of this set, 31 genes were also detected in the cell line expression data (Supporting Information Figure 5B (*available online* (26)), Table 7.2). Hence, paired analyses detected $> 90\%$ of all genes found with corresponding nonpaired analyses. In addition, approximately twice as many genes as found in the corresponding nonpaired analyses were detected (Figure 7.2, Supporting Information Figure 6 (*available online* (26))). Note that in the paired analysis fewer samples are included. Thus, paired analysis gives more robust results despite the lower sample size. Changing the threshold of FDR-adjusted p-values in the nonpaired integrative analysis from 0.05 to 0.15 (*data not shown*) did not alter this ratio.

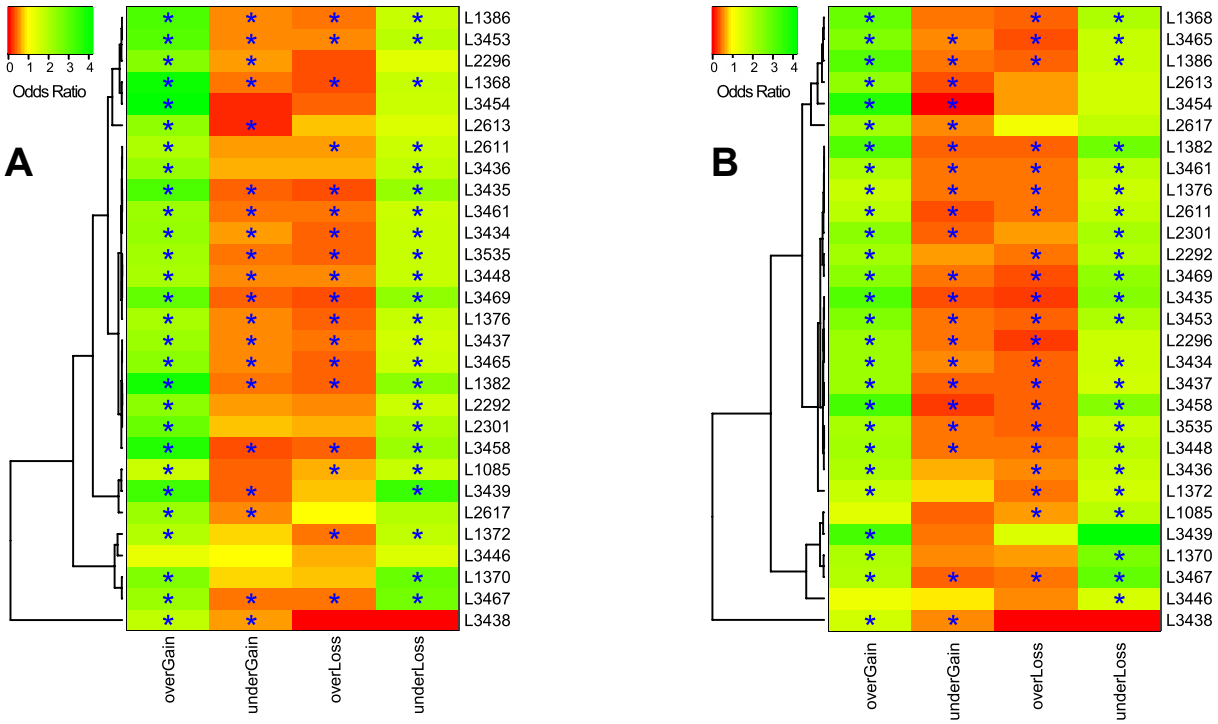


Figure 7.3: Dependence of gene copy number and gene expression data. The heatmaps depict odds ratios for the numbers of genes per sample which show gain and overexpression (overGain), gain and underexpression (underGain), loss and overexpression (overLoss), and loss and underexpression (underLoss). Chi-square tests, or, in case a group contained < 10 genes, Fisher's exact tests, were performed in order to evaluate whether the number of genes reported from the integrative analysis was higher than expected by chance. $*$ Bonferroni-corrected p-values < 0.05. *A*, Osteosarcoma biopsies versus MSCs, *B*, versus osteoblasts.

Symbol	Cytoband	CNA	CNA freq (%)	logFC
CLCC1	1p13.3	Gain	41.4	1.24
MCM4	8q11.21	Gain	37.9	1.35
AKR1C3	10p15.1	Loss	37.9	-1.94
AKR1C4	10p15.1	Loss	37.9	-1.34
ARHGAP22	10q11.22	Loss	37.9	-0.45
PGBD3	10q11.23	Loss	41.4	-0.82
ARID5B	10q21.2	Loss	48.3	-2.33
REEP3	10q21.3	Loss	48.3	-0.51
HERC4	10q21.3	Loss	51.7	-1.31
PBLD	10q21.3	Loss	48.3	-0.29
RUFY2	10q21.3	Loss	48.3	-0.20
KIAA1279	10q22.1	Loss	43.1	-0.57
SRGN	10q22.1	Loss	43.1	-2.26
AIFM2	10q22.1	Loss	44.8	-0.52
CHST3	10q22.1	Loss	48.3	-1.17
FAS	10q23.31	Loss	44.8	-0.42
PCGF5	10q23.32	Loss	37.9	-0.34
PPP1R3C	10q23.32	Loss	37.9	-2.89
AVPI1	10q24.2	Loss	37.9	-2.35
BLOC1S2	10q24.31	Loss	37.9	-0.51
CASC2	10q26.11	Loss	44.8	-0.18
FAM45A	10q26.11	Loss	39.7	-0.78
ERCC6	13q11.23	Loss	41.4	-0.52
WASF3	13q12.13	Loss	44.8	-2.43
C13orf33	13q12.3	Loss	48.3	-2.26
LHFP	13q14.11	Loss	48.3	-1.89
WBP4	13q14.11	Loss	55.2	-0.93
TSC22D1	13q14.11	Loss	58.6	-1.39
RCBTB1	13q14.2	Loss	58.6	-0.25
LATS2	13q21.11	Loss	44.8	-0.96
DCUN1D3	16p12.3	Loss	37.9	-1.39

Table 7.2: Candidate osteosarcoma driver genes. All frequencies and fold changes are mean values of both integrative analyses—osteosarcoma biopsies versus MSCs and osteosarcoma biopsies versus osteoblasts. For genes for which more than one probe was present on the array, the probe with the highest fold change was used. Cytoband: UCSC cytogenetic band, CNA: copy number aberration, CNA freq: copy number aberration frequency (for $n = 29$), logFC: log fold change in biopsies (negative means downregulation, positive means upregulation).

Genomic instability genes play a role in osteosarcoma progression

We calculated genome instability scores using the method of Carter *et al.* (25), which compares levels of gene expression of a previously defined genomic instability signature between samples in a dataset, for all osteosarcoma biopsies and different control tissues and cell lines (Figure 7.4A). The osteosarcoma biopsies showed highly variable scores,

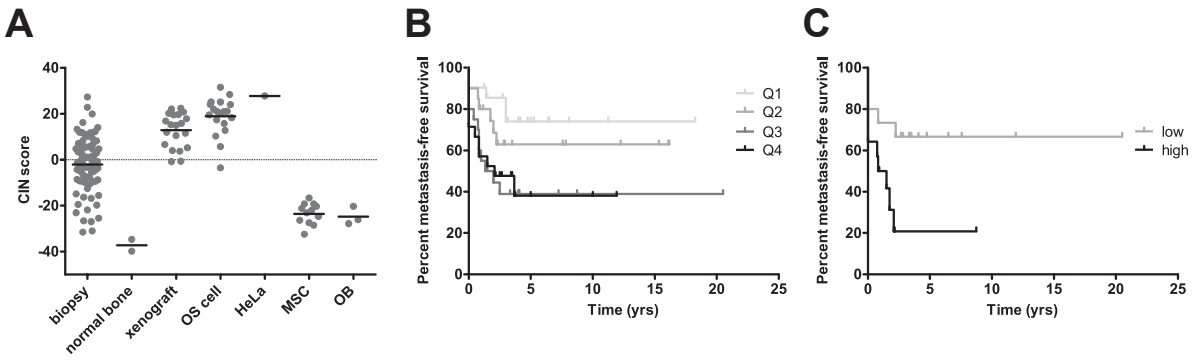


Figure 7.4: Genomic instability scores and metastasis-free survival. *A*, Genomic instability scores for high-grade osteosarcoma biopsies, normal bone, osteosarcoma xenografts and cell lines, the HeLa cell line, and mesenchymal stem cells (MSC) and osteoblasts (OB), as calculated by the method of Carter *et al.* (25). *B*, Metastasis-free survival Kaplan-Meier curves for four quartiles of genomic instability scores. *C*, Metastasis-free survival Kaplan-Meier curves for the total amount of genes with copy number gains and losses, using a cut-off based on the median amount of genes per sample showing copy number aberration.

whereas genomic instability scores for the controls, normal bone, MSCs, and osteoblasts were relatively low. High instability scores were detected for osteosarcoma xenografts, cell lines, and the HeLa cell line, in increasing order. This signature predicted for metastasis-free survival in osteosarcoma samples as well (Figure 7.4B), with high scores correlating with shorter metastasis-free survival (Logrank test for trend p-value = 0.0112). As expected, the total number of genes with copy number gains or losses, which is a direct measure of genomic instability from the SNP data, was predictive for progression as well (Logrank test p-value = 0.018, Figure 7.4C).

Candidate osteosarcoma driver genes

The 31 genes returned by the paired integrative analysis on clinical samples that also were differentially expressed in osteosarcoma cell lines are shown in Table 2, together with their chromosomal locations, aberration frequencies, and log fold changes. A total of 22/31 genes have been described to play a role in cancer. Interestingly, one third of these 22 genes have a role in cell cycle regulation, matching the importance of cell cycle and replication in osteosarcomagenesis as was found both using the genomic instability scores of the expression data and the overall chromosomal instability as detected in the copy number data (Figure 7.4).

Discussion

In this study, we report copy number and gene expression alterations in high-grade osteosarcoma prechemotherapy biopsies, and then integrate these data in order to detect osteosarcoma driver genes. Copy number analyses, which were obtained with high-density SNP microarrays, showed very high genomic instability in the osteosarcoma biopsies. The pattern of aberrations is in line with previous studies using aCGH and SNP arrays, which show recurrent gains in chromosome arms 1p, 6p, and 8q, and losses in chromosome 10. The previously reported recurrent amplification on chromosome arm 17p (3–6) is not listed, because we used a very strict cut-off for aberration frequency (35%). Aberration frequencies of 17% (4) and 26% (6) were previously found on chromosome arm 17p, and a distinct amplification in 17p with an aberration frequency of 21% can be seen in Figure 1. We chose such a high cut-off for recurrent aberrations in order to enrich for selected genetic events and exclude the numerous haphazard alterations that can be attributed to the high genomic instability of high-grade osteosarcoma. In addition, we previously determined that this cut-off, as compared with cut-offs of 15% and 50%, showed the most consistent results in subsequent network and pathway analyses on osteosarcoma cell line SNP data (*data not shown*). For genome-wide gene expression analyses, both MSCs and osteoblasts were used as control cells, and we only considered overlapping genes between both comparisons, in order to make sure the affected genes were differentially regulated in osteosarcoma when compared with its putative progenitor cells. This analysis identified a large number ($n = 1,639$) of probes encoding for differentially expressed genes. Many of these genes encode tissue type-specific proteins, as is shown in the GO term enrichment analysis, and appear as upregulated in osteosarcoma biopsies because the *in vitro* grown control cells, MSCs and osteoblasts, lack surrounding stroma and are nurtured under other conditions. Antigen processing and presentation as well as angiogenesis pathways were expected to be upregulated, as macrophages and other infiltrating cells are present in osteosarcoma tissue (11), and as angiogenesis plays a role in osteosarcoma progression (27). Nevertheless, most stroma-derived gene expression is filtered out by integration with copy number data, as this expression is not a result of underlying copy number changes. In addition to stroma-related gene sets, GO term analysis showed enrichment in apoptosis and signal transduction genes, which are probably altered in the osteosarcoma tumor cells and not in the stroma. Because genes with concordant changes in copy number and gene expression are likely to be enriched in drivers of tumorigenesis, we performed integrative analyses on both types of data.

We found a remarkable increase in significant differential genes in paired compared with nonpaired analysis, *i.e.* 101 versus 45. In general, paired integrative analysis was advantageous over nonpaired integrative analysis, identifying roughly twice as many genes, also when different aberration frequency cut-offs or less stringent cut-offs for significance were used in the nonpaired analysis. Nonpaired analysis as performed in Nexus software

compares the number of differentially expressed genes in a region of copy number aberration with the expected number of differentially expressed genes, which is based on the total number of differentially expressed genes over the whole genome. This method may be too rigorous, because an altered copy number region may encompass tissue-specific genes, which may not be expressed in the particular tumor tissue. These genes then have altered copy number, but no difference in expression. If an altered copy number region contains a relatively large number of such genes plus only a few candidate drivers, the entire region will be removed from the output of the analysis, which increases the amount of false negatives. Moreover, in the cancer gene expression profile, a large number of genes downstream of drivers, *i.e.* directly or indirectly regulated by drivers, or present in feedback loops will be differentially expressed. This increases the total number of differentially expressed genes, which again lowers the chance that a specific altered region is returned from the nonpaired integrative analysis as significantly affected. Furthermore, a single differentially expressed gene in a certain region of copy number alteration may still exert its driving function, and this driving function usually does not depend on the proportion of differentially expressed genes in the same region. Because of this, and because our method of paired integrative analysis is gene-based and not region-based, we did not perform a correction based on the total number of differentially expressed genes when compared with the affected copy number regions in the paired analysis in R, and this may be an additional reason why more genes are returned from the paired analysis. However, in all samples, except for one (L3438), the number of genes showing both copy number alteration and differential expression was higher than expected when compared with the numbers of copy number alterations and differentially expressed genes over the whole genome. This was significant for the vast number of samples (28/29, 23/29, 27/29, and 23/29, for combinations gain and overexpression, loss and underexpression in biopsies versus MSCs, and gain and overexpression, loss and underexpression in biopsies versus osteoblasts, respectively, as shown in Figure 7.3).

Genomic instability scores showed that the instability in osteosarcoma tissues ranges from a level comparable to that of the controls, to the high instability levels of repeatedly passaged tumors in xenografts and osteosarcoma cell lines. We demonstrated both on copy number data, as well as by applying a genomic instability gene signature to genome-wide gene expression data, that high genomic instability in osteosarcoma is correlated with poor metastasis-free survival. This suggests that genes playing a role in genomic instability may be potent drivers of osteosarcoma progression, as has been reported for various other tumor types (25). Paired integrative analysis confirmed this result, as one third of the genes with a possible role in tumorigenesis had a function connected to the cell cycle. Of these genes, *MCM4* showed gain and overexpression and was only detected by the paired integrative analysis. *MCM4* is part of the minichromosome maintenance complex, which functions as a replication helicase, with a role in maintaining genomic stability (28). This gene has been reported overexpressed in various tumor types (29–31).

Genes that were detected in both nonpaired and paired analyses were all deleted and underexpressed. *AVPI1*, or arginine vasopressin-induced 1, may be involved in cell cycling (32). *ERCC6* is involved in transcription-coupled nucleotide excision repair, which is a critical survival pathway protecting against cancer (33). *RCBTB1*, a candidate tumor suppressor, was recently shown to have growth inhibitory activity in osteosarcoma cells by regulating pathways of DNA damage/repair and apoptosis (34). *LATS2*, or large tumor suppressor homolog 2, plays a critical role in centrosome duplication, maintenance of mitotic fidelity, and genomic instability (35). Positive feedback between the p53 and Lats2 tumor suppressors prevents tetraploidization (36), which could be an initiating step in osteosarcomagenesis, leading to genomic instability (37, 38). Also, a role of Lats2 in quenching of the increased genomic instability of H-Ras-induced transformation has been identified (36). *DCUN1D3* encodes for a UVC-responsive protein involved in cell cycle progression and cell growth (39). Additional candidate genes with no direct role in cell cycle regulation include for example genes with a role in apoptosis (*AIFM2*, *BLOC1S2*, *FAS*) and metabolism (*AKR1C3* and *-4*). Some previously reported genes with a driver role in osteosarcoma were not identified, mainly because our high cut-off for recurrence. For example, *CDKN2A*, *MDM2*, and *E2F2* had recurrence frequencies of 28%, 17%, and 34%, respectively (in the dataset of 29 samples). *CDKN2A* and *MDM2* were not significantly differentially expressed, but *E2F2* was consistently significantly overexpressed with log fold changes > 1.50 in all analyses (biopsies and cell lines as compared with different controls). *TP53* and *RB1* aberrations were present in $> 35\%$ of all samples (38% and 69%, respectively). *TP53* was significantly downregulated in biopsies as compared with both controls, but not in the osteosarcoma cell line dataset. *RB1* showed significant downregulation when compared with MSCs, but not with osteoblasts, indicating a difference between these controls in *RB1* signaling. We set our cut-off for recurrence to 35% and only selected genes present both in osteosarcoma biopsies as well as in cell lines as compared with two different control sets, in order to select for the most important osteosarcoma drivers. Using this method, we were able to detect previously unreported driver genes.

In summary, we have shown that an individual gene-based paired integrative analysis of copy number and gene expression data performs better than a region-based nonpaired analysis. Several osteosarcoma candidate driver genes, especially genes playing a role in cell cycle progression, have been identified. Additional research, particularly functional studies, should reveal whether these genes are early or late drivers in osteosarcomagenesis.

References

[1] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[2] Cleton-Jansen AM, Buerger H, Hogendoorn PCW. Central high-grade osteosarcoma of bone: diagnostic and genetic considerations. *Current Diagnostic Pathology*, 11(6):390–399, 2005.

[3] Squire JA, Pei J, Marrano P, Beheshti B, *et al.* High-resolution mapping of amplifications and deletions in pediatric osteosarcoma by use of CGH analysis of cDNA microarrays. *Genes, Chromosomes and Cancer*, 38(3):215–225, 2003.

[4] Man TK, Lu XY, Jaeweon K, Perlaky L, *et al.* Genome-wide array comparative genomic hybridization analysis reveals distinct amplifications in osteosarcoma. *BMC Cancer*, 4(1):45, 2004.

[5] Atiye J, Wolf M, Kaur S, Monni O, *et al.* Gene amplifications in osteosarcoma-CGH microarray analysis. *Genes, Chromosomes and Cancer*, 42(2):158–163, 2005.

[6] Yen CC, Chen WM, Chen TH, Chen WYK, *et al.* Identification of chromosomal aberrations associated with disease progression and a novel 3q13.31 deletion involving LSAMP gene in osteosarcoma. *International Journal of Oncology*, 35(4):775–788, 2009.

[7] Kresse SH, Szuhai K, Barragan-Polania AH, Rydbeck H, *et al.* Evaluation of high-resolution microarray platforms for genomic profiling of bone tumours. *BMC Research Notes*, 3(1):223, 2010.

[8] Kuijjer ML, Namløs HM, Hauben EI, Machado I, *et al.* mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts. *BMC Medical Genomics*, 4(1):66, 2011.

[9] Mohseny AB, Szuhai K, Romeo S, Buddingh EP, *et al.* Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of Cdkn2. *The Journal of Pathology*, 219(3):294–305, 2009.

[10] Mohseny AB, Hogendoorn PCW. Concise review: mesenchymal tumors: when stem cells go mad. *Stem Cells*, 29(3):397–403, 2011.

[11] Buddingh EP, Kuijjer ML, Duim RAJ, Bürger H, *et al.* Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. *Clinical Cancer Research*, 17(8):2110–2119, 2011.

[12] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

[13] Lee H, Kong SW, Park PJ. Integrative analysis reveals the direct and indirect interactions between DNA copy number aberrations and gene expression changes. *Bioinformatics*, 24(7):889–896, 2008.

[14] Salari K, Tibshirani R, Pollack JR. DR-Integrator: a new analytic tool for integrating DNA copy number and gene expression data. *Bioinformatics*, 26(3):414–416, 2010.

[15] Soneson C, Lilljebjörn H, Fioretos T, Fontes M. Integrative analysis of gene expression and copy number alterations using canonical correlation analysis. *BMC Bioinformatics*, 11(1):191, 2010.

[16] Bicciato S, Spinelli R, Zampieri M, Mangano E, *et al.* A computational procedure to identify significant overlap of differentially expressed and genomic imbalanced regions in cancer datasets. *Nucleic Acids Research*, 37(15):5057–5070, 2009.

[17] Vaske CJ, Benz SC, Sanborn JZ, Earl D, *et al.* Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics*, 26(12):i237–i245, 2010.

[18] Louhimo R, Hautaniemi S. CNAmet: an R package for integrating copy number, methylation and expression data. *Bioinformatics*, 27(6):887–888, 2011.

[19] Pansuriya TC, Oosting J, Krenács T, Taminiau AH, *et al.* Genome-wide analysis of Ollier disease: is it all in the genes? *Orphanet Journal of Rare Diseases*, 6(2), 2011.

[20] Team RDC. R: a language and environment for statistical computing, reference index version 2.10.0. R Foundation for Statistical Computing, 2005.

[21] Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, 22(12):1540–1542, 2006.

[22] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):3, 2004.

[23] Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–1607, 2006.

[24] Ottaviano L, Schaefer KL, Gajewski M, Huckenbeck W, *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes, Chromosomes and Cancer*, 49(1):40–51, 2010.

[25] Carter SL, Eklund AC, Kohane IS, Harris LN, *et al.* A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nature Genetics*, 38(9):1043–1048, 2006.

[26] Supporting information Chapter 7. `onlinelibrary.wiley.com/doi/10.1002/gcc.21956/suppinfo`.

[27] Lee YH, Tokunaga T, Oshika Y, Suto R, *et al.* Cell-retained isoforms of vascular endothelial growth factor (VEGF) are correlated with poor prognosis in osteosarcoma. *European Journal of Cancer*, 35(7):1089, 1999.

[28] Aguilera A, Gómez-González B. Genome instability: a mechanistic view of its causes and consequences. *Nature Reviews Genetics*, 9(3):204–217, 2008.

[29] Freeman A, Morris LS, Mills AD, Stoeber K, *et al.* Minichromosome maintenance proteins as biological markers of dysplasia and malignancy. *Clinical Cancer Research*, 5(8):2121–2132, 1999.

[30] Alison MR, Hunt T, Forbes SJ. Minichromosome maintenance (MCM) proteins may be pre-cancer markers. *Gut*, 50(3):290–291, 2002.

[31] Majid S, Dar AA, Saini S, Chen Y, *et al.* Regulation of minichromosome maintenance gene family by microRNA-1296 and genistein in prostate cancer. *Cancer Research*, 70(7):2809–2818, 2010.

[32] Apweiler R, Martin MJ, O'Donovan C, Magrane M, *et al.* Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Research*, 39:D214–D219, 2011.

[33] Fousteri M, Mullenders LHF. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Research*, 18(1):73–84, 2008.

[34] Zhou X, Münger K. Clld7, a candidate tumor suppressor on chromosome 13q14, regulates pathways of DNA damage/repair and apoptosis. *Cancer Research*, 70(22):9434–9443, 2010.

[35] Visser S, Yang X. LATS tumor suppressor: a new governor of cellular homeostasis. *Cell Cycle*, 9(19):3892–3903, 2010.

[36] Aylon Y, Michael D, Shmueli A, Yabuta N, *et al.* A positive feedback loop between the p53 and Lats2 tumor suppressors prevents tetraploidization. *Genes & Development*, 20(19):2687–2700, 2006.

[37] Ganem NJ, Pellman D. Limiting the proliferation of polyploid cells. *Cell*, 131(3):437–440, 2007.

[38] Ganem NJ, Storchova Z, Pellman D. Tetraploidy, aneuploidy and cancer. *Current Opinion in Genetics & Development*, 17(2):157–162, 2007.

[39] Ma T, Shi T, Huang J, Wu L, *et al.* DCUN1D3, a novel UVC-responsive gene that is involved in cell cycle progression and cell growth. *Cancer Science*, 99(11):2128–2135, 2008.

Chapter 8

Frequent loss of heterozygosity and amplification in high-grade osteosarcoma: analysis of recurrent tumor suppressor genes

This chapter is based on the manuscript: Kuijjer ML, Liebelt F, Rydbeck H, Myklebost O, Meza-Zepeda LA, Szuhai K, Hogendoorn PCW, Cleton-Jansen AM. *In preparation*

Abstract

High-grade osteosarcoma is an aggressive primary bone tumor, with a peak incidence at adolescence. Osteosarcoma karyotypes are heterogeneous, and characterized by a high degree of genomic instability, rendering it difficult to detect recurrent driver genes in this tumor type. By performing Affymetrix SNP 6.0 microarray data analysis of 29 prechemotherapy biopsies, we observed that loss of heterozygosity (LOH) often cooccurs with copy number gains. We performed paired integrative analysis with genome-wide expression data in order to determine which genes show differential expression in regions of LOH and gain, and integrated data from biopsies with data from 12 cell lines, which resulted in the identification of 29 recurrent genes with LOH, gain, and overexpression. We validated LOH of candidate tumor suppressor genes by Sanger sequencing and screened for mutations in candidate genes, as osteosarcoma cells may have selected for LOH and amplification of mutated tumor suppressors. We did not identify recurrent mutations, suggesting that these genes do not have a tumor driving function. Fluorescence in situ hybridization (FISH) analysis of candidate gene *XRCC6BP1* showed that this gene was present on homogeneous staining regions (HSRs) in 1/2 cell lines. As we detected a large number of recurrent candidate oncogenes by paired integrative analysis of LOH, gain, and overexpression, it may be valuable to determine whether these candidate oncogenes are present on HSRs in osteosarcoma.

Background

High-grade osteosarcoma is the most frequent primary malignant bone tumor, affecting roughly five persons in a population of one million each year (1). The tumor is highly aggressive, leading to distant metastases in approximately 45% of all patients. Since the introduction of neoadjuvant chemotherapy in the 1970s, survival profiles have reached a plateau. In order to identify specific targets for therapy it is important to screen for recurrent driver genes in osteosarcoma. High-grade osteosarcoma karyotypes are characterized by a high level of genomic instability, often harboring numerous numerical and structural changes, and high degree of aneuploidy (2). This results in many frequently affected genes in osteosarcoma, which may not all be important drivers, and thus renders it difficult to determine which genes are true drivers of osteosarcomagenesis. Integration of genomic and transcriptomic data will filter out most bystander and tissue-specific genes, and can thereby result in a more specific list of candidate recurrent drivers. We previously detected novel osteosarcoma driver genes by integrating high-throughput copy number and gene expression data (3). Zygosity status can also be retrieved from SNP microarray data, which we describe in the present study.

Loss of heterozygosity (LOH) and allelic imbalance have been studied in osteosarcoma to quite some extent, and several recurrent regions have been described in detail (4–8).

Smida *et al.* reported that the amount of LOH negatively correlated with survival (9). This may be a readout of general genomic instability of the tumor, as for multiple human cancers (10), including osteosarcoma (3), it has been shown that genomic instability is predictive for survival, but LOH of specific regions may play an important role in tumorigenesis of osteosarcoma. Loss of heterozygosity caused by the loss of one allele may cause downregulation of the transcript, which may especially be relevant for tumorigenesis when an affected tumor suppressor gene shows haploinsufficiency (11). In osteosarcoma, *TP53*, *RB1*, and *PTEN* are frequently deleted (2), and these genes could drive tumorigenesis by haploinsufficiency, although one study found that the LOH state of *RB1* is not associated with prognosis (12). *CDKN2A*, another tumor suppressor often affected in osteosarcoma, also shows hemizygous losses which may have a role in tumorigenesis, although small homozygous deletions in this gene are also seen (13). *LSAMP* is frequently focally deleted in osteosarcoma, and may have a role as haploinsufficient tumor suppressor (6–8). Copy neutral LOH (CN-LOH), which is LOH without change in copy number, may play a role in tumorigenesis as well, as is for example shown in hematological malignancies (14). In a study of osteosarcoma samples on Affymetrix 10 K 2.0 SNP arrays, it was reported that 28% of LOH events result from CN-LOH (6). Regions of LOH accompanied by gains have not yet been discussed in high-grade osteosarcoma, but have been described in other cancer types, *e.g.* in lung cancer (15) and triple-negative breast cancer (16). Tumor cells could in theory select for a region of amplified LOH in case a mutated tumor suppressor with a gain-of-function or partial dominant negative function is affected, with deletion of the wild-type gene and amplification and overexpression of the mutated gene. Another advantage for the tumor of stretches of LOH accompanied by gains is that tumor suppressors with inactivating mutations and oncogenes can have tumor-promoting activities at the same time (17).

In the present study, we analyzed high-throughput SNP data of osteosarcoma pretreatment biopsies, and detected that LOH is often accompanied by copy number gains. By paired integrative analysis of LOH, copy number gain, and gene expression of osteosarcoma biopsy and cell line data, we identified 29 candidate driver genes, exhibiting both LOH and copy number gains. Gene set enrichment on genes in regions of LOH accompanied by gains and overexpression of the transcript returned pathways important in tumorigenesis and genomic instability. We validated a selection of candidate tumor suppressor genes by Sanger sequencing and Fluorescence In Situ Hybridization (FISH). Mutation analysis of a selection of candidate tumor suppressors did not reveal any recurrent mutations. Further studies need to be performed to determine the role of drivers in these regions.

Methods

SNP microarray data analysis

Previously published Affymetrix Genome-Wide Human SNP 6.0 arrays were used for SNP microarray data analysis of high-grade osteosarcoma pretreatment biopsies (GEO accession number GSE33383) and of high-grade osteosarcoma cell lines (GEO accession number GSE36003). Microarray data preprocessing was performed as described in Pansuriya *et al.* (18). We previously described quality control and detection of aberrant regions (3). We used recommended settings for detection of LOH and allelic imbalance in Affymetrix SNP 6.0 data—a minimum LOH length of 500 kb, a homozygous frequency threshold of 95%, a homozygous value threshold of 0.8, and a heterozygous imbalance threshold of 0.4. High gain, gain, and losses were defined using $\log_2$ ratio cut-offs of 0.6, 0.2, and -0.2, respectively, which are slightly more conservative cut-offs than recommended by the software (0.6, 0.18, and -0.18 for Affymetrix SNP 6.0 data). We selected 29 patients for which gene expression microarray data were available, so that we could perform a paired integrative analysis. Of the 19 cell lines, 12 passed our quality control (143B, HAL, HOS, IOR/MOS, IOR/OS10, IOR/OS15, IOR/SARG, KPD, MG-63, MNNG-HOS, OSA, and SAOS-2). For all cell lines, expression data was available. Aberration frequency cut-offs of 5% (at least 2 samples out of 29) and of 15% (at least 2 samples out of 12) were used to detect recurrent regions in biopsies and in cell lines, respectively.

Genome-wide gene expression microarray data analysis

Genome-wide gene expression Illumina Human-6 v2.0 microarray data were previously published (GEO accession number GSE33383 for biopsies, GSE42351 for osteosarcoma cell lines). Microarray data processing and quality control in the statistical language R version 2.14 (19) were performed as described previously (20). Mesenchymal stem cells (MSCs, $n = 12$) and osteoblasts ($n = 3$) were used as control samples, as described by Kuijjer *et al.* (3) (GEO accession number GSE33383).

Paired integrative analyses

A detailed description of the paired integrative analysis can be found in Kuijjer *et al.* (3). For this study, we generated different binary files, including all genes that showed both LOH and copy number loss (1) or not (0), and LOH and copy number gain (1) or not (0). Gene expression data were normalized against average gene expression of the corresponding probes over all control samples (MSCs or osteoblasts). Different from our previous study, we included all genes—not only the subset of genes with significant differential expression. Genes were determined to be affected when frequencies of recurrent aberrations were higher than 5% and log fold changes > 1. Finally, only overlapping genes between

analyses with both control samples were considered of interest.

GO term enrichment

Gene set enrichment was performed using Bioconductor package *topGO* (21). Lists of significantly affected genes were compared with all genes eligible for the analysis. GO terms with Fisher's exact p-values < 0.001, as calculated by the *weight01* algorithm from *topGO*, were defined significant.

Other statistical analyses

Comparisons between the number of genes with both LOH and loss, and with both LOH and gain were performed using Pearson's chi-square test. All p-values were below 0.001.

Primer design

Primers for PCR amplifications were designed with a universal M13 tail in order to be able to use one set of universal primers for all sequencing reactions (M13 tail forward 5'-TGTAAAACGACGGCCAGT-3' and reverse 3'-CAGGAAACAGCTATGACC-5'). In order to first validate the LOH detected with the SNP arrays, we selected Affymetrix SNP probes for *XRCC6BP1*, *RASD1*, and *LLGL1* according to the population frequencies of the specific SNPs. Population frequency data from Affymetrix validation studies (`www.Affymetrix.com`) were assessed in order to select probes with frequencies close to an even distribution (50%/50%). We determined the number of SNPs to be evaluated for each gene by minimizing the chance of false positive homozygosity to less than 5%. Primers were designed up- and downstream of the SNPs using Primer 3 (`www.Primer3.com`) and the UCSC genome browser (`www.genome.ucsc.edu`). Primer sequences can be found in Additional Table 8.1. For mutation analysis primers were designed for the exons of *XRCCBP1*, *PLEKHO1*, and *TCC19* using Primer 3 (`www.Primer3.com`) and the UCSC genome browser (`www.genome.ucsc.edu`) (Additional Table 8.2).

Sanger sequencing

The procedure for PCR amplification is described in Rozeman *et al.* (22). The following PCR protocol was used: 5min at 95°C, 3 cycles of 10sec at 95°C and 10sec of 60°C, followed by 10sec of 72°C. Sequencing was performed at Macrogen (Macrogen Europe, Amsterdam, the Netherlands), Baseclear (Leiden, the Netherlands), and the Leiden Genome Technology Center (Leiden, the Netherlands). Sequences were analyzed with the Mutation Surveyor software, Softgenics (State College, PA) and Chromas software (Technelysium Pty Ltd, Helensvale, Australia).

Fluorescent in situ hybridization (FISH)

Metaphase preparations of osteosarcoma cell lines KPD and SAOS-2 and of a control cell line were obtained using colcemid as in Pajor *et al.* (23). The BAC probe for *XRCC6BP1* (BAC/Fosmid ID RP1160O7) was ordered from the BacPac Resource Centre at Children's Hospital Oakland Research Institute (Oakland, CA) and was labeled with biotin-16-2'-deoxyuridine-5'-triphosphate (Bio-16-dUTP) using a Nick translation method. The centromere probe for chromosome 12 (23) was labeled with digoxigenin-11-dUTP. For immunodetection, the following antibodies were used: streptavidin-Texas Red ($1:100$), mouse-anti-digoxin ($1:1,000$), goat-anti-streptavidin-bio ($1:100$), rabbit-anti-mouse-FITC ($1:1,000$), and goat-anti-rabbit-FITC ($1:100$). FISH was scored by counting red and green probes in 50 metaphase and 50 interphase nuclei per cell line.

Results

LOH and allelic imbalance in osteosarcoma biopsies

We set out to determine recurrent LOH in high-grade osteosarcoma. From SNP data analysis, we could demonstrate that LOH and allelic imbalance was detected less frequently than CN gains and losses (Figure 8.1). Recurrent LOH (frequency $> 5\%$) was detected



Figure 8.1: This figure shows the distribution of frequencies of LOH (blue) and allelic imbalance (purple) on a background of frequencies of copy number gains (green) and losses (red) for the 29 osteosarcoma biopsies.

for 9.4% of all analyzed genes. The highest percentage of recurrent LOH detected in

this dataset was 35%. Recurrent allelic imbalance was seen in 0.14% (23 genes), while recurrent total allelic loss was detected in 0.16% (25 genes, including *TP53* and *RB1*) of all analyzed genes, with highest recurrent frequencies of 7% and 55%, respectively.

LOH is often accompanied by copy number gains

On average, 0.05% of all genes show LOH and loss of DNA. A chi-square test demonstrated that LOH and loss occurred less frequent than expected (log odds -1.55, p-value $<$ 0.0001). Also copy neutral LOH occurred less frequent than expected (log odds -1.75, p-value $<$ 0.0001). Based on a comparison of the allelic ratio overview of the genome with CN gains and losses, LOH appears to often cooccur with gain at the other allele (Figure 8.1). On average, 1.10% of all genes show LOH and gain at the other allele in the same sample. Chi-square test verified that LOH accompanied by copy number gains indeed occurred more frequently than expected (log odds 2.44, p-value $<$ 0.0001).

Integration of LOH, gain, and differential expression

In order to identify genes present in regions of LOH and gain (LOH-gain) which also were differentially expressed and hence may have a tumor driving function, we performed paired integrative analyses of LOH-gain and expression in the dataset of osteosarcoma pretreatment biopsies. Paired integrative analysis of these biopsies as compared with MSCs resulted in 148 up- and 17 downregulated genes in combination with LOH-gain, while the analysis where osteoblasts were used as a control resulted in 135 up- and 9 downregulated genes in combination with LOH-gain. Of these affected genes, 114 upregulated and 5 downregulated genes overlapped between both analyses.

The same approach was taken for the analysis of high-grade osteosarcoma cell line data. This analysis returned 137 up- and 44 downregulated genes in combination with LOH-gain in osteosarcoma compared with MSCs, and 134 up- and 35 genes when compared with osteoblasts. In total, 97 upregulated genes and 20 downregulated genes overlapped. Of the 119 genes being over- and underexpressed together with LOH-gain in osteosarcoma biopsies as compared with MSCs and osteoblasts, 29 showed recurrent LOH-gain in combination with significant differential overexpression in both analyses of osteosarcoma cell lines (Figure 8.2A). No genes showing LOH and CN gain together with downregulation overlapped (Figure 8.2B).

Involvement of cell cycle pathways

GO term enrichment was performed on the 29 affected genes obtained with the analysis of biopsies and cell lines. This resulted in three significant GO terms—S-phase of mitotic cell cycle (GO:0000084), double-strand break repair via non-homologous end-joining (GO:0006303), and M/G1 transition of mitotic cell cycle (GO:0000216), including genes

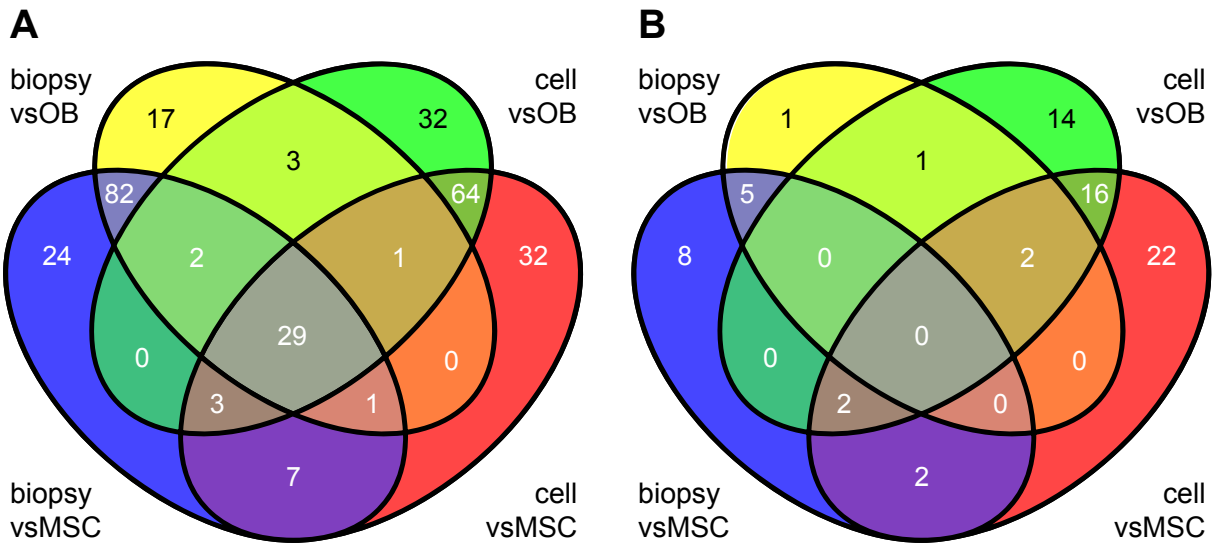


Figure 8.2: Depicted are the numbers of returned genes from the paired integrative analysis on LOH with gain and *A*, upregulation or *B*, downregulation, as compared with MSCs or osteoblasts (OB), in both osteosarcoma biopsies and cell lines.

such as *CDK4*, *MCM4*, and *XRCC6BP1* (Table 8.1). Literature review (`www.genecards.org`) indicated that 17/29 genes may have an oncogenic role and 2/29 a tumor suppressive role.

GO ID	Term	Ann	Sign	Exp	weight01	Genes affected
0000084	S phase of mitotic cell cycle	114	4	0.26	0.00013	*CDK4*, *MCM4*, *PSMB4*, *PSMD4*
0006303	double-strand break repair via NHEJ	10	2	0.02	0.00023	*PRKDC*, *XRCC6BP1*
0000216	M/G1 transition of mitotic cell cycle	67	3	0.16	0.00049	*MCM4*, *PSMB4*, *PSMD4*

Table 8.1: GO terms significantly enriched for genes with LOH-gain and overexpression. GO ID: GO-term ID, Term: GO term, Ann: number of annotated genes, Sig: number of significant genes, Exp: number of genes expected to be significant, weight01: p-value obtained with *weight01* algorithm, NHEJ: non-homologous end-joining.

Validation by Sanger sequencing

Tumor suppressor genes which are present in a region of LOH and gain may be particularly interesting, because a tumor cell could select for a mutant allele with a partial dominant negative or altered function. We therefore set out to identify mutations in tumor suppressor genes present in these regions of LOH and gain. Yet, false positive regions of LOH may be returned from SNP data analysis in regions of high CN amplification as a technical artifact. Hence, we validated regions of LOH and gain by Sanger sequencing. For validation, we selected the candidate tumor suppressor gene *XRCC6BP1*. In addition,

we chose to validate the gene *LLGL1*, which showed LOH and gain in the SNP data of cell line IOR/OS15 and allelic imbalance and gain in cell line IOR/SARG. We also validated *RASD1*, which showed recurrent LOH and gain, but downregulation when compared with osteoblasts. We validated LOH in the cell lines which showed LOH and gain in the particular genes in the SNP data analysis. The selected genes harbored homozygous as well as heterozygous SNPs when analyzed on normal blood donor DNA. Sequencing of the selected SNPs in and around *XRCC6BP1* in the cell lines KPD and SAOS-2, and of the SNPs in and around *RASD1* in cell lines 143B, HOS, and IOR/OS15 and the diagnostic biopsy L2613 revealed only homozygous SNPs. For *LLGL1*, we detected homozygosity in cell line IOR/OS15, but heterozygosity in cell line IOR/SARG, which was detected as allelic imbalance in SNP microarray data analysis. The probabilities for obtaining false positive results in the Sanger sequencing validation were 0.001, 0.037, and 0.019 for *XRCC6BP1*, *RASD1*, and *LLGL1*, respectively. These findings therefore confirm the detection of homozygosity by the SNP microarray data analysis.

Nature of the amplification

The copy number state of the *XRCC6BP1* locus in the affected cell lines was analyzed by FISH (Figure 8.3). For KPD, 42/50 metaphase cells had four copies of chromosome 12, of which two were negative for the *XRCC6BP1* probe. In addition, these cells showed more than ten homogeneous staining regions (HSR) for *XRCC6BP1*. 8/50 cells had only two copies of chromosome 12, of which one harbored *XRCC6BP1*, and showed 5–10 HSRs per cell. In 50/50 SAOS-2 metaphases, in contrast, we detected two copies of chromosome 12 harboring the *XRCC6BP1* locus, and two additional chromosomes without a chromosome 12 centromere, but with signals for *XRCC6BP1*. These results do not prove homozygosity of the locus, but do illustrate the different levels of amplification in the different cell lines. These amplifications identified with FISH corresponded to results from the SNP data analysis, as we detected a high gain in KPD and a normal gain in SAOS-2.

Mutation analysis of selected genes

We performed Sanger sequencing for the entire coding region of *XRCC6BP1*, one of the two candidate tumor suppressor genes with recurrent LOH-gain and overexpression, but did not identify any mutation in this gene, indicating that the wild-type allele is amplified in the osteosarcoma cell lines KPD and SAOS-2, which harbor the region of LOH and gain. We therefore expanded our list of candidate genes with tumor suppressor genes showing upregulation when compared with MSCs only (2/11 genes may have a tumor suppressive function), and when compared with osteoblasts only (3/5 genes may have a tumor suppressive function). The coding region of *PLEKHO1*, which showed overexpression only when compared with osteoblasts, was sequenced and analyzed for mutations in cell lines which showed this aberration—HOS, IOR/MOS, and IOR/SARG. No mutations

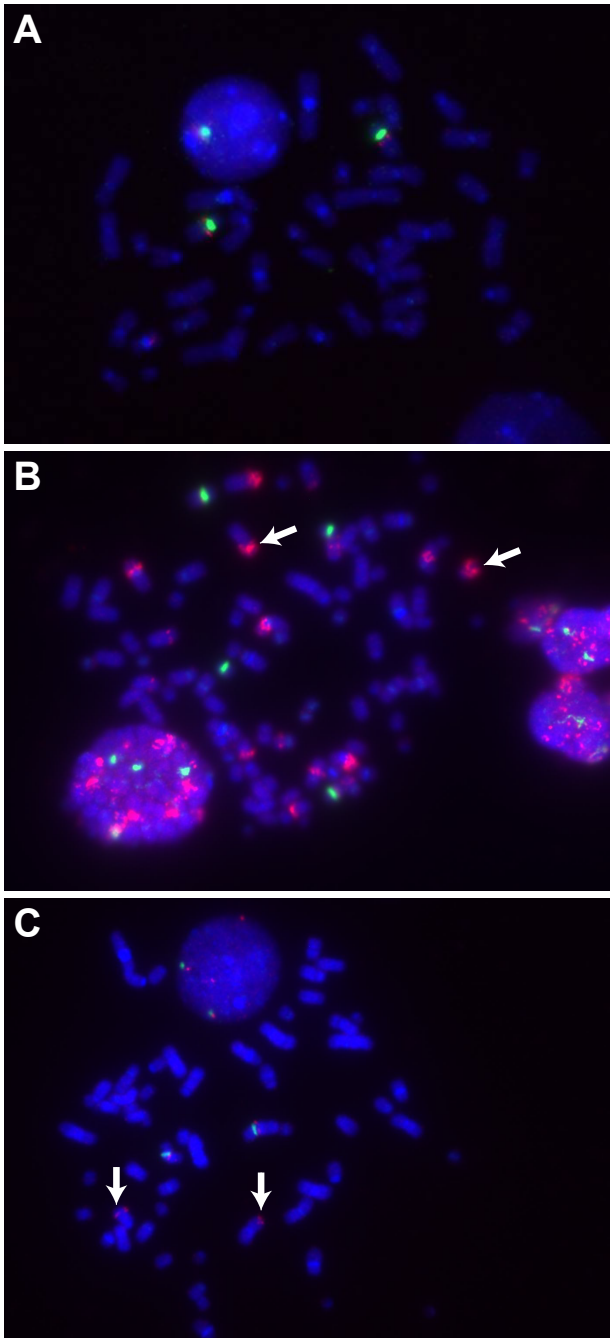


Figure 8.3: FISH depicting *A*, a control metaphase cell, *B*, a metaphase cell of KPD with arrows indicating examples of HSRs, and *C*, a metaphase cell of SAOS-2 with arrows indicating *XRCC6BP1* alleles not located on chromosome 12. Green: probe for the centromere of chromosome 12, red: BAC probe for *XRCC6BP1*.

could be identified, but we were not able to sequence the first exon of this gene. Mutation analysis of *TTC19*, which showed overexpression compared to MSCs, revealed a point mutation in exon 7 in the cell line IOR/OS10, but no mutations in the other affected cell lines (143B, HOS, IOR/OS15, IOR/SARG) and in two additionally analyzed diagnostic biopsies showing LOH and gain (L3437, L3469) were detected, indicating that the mutation in *TTC19* is not recurrent in osteosarcoma. The mutation, R274G, however, was predicted to be possibly damaging with a score of 0.752 (sensitivity of 0.85, specificity of 0.92) by PolyPhen-2 (24).

Discussion

By analysis of high-grade osteosarcoma high-throughput genomic data, we demonstrated that, in osteosarcoma, recurrent LOH happens less frequently than copy number aberrations such as gains and losses. Interestingly, we found that LOH was more often accompanied by copy number gains than expected by chance. Tumor suppressor genes showing overexpression in recurrent regions of gain and LOH may be drivers if these genes harbor mutations leading to a gain-of-function or partial dominant negative function. We thus screened for mutations in candidate tumor suppressor genes in these regions. Of the 29 genes that were recurrently affected in all comparisons, two may have a possible tumor suppressive role—*XRCC6BP1* and *PRKDC*. We performed mutation analysis for *XRCC6BP1*, or XRCC6 binding protein 1 / Ku70 binding protein 3, which is involved in non-homologous end-joining (NHEJ) of DNA double strand breaks (25). *XRCC6BP1* has been reported to be amplified and overexpressed in an alternatively spliced isoform in human gliomas, which may interfere with the normal function of the DNA-PK complex (26). In regions of LOH-gain in osteosarcoma cell lines, *XRCC6BP1* did not harbor any recurrent mutations. We did not screen for mutations in *PRKDC*, or DNA-PK catalytic subunit, a Ser/Thr kinase which also plays a role in NHEJ (27), because of its size (over 13kb). We did analyze two additional genes, *PLEKHO1* and *TTC19*, which showed recurrent LOH and gain, but which were overexpressed only in comparison with osteoblasts, or only in comparison with MSCs, respectively. No mutations in *PLEKHO1* (pleckstrin homology domain containing, family O member 1), a gene with a role in regulation of the actin cytoskeleton (28) with an inhibitory effect on PI3K/Akt signaling (29), were detected, but we were unable to sequence exon 1 of this gene. We did detect a mutation in *TTC19* (tetratricopeptide repeat domain 19), of which the protein is reported to be involved in oxidative phosphorylation in mitochondria (30), and which may play a role in cytokinesis as well (31). The mutation detected results in a arginine to glycine substitution at codon 274 of the TTC19 protein and was predicted as possibly damaging. However, the mutation was only found in 1/6 samples analyzed, and is therefore not recurrent. A shortcoming of these mutation analyses is that only exons were sequenced, and mutations in *e.g.* intronic regions may also affect the protein function, for example

by affecting alternative splicing. We thus cannot exclude that these genes harbor any recurrent mutations.

A weakness of this study is that we did not have paired control samples available for the SNP microarray data analysis. Using paired control samples helps in avoiding the false-positive regions of LOH which are detected when comparing tumor samples with an independent set of controls, because of patient-specific inherited segments of homozygosity (32). A second limitation in the analysis of these data is that high-grade osteosarcoma is extremely genomically unstable. Copy number aberrations that are returned by data analysis are relative changes against the background copy number state of the samples, and therefore true copy number states are not uncovered (33). Regions of copy number gain could, *e.g.* in a tetraploid background, consequently represent even higher gains than what is expected based on assumption of a near-diploid background. In the case of an unbalanced gain, the detection of the other allele may be low, which can lead to the detection of a false-positive region of LOH. Because of these considerations, we validated a selection of genes in regions of recurrent LOH-gain by Sanger sequencing. All regions we tested were indeed detected as homozygous for the all SNPs. However, in a highly amplified region, Sanger sequencing may also not be sensitive enough to detect the sequence of an allele of which only one copy is present. We therefore cannot conclude that these regions are actually homozygous, although for regions of low amplification this is probably the case.

FISH analysis of *XRCC6BP1* revealed four copies of chromosome 12 in both cell lines, of which two harbored the *XRCC6BP1* locus and two not. In addition, cell line KPD showed numerous homozygous staining regions. These results could be an indication of LOH and amplification of the other allele with intrachromosomal HSRs, but allele-specific FISH should be performed to clarify whether these are true cases of LOH. Nevertheless, FISH validated the copy number states that were detected by SNP microarray data analysis, as the gain detected in SAOS-2 was represented by four copies of the gene, and the high gain detected in KPD was represented by > 10 copies of the *XRCC6BP1* locus in FISH, thereby confirming the detection algorithm for copy number gain we used was appropriate for these samples. Chromothripsis, or chromosome scattering, is reported to be present in bone tumors with a frequency of at least 25% (34). In chromothripsis, part of the genome generally oscillates between two states, with the higher copy number state retaining heterozygosity and the lower copy number state showing LOH. The regions of LOH-gain we detected in the osteosarcoma SNP data could represent chromothripsis, since it would be possible that small oscillating regions were not detected due to the density of the probes targeting SNPs on the microarray and the detection algorithm. Such regions may be returned as larger regions of CN gain harboring LOH. A characteristic of chromothripsis is the presence of double minutes—small circular extra-chromosomal DNA fragments, which may be highly amplified in the tumor cell, and which frequently harbor oncogenes in cancer cells (35). The HSRs which were detected in cell line KPD

may represent chromosomal integration of double minutes, especially considering 17/29 genes detected by our analysis are possible oncogenes. It would thus be interesting to characterize whether the regions that we detected are recurrent HSRs or double minutes, and what the function of these oncogenes is in tumorigenesis of osteosarcoma.

References

[1] Raymond AK, Ayala AG, Knuutila S. Conventional osteosarcoma. In Fletcher C, Unni K, Mertens F, editors, Pathology and genetics of tumours of soft tissue and bone, 264–270. IARC Press, 2002.

[2] Cleton-Jansen AM, Buerger H, Hogendoorn PCW. Central high-grade osteosarcoma of bone: diagnostic and genetic considerations. *Current Diagnostic Pathology*, 11(6):390–399, 2005.

[3] Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, *et al.* Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes, Chromosomes and Cancer*, 51(7):696–706, 2012.

[4] Johnson-Pais TL, Nellissery MJ, Ammerman DG, Pathmanathan D, *et al.* Determination of a minimal region of loss of heterozygosity on chromosome 18q21.33 in osteosarcoma. *International Journal of Cancer*, 105(2):285–288, 2003.

[5] Deshpande AM, Akunowicz JD, Reveles XT, Patel BB, *et al.* PHC3, a component of the hPRC-H complex, associates with E2F6 during G0 and is lost in osteosarcoma tumors. *Oncogene*, 26(12):1714–1722, 2007.

[6] Yen CC, Chen WM, Chen TH, Chen WYK, *et al.* Identification of chromosomal aberrations associated with disease progression and a novel 3q13.31 deletion involving LSAMP gene in osteosarcoma. *International Journal of Oncology*, 35(4):775–788, 2009.

[7] Kresse SH, Ohnstad HO, Paulsen EB, Bjerkehagen B, *et al.* LSAMP, a novel candidate tumor suppressor gene in human osteosarcomas, identified by array comparative genomic hybridization. *Genes, Chromosomes and Cancer*, 48(8):679–693, 2009.

[8] Pasic I, Shlien A, Durbin AD, Stavropoulos DJ, *et al.* Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Research*, 70(1):160–171, 2010.

[9] Smida J, Baumhoer D, Rosemann M, Walch A, *et al.* Genomic alterations and allelic imbalances are strong prognostic predictors in osteosarcoma. *Clinical Cancer Research*, 16(16):4256–4267, 2010.

[10] Carter SL, Eklund AC, Kohane IS, Harris LN, *et al.* A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nature Genetics*, 38(9):1043–1048, 2006.

[11] Berger AH, Knudson AG, Pandolfi PP. A continuum model for tumour suppression. *Nature*, 476(7359):163–169, 2011.

[12] Heinsohn S, Evermann U, Stadt UZ, Bielack S, *et al.* Determination of the prognostic value of loss of heterozygosity at the retinoblastoma gene in osteosarcoma. *International Journal of Oncology*, 30(5):1205–1214, 2007.

[13] Mohseny AB, Tieken C, van der Velden PA, Szuhai K, *et al.* Small deletions but not methylation underlie CDKN2A/p16 loss of expression in conventional osteosarcoma. *Genes, Chromosomes and Cancer*, 49(12):1095–1103, 2010.

[14] O'Keefe C, McDevitt MA, Maciejewski JP. Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies. *Blood*, 115(14):2731–2739, 2010.

[15] Harris T, Pan Q, Sironi J, Lutz D, *et al.* Both gene amplification and allelic loss occur at 14q13.3 in lung cancer. *Clinical Cancer Research*, 17(4):690–699, 2011.

[16] Ha G, Roth A, Lai D, Bashashati A, *et al.* Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. *Genome research*, 22(10):1995–2007, 2012.

[17] Bacolod MD, Schemmann GS, Giardina SF, Paty P, *et al.* Emerging paradigms in cancer genetics: some important findings from high-density single nucleotide polymorphism array studies. *Cancer Research*, 69(3):723–727, 2009.

[18] Pansuriya TC, Oosting J, Krenács T, Taminiau AH, *et al.* Genome-wide analysis of Ollier disease: is it all in the genes? *Orphanet Journal of Rare Diseases*, 6(2), 2011.

[19] Team RDC. R: a language and environment for statistical computing, reference index version 2.14.0. R Foundation for Statistical Computing, 2011.

[20] Buddingh EP, Kuijjer ML, Duim RAJ, Bürger H, *et al.* Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. *Clinical Cancer Research*, 17(8):2110–2119, 2011.

[21] Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–1607, 2006.

[22] Rozeman LB, Hameetman L, Cleton-Jansen AM, Taminiau AHM, *et al.* Absence of IHH and retention of PTHrP signalling in enchondromas and central chondrosarcomas. *The Journal of Pathology*, 205(4):476–482, 2005.

[23] Pajor L, Szuhai K, Mehes G, Kosztolanyi G, *et al.* Combined metaphase, interphase cytogenetic, and flow cytometric analysis of DNA content of pediatric acute lymphoblastic leukemia. *Cytometry*, 34(2):87–94, 1998.

[24] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, *et al.* A method and server for predicting damaging missense mutations. *Nature Methods*, 7(4):248–249, 2010.

[25] Yang J, Yang D, Sun Y, Sun B, *et al.* Genetic amplification of the vascular endothelial growth factor (VEGF) pathway genes, including VEGFA, in human osteosarcoma. *Cancer*, 117(21):4925–4938, 2011.

[26] Fischer U, Hemmer D, Heckel D, Michel A, *et al.* KUB3 amplification and overexpression in human gliomas. *Glia*, 36(1):1–10, 2001.

[27] Chan DW, Chen BPC, Prithivirajsingh S, Kurimasa A, *et al.* Autophosphorylation of the DNA-dependent protein kinase catalytic subunit is required for rejoining of DNA double-strand breaks. *Genes & Development*, 16(18):2333–2338, 2002.

[28] Canton DA, Olsten MEK, Kim K, Doherty-Kirby A, *et al.* The pleckstrin homology domain-containing protein CKIP-1 is involved in regulation of cell morphology and the actin cytoskeleton and interaction with actin capping protein. *Molecular and Cellular Biology*, 25(9):3519–3534, 2005.

[29] Tokuda E, Fujita N, Oh-hara T, Sato S, *et al.* Casein kinase 2–interacting protein-1, a novel Akt pleckstrin homology domain-interacting protein, down-regulates PI3K/Akt signaling and suppresses tumor growth in vivo. *Cancer Research*, 67(20):9666–9676, 2007.

[30] Ghezzi D, Arzuffi P, Zordan M, Da Re C, *et al.* Mutations in TTC19 cause mitochondrial complex III deficiency and neurological impairment in humans and flies. *Nature Genetics*, 43(3):259–263, 2011.

[31] Sagona AP, Nezis IP, Pedersen NM, Liestøl K, *et al.* PtdIns(3)P controls cytokinesis through KIF13A-mediated recruitment of FYVE-CENT to the midbody. *Nature Cell Biology*, 12(4):362–371, 2010.

[32] Heinrichs S, Li C, Look AT. SNP array analysis in hematologic malignancies: avoiding false discoveries. *Blood*, 115(21):4157–4161, 2010.

[33] Gardina PJ, Lo KC, Lee W, Cowell JK, *et al.* Ploidy status and copy number aberrations in primary glioblastomas defined by integrated analysis of allelic ratios, signal ratios and loss of heterozygosity using 500K SNP Mapping Arrays. *BMC Genomics*, 9(1):489, 2008.

[34] Stephens PJ, Greenman CD, Fu B, Yang F, *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*, 144(1):27–40, 2011.

[35] Forment JV, Kaidi A, Jackson SP. Chromothripsis and cancer: causes and consequences of chromosome shattering. *Nature Reviews Cancer*, 12(10):663–670, 2012.

Additional Tables

Gene	SNP ID	Population frequencies (%)	Forward primer	Reverse primer
XRCC6BP1	SNP_A_8453712	45/55	GCAAGAACCTGTCTCTGAAAAA	GCTGCATAGTATTCCGTGGTG
	SNP_A_4194973	45/55	AGACAGTGGTGCAGCTGAGA	GACCACACGGGCTGTTTTTAT
	SNP_A_1846441	50/50	AACCCCAGAGAAAAACACCA	GTCCCAAGATGCATTGCTTT
	SNP_A_2291513	45/55	TCTGGCAGTAATGTGGTGGT	ATTGCTGCTAAGCCAAGGAC
	SNP_A_2063148	45/55	TCTGAGCCTAAAACCCAGGA	GCTGGGCAGCTGACTCTAAT
	SNP_A_8461882	41/59	AAGCAGGGAAACAGGCTACC	CACACCACAGCTGGCAGAATC
	SNP_A_8379343	48/52	GGGCTGATGTGGTCTAGGAG	CCCTGCACAGATGTCTACCC
	SNP_A_1818707	50/50	AGGTGGGAATATGAAGTTCAGTG	GAGCCACAAGGGTGAGAAGT
RASD1	SNP_A_8352439	46/54	TGAATCCTGCCTTTCCCATA	CATCCATATAGTTGCTGAAATGC
	SNP_A_8552537	47/53	CCATGAACCTTTTGGAAGGA	CTTCATGATGATGGAAGCTCTG
	SNP_A_2173899	25/75	CCTCCCTCCTGCTTCTTCTT	TGATCAGTGACAACCATCACA
	SNP_A_8383260	27/73	CAAGTGTCCATTGCCTGATG	GTGTCCGGCTTCTCTCACTC
	SNP_A_8307109	27/74	TTGATGCCATCTCTCAGCAC	AGTGTCCCCAGCAGTGTCTT
	SNP_A_1785268	44/56	TTCCAAGGAGCTGGAAGTTG	AGGCACCTTATCCCTTCTCC
	SNP_A_2186302	10/90	GCCTTGGTTGTCTCATTTTTG	GCCCTTACCAGTCCATTCCT
	SNP_A_8609757	25/75	GATTTGCAAGGTGTGAGACG	GTGGGAAATTTCAACCCAGA
	SNP_A_1953953	20/80	GTTAGTGGCCCACCCACTTT	CCAAATGAAGCCAGGGTCTA
LLGL1	SNP_A_8649983	30/70	AAGTTTGGCCTGAAGCTGTG	TCAGCTCCGTGTGTCTATGG
	SNP_A_8287520	40/60	GGGAAGGTCCTGGATTTGTT	CAGGCATGTGAGGTATGTGG
	SNP_A_8651023	20/80	CTGTGCATAGGCAGGGTTG	CTGGGTTGGTACTCCCCTTT
	SNP_A_8688092	40/60	"	"
	SNP_A_1919461	15/85	TCCCAAAGTGCTGGGATTAC	GCAAGGAAATGGCTGTGGTA
	SNP_A_8330336	40/60	ATCATACCACTGCCCTCCAG	CCAGACTCATGGATGCAGAA
	SNP_A_2043066	20/80	"	"
	SNP_A_4253141	20/80	"	"

Additional Table 8.1: Primer sequences to validate LOH.

Gene	Exon	Forward primer	Reverse primer
PLEKHO1	2	TGAAAACCTTTCCGAAGTGG	GCAGATGAGATGGGGGTAGA
	3	CCTCTCCTCAGGCTTCCTCT	TGCCTGGAAAGAAGGAAATG
	4	GGTGAGGCACCACCTCTAAA	TGGTGGAGGAGCGAGTAAAC
	5	TGTCCAGTAAATCCCCTTGC	CCTAATGGGCGCTGAATAAA
	6.1	CTGATGACAGGTTCCCCACT	AAGGTCGGGAGAGACTGCTT
	6.2	CTGAGAGCTTTCGGGTTGAC	TCCAATTCGATGATGCCTCT
	6.3	AGGTTCAGGGACTGGGAGAT	TACGAGGGGCATATGGAAAG
XRCC6BP1	1	CGGGAGGGAGGTTACCTTT	CAGACCCATTCTGTGGAACC
	2	CGCCTCAACCTCCTAAAGTG	GTTTTCAGCAGCCAGACCTC
	3	TATGGTGGAGGGTCTCCTGA	GCATGTGGAAGATGCTCAAA
	4	TCACTGAACTTCTTTTATTTTGGTG	CCGAAATTCAAGACTAAGGTAGAA
	5	GAGCATGAGCGTTTATTTCTTTT	ACACTCTGGAGGGGAAGTGA
	6.1	GAGCCTCATACTTTTCCTTCTTTTT	TGCCTTGGAGTTTAAAGCAG
	6.2	AAACAGAGAAGACTGTGATTCTAGC	CAACAGCTCAATAAGTATCCTACAATG
TTC19	1.1	CAACTGCGCTGTACCGTAAAT	CAGGATCCTCCACAGGTAGG
	1.2	GGCAACACTACGGCCATC	AGCTCAGGAGCCGGAACAT
	1.3	AGGGCGAGACGGAGTGAC	GAAGGGGCTCTGAGGTCAT
	2	GATGACCTCAGAGCCCCTTC	TAGAGTCGGAAAAGCCTGGA
	3	CAGTTGGGATGTACAGTTGCAT	CCAACCTTCCTCATCAGTGG
	4	TTGAGGGTGAAAGCAAAAGG	TCCCTTGAAGCTACTCCTTCAT
	5	GGGGCCCAATTAAAAGAAAA	CTCCACCTTTCCTGACCAAA
	6	CCACCGTCAGTCTGGAAGTT	GACACCCAATTTCTGGGAGA
	7	CAACAAGAGCGAAACTCCATC	GAAAAGGCAATGCCCAGATA
	8	TGGGTCCTGGTAACAACCAT	GGACCATCTGCTGATCCTGT
	9	TTGGATGCACTCCACATTAAA	CTTGCCCTCCCTACATACCA
	10	CACCAGCTTGTCGCTTCATA	ATGCCCAGAAAACTCCAGTG

Additional Table 8.2: Primer sequences for mutation analysis.

Chapter 9

Concluding remarks and future perspectives

High-grade osteosarcoma is a primary bone tumor with complex genetic alterations, for which targeted therapy is lacking. The aim of this thesis was to use high-throughput molecular data analysis of high-grade osteosarcoma specimens and model systems, in order to learn more on osteosarcomagenesis and to find possible ways to inhibit this process. **Chapter 1** and **Chapter 2** give an introduction and literature review on microarray data analysis of high-grade osteosarcoma.

In **Chapter 3** we provide a rationale for the use of model systems for osteosarcoma. It describes differential expression for the clinical parameters sex, tumor location (femur, humerus, fibula/tibia), response to preoperative chemotherapy (poor responders, or Huvos grade 1–2, versus good responders, or Huvos grade 3–4), and histological subtype (osteoblastic, chondroblastic, fibroblastic osteosarcoma). Importantly, as we describe in a previous study performed by our group (1), no significantly differentially expressed genes were detected between poor and good responders to preoperative chemotherapy, even though a substantial amount of tumor samples was analyzed (see also Chapter 2). Several publications do report differences between poor and good responders, but used relatively small sample sizes, and did not apply correction for multiple testing. An analysis of gene expression profiles of the three described histological subtypes showed that these differed significantly. Sets of fibroblastic- and of chondroblastic osteosarcoma-specific genes were determined, and were enriched in genes with a role in cellular growth and proliferation and in the chondroid extracellular matrix, respectively. Using nearest shrunken centroids classification, an expression signature consisting of 24 probes that could predict for histological subtype was generated. This profile was validated on an independent dataset of osteosarcoma and control samples. Interestingly, this prediction profile was able to classify histological subtypes of the primary tumor from which the tested osteosarcoma xenografts and cell lines were derived, even though such material often lacks extracellular matrix. This implicates that the mRNA expression profiles of these model systems are representative for the primary tumor, and favor the use of osteosarcoma xenografts and cell lines in studying osteosarcoma biology. This is of particular importance given the rarity of this tumor as well as the difficulties to obtain adequate tissue.

Targets for treatment of high-grade osteosarcoma

Different comparative analyses of the various types of data that were available led to the discovery of a number of particular ways to target this tumor. In **Chapter 4**, we compared patients who developed metastases within five years with patients who did not develop metastases within this time frame. The list of significantly differentially expressed genes was enriched in macrophage-associated genes expressed by infiltrating cells (approximately 50% of all genes), which were all overexpressed in patients who did not develop metastases. Tumor-associated macrophages (TAM) of both the M1- (antitumor) and M2 (protumor)-type were quantified with IHC in additional cohorts. The total count

of M1- and M2-type macrophages was significantly correlated with a better overall survival. This is in contrast with most epithelial tumor types, which often show a correlation between (M2-type) infiltrating macrophages and poor survival. However, mesenchymal tumor cells may not need the guidance of the infiltrate to metastasize, and tumor-associated macrophages may have a more antitumorigenic role in osteosarcoma (2). Moreover, macrophages are plastic cells, and it could be that M2-type macrophages polarize towards M1-type macrophages after chemotherapy due to the release of danger signals by the dying tumor cells. The results of Chapter 4 provide a rationale for adjuvant treatment of high-grade osteosarcoma patients with macrophage-activating and recruiting agents, such as liposomal muramyl tripeptide phosphatidylethanolamine (L-MTP-PE). This drug has been previously shown to increase overall survival in canine and in human osteosarcoma (2, 3), although interpretation of the results obtained from the latter study has been difficult, due to the 2x2 factorial design; standard adjuvant chemotherapy treatment plus L-MTP-PE and/or ifosfamide. This study showed that two additional drugs did not show significant interaction (p-value = 0.101) and therefore the treatment arms were pooled. A significant difference was then found for overall survival, but not for event-free survival (EFS). In an unpooled analysis, EFS for patients treated with L-MTP-PE and ifosfamide was significantly improved when compared with patients treated with ifosfamide alone, but EFS arms of patients with only the standard adjuvant chemotherapy and patients who received L-MTP-PE without ifosfamide were not significantly different (3). Further testing of this drug in osteosarcoma is therefore necessary, and this will, in the near future, be initiated in a phase II study in patients with metastatic and/or relapsed osteosarcoma.

In addition to differences between high-grade osteosarcoma tumors with different clinical features, we studied common gene expression changes between sets of tumors and control samples. In **Chapter 5**, mRNA expression in osteosarcoma cell lines was compared with expression in osteosarcoma progenitors. Global pathway analyses pointed to differences in mRNA expression of the IGF1R pathway. Specifically genes negatively regulating this pathway upstream the IGF1 receptor showed downregulation in osteosarcoma, of the highest degree (*i.e.* the highest negative fold changes in the dataset). We therefore hypothesized that this pathway can be inhibited at the receptor level, and that this may inhibit growth of these tumors. Osteosarcoma cell lines were treated with a dual kinase inhibitor OSI-906, which inhibits both the insulin receptor (IR) and IGF1R, as IR can take over downstream signaling in case IGF1R is blocked, thereby inducing resistance to single IGF1R targeting (4, 5). Inhibition with OSI-906 resulted in an inhibition of proliferation of 3/4 osteosarcoma cells, and may therefore be a promising drug for treatment in addition to adjuvant chemotherapy. Other pathways with a role in bone development, namely canonical Wnt signaling (6) and TGFβ/BMP signaling (7), have been reported to play a role in osteosarcomagenesis. Because of the role of IGF1R signaling in bone development and growth, it is not surprising that this pathway is deregulated in osteosarcoma. Notably, in a recent case–parent study, two SNPs in the growth hormone

(GH)/IGF1 pathway (in *IGF2R* and *IGFALS*) were significantly associated with osteosarcoma incidence (8). Interestingly, one of the very few genes which were overexpressed in patients developing metastases within 5 years (Chapter 4) was the growth hormone receptor (GHR), which was also frequently amplified (in 34% of all samples). IGF1 synthesis is largely dependent on growth hormone signaling (9), and an association between osteosarcoma and height/growth has been reported. The specific roles of the GH/IGF1 axis in osteosarcoma tumor growth and metastasis remain to be elucidated. The osteosarcoma cell line panel shows a variable expression of *GHR*, and includes four cell lines with high expression of *GHR*. These cell lines could be utilized to further experimentally examine this pathway in osteosarcoma.

Chapter 6 described Ser/Thr kinome profiling analysis of two osteosarcoma cell lines using a peptide microarray. Although it is not yet possible to directly infer what kinase caused differential phosphorylation of the identified peptides, by pathway analysis we detected hyperphosphorylation directly downstream of Akt, pointing to active PI3K/Akt signaling. We treated osteosarcoma cells with MK-2206, an Akt inhibitor, which inhibited proliferation of 2 out of 3 cell lines. Inhibition of the PI3K/Akt signaling pathway may therefore also be a possible target for treatment of these tumors.

The effects of IGF1R and Akt inhibitors on osteosarcoma cell proliferation should be studied further, in order to determine why these cells stop proliferating after treatment (this may for example be ascribed to induction of apoptosis). In addition, the drugs need to be tested in combination with chemotherapy, in order to check for synergy and also in order to rule out toxicity of a combined treatment, as targeted treatment using a single drug against these signal transduction pathways will probably not be able eliminate all osteosarcoma tumor cells, and patients may develop resistance to targeted treatment. In addition, if targeted treatments will be used to treat osteosarcoma patients, the genomic and mutational status of associated pathway players should be determined, because patients with downstream aberrations may be insensitive to treatment, as was shown in Chapter 6 for the 143B cell line, which is insensitive to Akt inhibitor MK-2206, most probably due to its oncogenic transformation of *KRAS*.

Integrative analysis and genomic instability

A main finding of this thesis regards genomic instability, which appears to be affected in all data types studied—the genome, transcriptome, and kinome. In **Chapter 6**, an integrative analysis shows that pathways with a role in genomic stability are enriched for overexpressed genes, as well as for hyperphosphorylation of peptides, implying that not only gene expression, but also kinase activity is deregulated in these pathways.

In addition to complementing two different data types as was performed in Chapter 6, integrative analysis can also be performed by applying intersections of both data types. This is useful in the analysis of copy number and gene expression data, because differently

from kinase activity and transcript expression, the influence of DNA copy number of a specific gene on the mRNA expression levels of that particular gene is more direct than the influence of kinase activity on gene expression, since kinases usually act quite far upstream of transcription factors in a specific pathway. Different methods for performing integrative analyses on gene expression and copy number data exist, as is described in **Chapter 7** of this thesis. We tested the performance of two of such methods—a non-paired and a paired analysis—on the high-grade osteosarcoma dataset. In the nonpaired analysis, genes showing significant differential expression as compared with the control samples were returned when present in recurrent regions with copy number alterations that contained a higher number of significantly differentially expressed than expected by chance. For the paired analysis, a new approach was developed in statistical language R. This method determined cooccurrence of copy number changes and significant differential expression. A comparison of both methods on osteosarcoma data illustrated that the paired analysis returned more genes with biological relevance over a larger number of regions, even though fewer samples were used for this analysis, since complete RNA expression–DNA copy number pairs were available for 29/32 cases. By using a conservative approach and by combining the results of different paired analyses, we identified 31 candidate osteosarcoma drivers with high frequency of occurrence and significant differences in expression. While most of these genes were not yet reported in osteosarcoma, more than two-thirds of the genes have been described to play a role in cancer. A large number of our candidate genes had a role in cell cycle regulation, stressing the possible role of genomic instability in driving osteosarcoma progression. This was further evaluated by calculating genomic instability scores, which showed that higher genomic instability correlated with poorer metastasis-free survival. In addition, a negative correlation between the total amount of copy number aberrations and metastasis-free survival was detected. We determined correlation with metastasis-free survival, and not overall survival, because of the limited follow-up available. Metastasis-free survival, however, highly correlates with overall survival, as only a small percentage of patients with resectable metastases survive.

Some issues can be raised with regard to the selection of the method we applied to identify candidate driver genes. For determining recurrent copy number aberrations, we used a cut-off for frequency. This will result in the detection of mostly broad events (*e.g.* the amplification of an entire chromosome arm). Focal events may be detected as well, but this method of analysis does not directly pinpoint specific targets (alterations with selective benefits) of recurrent copy number aberrations. Focal events are most often determined by the identification of the minimal common region of overlap of a copy number aberration, but this approach is prone to misidentification of the driver gene, especially when the recurrence frequency of the driver event is low (10), which is one of the reasons why we did not take this approach. Even though broad events do have a higher prevalence in most cancer types (low-level aberrations affecting a chromosome arm or an entire chromosome), determining focal alterations may have great power to identify

important genes in cancer (11). Importantly to note is also that both types of events may have different biological consequences. A method exist which can distinguish broad and focal events from each other and from background (*i.e.* passenger) events (GISTIC, Beroukhim *et al.* (12)). It would be therefore interesting to perform this method to the osteosarcoma data set, in order to also identify significant recurrent focal aberrations. Results obtained with this analysis could return less frequently occurring aberrations, which are specifically selected for in a subset of the tumors. Using our integrative method, which does not filter out passenger copy number alterations statistically (as is done in GISTIC), but which integrates copy number with expression data, we were able to identify those significantly differentially expressed genes of which a large part (at least 35%) of all tumors could be explained by an underlying copy number aberration. This approach gives us more information on osteosarcomagenesis in general (*i.e.* genes are affected in a high percentage of osteosarcoma samples). In addition to our list of new candidate osteosarcoma drivers, determining significant focal copy number events could provide us with some additional possible targets for treatment of a subset of osteosarcoma patients.

Candidate drivers need to be validated in an experimental setting. For the detected amplified and overexpressed genes this can be done by shRNA studies, but the effects of deleted tumor suppressors are more challenging to validate, especially because affecting a single gene will probably not be sufficient to stop tumor growth in cells with large amounts of aberrations. Because the majority of osteosarcomas do not have a known benign or less malignant precursor lesion, it is difficult to study tumor evolution, and to discriminate between early and late events in tumorigenesis of osteosarcoma. The detection of early drivers is important—it will reveal the first steps a normal cell takes in order to become tumorigenic, and these findings may be used in for example diagnostics. Genomic instability appears to play a major role in osteosarcoma, and in at least 25% (13), but probably in total approximately 50% of all high-grade osteosarcomas, this can be explained by chromothripsis. However, how chromothripsis exactly occurs, and what happens in the other half of osteosarcomas is yet unknown. By using the right model systems, these mechanisms can be studied. This is for example ongoing in studies which make use of the injection of different passages of transformed MSCs in mice and zebrafish (14, 15) upon which their tumorigenicity can be assessed. Furthermore, conditional transgenic mouse models are useful tools to follow osteosarcomagenesis from the normal cell to the fully malignant tumor.

Finally, in **Chapter 8**, loss-of-heterozygosity (LOH) calls were integrated with copy number calls and expression. This approach identified a high cooccurrence of LOH and copy number gains. Such regions may harbor mutated tumor suppressor genes. Because the detected LOH may have been a technical artifact, we validated the LOH of a subset of tumor suppressors by Sanger sequencing and fluorescence in situ hybridization (FISH). Sanger sequencing may not be sensitive enough to pick up LOH in regions of high amplification, but using FISH we confirmed regions of low level gains. In these regions,

the LOH which was detected with both the SNP microarray and Sanger sequencing was probably not an artifact. Mutation analysis of a subset of genes did, however, not detect any recurrent mutations. In order to improve the accuracy of mapping regions of LOH in osteosarcoma, an analysis which includes paired tumor–control samples should be performed.

FISH analysis detected homozygous staining regions (HSRs) in samples with high level gains. Oncogenes may be present in HSRs, and could be associated with chromothripsis. It will therefore be of interest to perform FISH on these regions of high amplification, to validate whether these oncogenes are indeed highly amplified. Targeting a specific oncogene that is highly amplified could be beneficial in osteosarcoma, as the tumor may be addicted to such an oncogene. In osteosarcoma, screening with high-throughput methods for such events may detect possible patient-specific treatment options.

Summary

In summary, by high-throughput data analysis of pretreatment biopsies of a relatively large, homogeneous cohort of osteosarcoma patients, which was collected as a collaborative effort by EuroBoNeT, we discovered a protective role of macrophages against the development of metastases. In addition, the IR/IGF1R and PI3K/Akt signaling pathways were discovered as potential targets for treatment. By integrative genomic analyses, the genomic complexity of this tumor was confirmed, and a correlation of genomic complexity with metastasis-free survival was identified. A conservative integrative approach to filter out passenger genes from driving events resulted in a list of mostly new, highly frequent candidate drivers in osteosarcoma. Most convincingly, genes playing a role in the maintenance of genomic stability have a considerable driving role in osteosarcomagenesis, as these pathways were affected in all three data types studied (mRNA expression, copy number data, and the kinome screen). Figure 9.1 summarizes the results obtained from these studies. This thesis provides the first steps in unraveling the genomic and transcriptomic landscape of this highly genomically unstable tumor. The research of osteosarcoma genomics at an even higher resolution (Next Generation Sequencing) will, together with the proposed future studies discussed in this chapter, help to better understand this highly genomically unstable tumor, and will provide indispensable knowledge on cancer evolution, diagnostics, prognostics, and targeted therapies.

References

[1] Cleton-Jansen AM, Anninga JK, Briaire-de Bruijn IH, Romeo S, *et al.* Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways. *British Journal of Cancer*, 101(11):1909–1918, 2009.

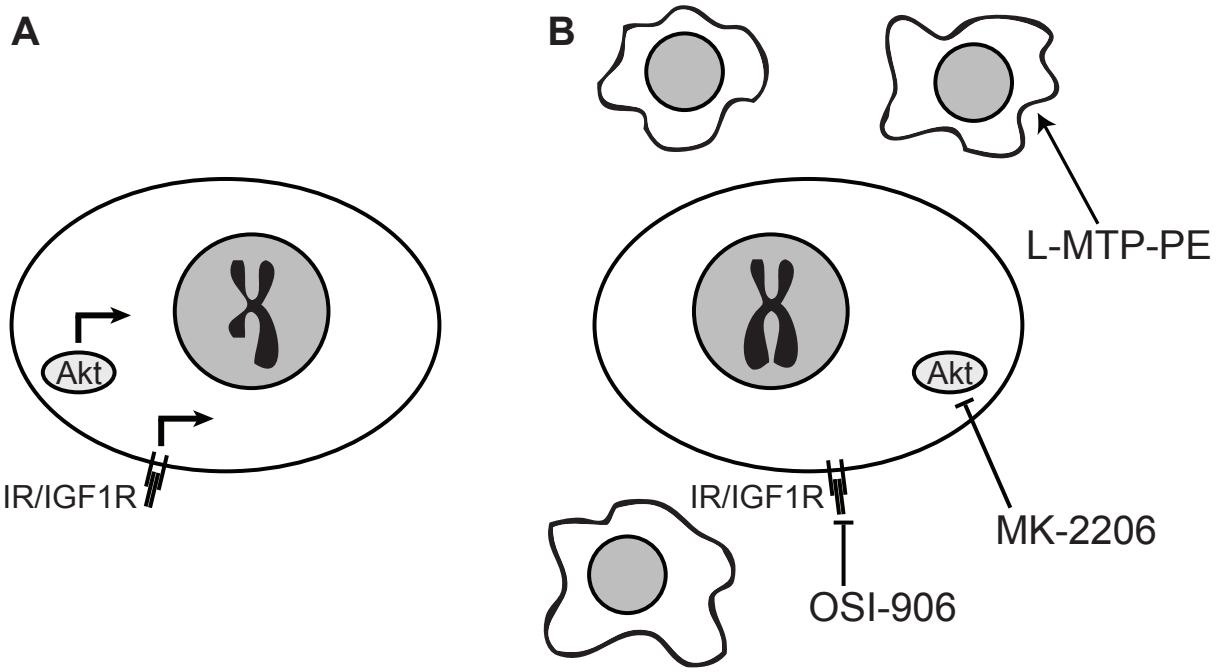


Figure 9.1: Schematic overview of the results described in this thesis. *A*, Genomic instability and low macrophage count in the primary tumor are associated with poor prognosis and IGF1R and Akt signaling pathways are active in osteosarcoma. *B*, Low genomic instability scores are associated with better prognosis. Possible ways to intervene tumor progression are activation of macrophages with L-MTP-PE and inhibition of IR/IGF1R and Akt signaling with OSI-906 and MK-2206, respectively.

[2] Cleton-Jansen AM, Buddingh EP, Lankester AC. Immunotherapy: is it different for sarcomas? *Oncoimmunology*, 1(2):255–257, 2012.

[3] Kager L, Pötschger U, Bielack S. Review of mifamurtide in the treatment of patients with osteosarcoma. *Therapeutics and Clinical Risk Management*, 6:279–286, 2010.

[4] Fulzele K, DiGirolamo DJ, Liu Z, Xu J, *et al.* Disruption of the insulin-like growth factor type 1 receptor in osteoblasts enhances insulin signaling and action. *Journal of Biological Chemistry*, 282(35):25649–25658, 2007.

[5] Garofalo C, Manara MC, Nicoletti G, Marino MT, *et al.* Efficacy of and resistance to anti-IGF-1R therapies in Ewing's sarcoma is dependent on insulin receptor signaling. *Oncogene*, 30(24):2730–2740, 2011.

[6] Cai Y, Mohseny AB, Karperien M, Hogendoorn PCW, *et al.* Inactive Wnt/β-catenin pathway in conventional high-grade osteosarcoma. *The Journal of Pathology*, 220(1):24–33, 2010.

[7] Mohseny AB, Cai Y, Kuijjer ML, Xiao W, *et al.* The activities of Smad and Gli mediated signalling pathways in high-grade conventional osteosarcoma. *European Journal of Cancer*, 48(18):3429–3438, 2012.

[8] Musselman JRB, Bergemann TL, Ross JA, Sklar C, *et al.* Case-parent analysis of variation in pubertal hormone genes and pediatric osteosarcoma: a Children's Oncology Group (COG) study. *International Journal of Molecular Epidemiology and Genetics*, 3(4):286, 2012.

[9] Khandwala HM, McCutcheon IE, Flyvbjerg A, Friend KE. The effects of insulin-like growth factors on tumorigenesis and neoplastic growth. *Endocrine Reviews*, 21(3):215–244, 2000.

[10] Mermel CH, Schumacher SE, Hill B, Meyerson ML, *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biology*, 12(4):R41, 2011.

[11] Beroukhim R, Mermel CH, Porter D, Wei G, *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature*, 463(7283):899–905, 2010.

[12] Beroukhim R, Getz G, Nghiemphu L, Barretina J, *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proceedings of the National Academy of Sciences*, 104(50):20007–20012, 2007.

[13] Stephens PJ, Greenman CD, Fu B, Yang F, *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*, 144(1):27–40, 2011.

[14] Mohseny AB, Hogendoorn PCW, Cleton-Jansen AM. Osteosarcoma models: from cell lines to zebrafish. *Sarcoma*, 2012:417271, 2012.

[15] Mohseny AB, Xiao W, Carvalho R, Spaink HP, *et al.* An osteosarcoma zebrafish model implicates Mmp-19 and Ets-1 as well as reduced host immune response in angiogenesis and migration. *The Journal of Pathology*, 227(2):245–253, 2012.

Chapter 10

Nederlandse samenvatting

Het hooggradig osteosarcoom is een maligne primaire bottumor, welke met name bij adolescenten en jonge volwassenen voorkomt op de plaats waar tijdens de pubertijd snelle botgroei plaatsvindt. Het is een zeer agressieve tumor, welke in 45% van de patiënten uitzaait, meestal naar de longen. De 5-jaars overleving van osteosarcoom patiënten is ongeveer 60–70%. Behandeling van het hooggradig osteosarcoom bestaat uit chemotherapie en operatieve verwijdering van de tumor. Een gerichte behandeling tegen specifiek osteosarcoom cellen, zoals dit bijvoorbeeld bestaat in de vorm van tamoxifen tegen oestrogeenreceptor-positieve borsttumoren, bestaat niet. Osteosarcoom tumor cellen hebben vele en complexe afwijkingen in het DNA. In dit proefschrift is zogenaamde high-throughput moleculaire data analyse gebruikt, om genoomwijd het osteosarcoom op verschillende niveaus, zoals DNA en mRNA, te kunnen bestuderen, met als doel meer over deze tumor te weten te komen en eventuele gerichte behandelingen tegen het osteosarcoom te kunnen identificeren. In de inleidende hoofdstukken van dit proefschrift, **hoofdstuk 1** en **hoofdstuk 2**, worden de verschillende microarray technieken—SNP-, genexpressie- en kinoomprofilering—besproken die gebruikt zijn in dit proefschrift, en wordt in een literatuuronderzoek een samenvatting gegeven van de tot nu toe gepubliceerde studies waarin dit soort technieken gebruikt zijn om het osteosarcoom te bestuderen.

Hoofdstuk 3 betreft mRNA expressie data analyse van pre-operatieve osteosarcoom biopten en modellen van het osteosarcoom, zoals cellijnen en diermodellen. Dit hoofdstuk beschrijft differentiële expressie tussen osteosarcoom biopten van verschillende groepen patiënten, bijvoorbeeld geslacht, de locatie van de primaire tumor in het lichaam van de patiënt, de reactie op de pre-operatieve chemotherapie en het histologische subtype van de tumor. Een opmerkelijke bevinding is dat er geen verschil in expressie bestaat tussen biopten van patiënten met een goede of slechte reactie op pre-operatieve chemotherapie. In tegenstelling tot een aantal publicaties waarin wel verschillen in genexpressie beschreven worden, is in onze studie gecorrigeerd voor herhaaldelijk testen—een statistische methode die toegepast moet worden als meerdere hypotheses worden getest, zoals het geval is bij het analyseren van microarray data. De meest voorkomende histologische

subtypes van het conventionele osteosarcoom—osteoblastisch, chondroblastisch en fibroblastisch osteosarcoom—vertoonden onderling verschillende genexpressieprofielen. Het expressieprofiel van het fibroblastaire osteosarcoom was verrijkt met genen die een rol spelen bij groei en proliferatie, terwijl het profiel specifiek voor het chondroblastaire osteosarcoom verrijkt was met genen die een rol spelen bij de chondroïde extracellulaire matrix van deze tumorcellen. Met behulp van een classificatiemethode werd een profiel van 24 probes (die met bepaalde genen corresponderen) ontwikkeld, welke het histologische subtype van de pre-operatieve biopten kon bepalen. Dit profiel werd vervolgens toegepast op genexpressie data verkregen uit osteosarcoomcellen en diermodellen en kon de histologische subtypes van de originele tumor waaruit deze modellen waren ontstaan correct classificeren. Deze modellen hebben minder of geen extracellulaire matrix, de eigenschap waarop de verschillende histologische subtypes van het osteosarcoom van elkaar onderscheiden worden. Dit impliceert dat genexpressieprofielen van deze osteosarcoommodellen nog steeds representatief zijn voor de primaire tumor waaruit deze ontwikkeld zijn, en is daarom een beweegreden om deze modellen te gebruiken in onderzoek naar het osteosarcoom indien er niet genoeg primair materiaal beschikbaar is.

Gerichte therapieën tegen het hooggradig osteosarcoom

Met behulp van verschillende analyses van verscheidene datasets zijn een aantal specifieke manieren ontdekt om deze tumor te bestrijden. In **hoofdstuk 4** zijn genexpressieprofielen van twee groepen patiënten met elkaar vergeleken—patiënten welke binnen 5 jaar uitzaaiingen ontwikkelden en patiënten bij wie in een periode van 5 jaar geen uitzaaiingen gevonden werden. De lijst van genen die significant verschilden in expressie was verrijkt met macrofaag-geassocieerde genen, welke door tumor infiltrerende cellen tot expressie gebracht werden. Deze genen vertoonden allen overexpressie in de patiëntgroep zonder metastasen. Het totale aantal tumor-geassocieerde macrofagen van type M1 (anti-tumor) en type M2 (pro-tumor) associeerde in aanvullende cohorten met een betere prognose. De resultaten van hoofdstuk 4 verschaffen een reden om osteosarcoompatiënten naast chemotherapie tevens met macrofaag-activerende/aantrekkende middelen te behandelen. Een voorbeeld hiervan is het medicijn liposomal muramyl tripeptide phosphatidylethanolamine (L-MTP-PE), wat in een fase III trial de prognose van patiënten met osteosarcoom kon verbeteren.

Naast het testen van verschillen in genexpressie tussen groepen van hooggradig osteosarcoom biopten met verschillende klinische karakteristieken, is genexpressie van het osteosarcoom ook vergeleken met controles. In **hoofdstuk 5** zijn verschillen in genexpressie bepaald tussen osteosarcoomcellijnen en voorlopercellen van het osteosarcoom. Een globale pathway analyse wees op verschillen in de IGF1R signaaltransductieroute, welke een rol speelt bij botgroei. Negatieve regulatoren van de IGF1 receptor waren sterk downgereguleerd in de osteosarcoomcellen, wat zou kunnen duiden op een verhoogde activiteit van

deze signaalstransductieroute. Deze route werd daarom vervolgens in osteosarcoomcellen geïnhibeerd met kinaseremmer OSI-906, welke niet alleen IGF1R, maar ook de insuline receptor kan remmen, wat noodzakelijk is om resistentie tegen remming van IGF1R tegen te gaan. Inhibitie met OSI-906 resulteerde in verlaagde proliferatie in drie van de vier geteste osteosarcoomcellijnen. OSI-906 zou derhalve een veelbelovend geneesmiddel kunnen zijn voor de behandeling van het osteosarcoom, naast de gebruikelijke chemotherapie.

In **hoofdstuk 6** wordt Serine/Threonine kinoomprofilering van twee osteosarcoomcellijnen beschreven. Voor deze studie is een peptide microarray gebruikt, welke peptides bevat die gefosforyleerd kunnen worden door kinases die aanwezig zijn in de cellysaten. Met behulp van een pathway analyse werd ontdekt dat de PI3K/Akt signaaltransductieroute verrijkt was met hyperfosforylering van moleculen direct downstream van Akt. Deze signaaltransductieroute speelt een rol bij celdeling en celgroei en kan geremd worden met verschillende medicijnen. Osteosarcoomcellijnen werden behandeld met de Akt remmer MK-2206, welke proliferatie van twee van de drie geteste cellen kon remmen. Inhibitie van de PI3K/Akt signaaltransductieroute is dan ook een mogelijke manier om deze tumor gericht te kunnen behandelen.

Genomische instabiliteit

Een tweede bevinding van dit proefschrift betreft genomische instabiliteit, welke zijn stempel drukt op verschillende niveaus—DNA, mRNA en kinaseactiviteit—in de tumorcel. De geïntegreerde analyse beschreven in **hoofdstuk 6**, toegepast op signaaltransductieroutes van verschillende types data, laat zien dat signaaltransductieroutes die een rol spelen in het behouden van genomische stabiliteit verrijkt zijn in zowel overexpressie van genen als hyperfosforylering van peptiden. Geïntegreerde analyse kan ook uitgevoerd worden op genniveau, door naar genen te kijken die aangedaan zijn in de te bestuderen data types. Deze methode kan gebruikt worden bij het combineren van DNA en genexpressie data, omdat de hoeveelheid DNA een directe invloed heeft op de expressie van het betreffende gen. In **hoofdstuk 7** worden twee methoden voor geïntegreerde analyse besproken—de gepaarde en de ongepaarde geïntegreerde analyse—en toegepast op data van osteosarcoom biopten en cellijnen. In de ongepaarde analyse zijn alle genen meegenomen die significant differentiële expressie vertonen ten opzichte van de controle celculturen. Significant differentieel tot expressie komende genen die zich bevinden in gebieden met genomische veranderingen, welke een hoger aantal van zulke genen bevatten dan verwacht, zijn hier kandidaatgenen. In de gepaarde analyse wordt van alle genen die significant differentieel tot expressie komen bepaald of deze in hetzelfde weefsel of in dezelfde cellijn tevens in een gebied van amplificatie of deletie liggen. Dit zijn vervolgens kandidaat kanker-drijvende genen. Met gepaarde analyse werden meer kandidaatgenen gedetecteerd in de bestudeerde osteosarcoomdataset dan met de ongepaarde analyse. Een conservatieve aanpak waarin alleen kandidaatgenen meegenomen werden die zowel in biopten als cellijnen, en vergele-

ken met verschillende controle celculturen gevonden waren, resulteerde in een lijst van 31 osteosarcoom kandidaatgenen, welke in minstens 35% van alle osteosarcomen aangedaan waren. Het overgrote deel van deze kandidaatgenen was voorheen nog niet ontdekt in osteosarcoma, en meer dan twee derde van de genen speelt een mogelijke rol in kanker. Een groot aantal van deze kandidaatgenen is belangrijk voor reguleren van bijvoorbeeld de celcyclus, wat aangeeft dat het verlies van genomische stabiliteit een rol speelt in het osteosarcoom. Deze bevinding werd verder geëvalueerd door het berekenen van bepaalde genomische instabiliteit scores, waaruit bleek dat hogere genomische instabiliteit gecorreleerd was met slechtere prognose. Tevens werd een negatieve correlatie tussen het totale aantal DNA kopieveranderingen en prognose waargenomen.

Tenslotte is in **hoofdstuk 8** loss-of-heterozygosity (LOH) data geïntegreerd met DNA kopie en genexpressie data. LOH gaat veelal gepaard met DNA amplificatie. Zulke gebieden zouden tumor suppressor genen kunnen bevatten, welke door bijvoorbeeld een dominant negatieve mutatie voordeel kunnen hebben van de amplificatie. Door hoge amplificatie kan echter vals positieve LOH gedetecteerd worden, vandaar dat LOH voor een aantal tumor suppressor genen gevalideerd werd met behulp van Sanger sequencing en fluorescence in situ hybridization (FISH). Sanger sequencing detecteerde geen heterozygositeit in deze genen, maar deze methode is wellicht óók niet sensitief genoeg om heterozygositeit te kunnen detecteren in het geval van een hoge amplificatie van één allel. Met behulp van FISH werden zowel hoge amplificatie als lage amplificatieniveaus in verschillende samples weergegeven, welke overeenkwamen met de SNP data. LOH in lage amplificatieniveaus zijn waarschijnlijk geen artifact, en tumor suppressorgenen die in deze gebieden liggen zouden interessante mutatie kunnen bevatten. Mutatieanalyse van een aantal tumor suppressorgenen wees echter niet op mutaties in deze genen. Zogenaamde homozygous staining regions (HSRs) werden met behulp van de FISH techniek gedetecteerd in osteosarcoomcellen met hoge amplificatieniveaus. Deze hoog geamplificeerde gebieden zijn wellicht geassocieerd met chromothripsis (het uiteenvallen van chromosomen) en bevatten mogelijk oncogenen waartegen gerichte, tumor-specifieke therapieën gebruikt zouden kunnen worden.

Conclusie

In dit proefschrift is 'high-throughput' data analyse beschreven van een relatief grote microarray dataset bestaande uit osteosarcoom pre-operatieve biopten, cellijnen en xenotransplantaten, welke beschikbaar werd gesteld door het FP6 netwerk EuroBoNeT. Met behulp van data analyse is gevonden dat macrofaag-activerende middelen en IGF1R en Akt remmers mogelijke adjuvante therapieën kunnen zijn in de behandeling van het osteosarcoom. Geïntegreerde genomische analyses wezen op de genomische complexiteit van deze tumor, en de rol van genomische instabiliteit ten opzichte van agressiviteit van het osteosarcoom. Een conservatieve benadering om zogenaamde passagiergenen weg te filteren,

zodat genen die een belangrijke rol spelen bij tumorgenese gedetecteerd kunnen worden resulteerde in een lijst van 31 kandidaatgenen die frequent in het hooggradig osteosarcoom aangedaan zijn op DNA en expressieniveau, waaronder meerdere genen die een rol spelen bij het behouden van genomische stabiliteit. Met dit proefschrift zijn de eerste stappen voor het in kaart brengen van het genoom- en transcriptoomlandschap van het hooggradig osteosarcoom in gang gezet. Deze informatie is onmisbaar voor het begrijpen van deze zeer genomisch instabiele tumor en voor het ontwikkelen van diagnostische/prognostische methoden en nieuwe therapieën.

Curriculum Vitae

Marieke Lydia Kuijjer was born on February 5, 1982, in Zaanstad, the Netherlands. She attended preuniversity school (gymnasium) at the Sint Adelbert College in Wassenaar. After graduating with honors in 2000, she enrolled in the Bachelor's programme Mathematics and Statistics at Leiden University. She switched to Biomedical Sciences in 2003, and received her Propaedeutics with distinction. During the Bachelor's phase of her studies, Marieke followed an Erasmus exchange program at Karolinska Institutet in Stockholm, Sweden. For her Bachelor's thesis, she studied protein dynamics using confocal microscopy under the supervision of C.R. Jost, PhD, at the department of Molecular Cell Biology, Leiden University Medical Center. During her Master's programme, she received two scholarships to follow Master's courses at Karolinska Institutet. As a trainee at the department of Human Genetics, Leiden University Medical Center, she integrated microarray data of a panel of cancer cell lines under the supervision of J.M. Boer, PhD. For her Master's thesis, she studied Wnt signaling in bone formation at the department of Molecular Cell Biology of the same institute, under the supervision of D.J. de Gorter, PhD, and P. ten Dijke, PhD. Marieke received her Master of Science degree in December 2008 and started her PhD education in the same month. The results obtained during her PhD education are described in this thesis. In May 2013, Marieke started as a postdoctoral fellow in the laboratory of J. Quackenbush, PhD, at the department of Biostatistics and Computational Biology of the Dana-Farber Cancer Institute, Boston (MA), USA.

List of publications

- Pahl JHW, Santos SJ, Kuijjer ML, Boerman GH, Sand LGL, Szuhai K, Cleton-Jansen AM, Egeler RM, Bovée JVGM, Schilham MW, Lankester AC. Expression of the immune regulation antigen CD70 in osteosarcoma. *Submitted*

- Buddingh EP, Ruslan SEN, Reijnders CMA, Kuijjer ML, Roelofs H, Hogendoorn PCW, Egeler RM, Cleton-Jansen AM, Lankester AC. Mesenchymal stromal cells derived from healthy donors and osteosarcoma patients do not transform during long-term culture. *Submitted*

- Kansara E, Leong HS, Lin DM, Popkiss S, Pang P, Garsed DW, Walkley CR, Cullinane C, Ellul J, Haynes NM, Hicks R, Kuijjer ML, Cleton-Jansen AM, Hinds PW, Smyth MJ, Thomas DM. Senescence-related immunologic functions of the RB1 tumor suppressor in radiation-induced osteosarcoma. *Submitted*

- Kuijjer ML, van den Akker BEWM, Hilhorst R, Mommersteeg M, Buddingh EP, Serra M, Bürger H, Hogendoorn PCW, Cleton-Jansen AM. Kinome and mRNA expression profiling of high-grade osteosarcoma identifies genomic instability, and reveals Akt as potential target for treatment. *Submitted*

- de Vos van Steenwijk PJ, Ramwadhdoebe TH, Goedemans R, Doorduijn E, van den Ham JJ, Gorter A, van Hall T, Kuijjer ML, van Poelgeest MIE, van der Burg SH, Jordanova ES. Tumor infiltrating CD14 positive myeloid cells work side by side with T cells to prolong the survival in patients with cervical carcinoma. Accepted for publication in *International Journal of Cancer*

- Kuijjer ML, Peterse EFP, van den Akker BEWM, Briaire-de Bruijn IH, Serra M, Meza-Zepeda LA, Myklebost O, Hassan AB, Hogendoorn PCW, Cleton-Jansen AM. IR/IGF1R signaling as potential target for treatment of high-grade osteosarcoma. Accepted for publication in *BMC Cancer*

- Kuijjer ML, Cleton-Jansen AM, Hogendoorn PCW. Genome-wide analyses on high-grade osteosarcoma; making sense of a most genomically unstable tumor. *Review. International Journal of Cancer.* 2013 Feb 20;Epub

- Namløs HM, Meza-Zepeda LA, Barøy T, Østensen IHG, Kresse SH, Kuijjer ML, Serra M, Bürger H, Cleton-Jansen AM, Myklebost O. Modulation of the Osteosarcoma Expression Phenotype by MicroRNAs. *PloS One.* 2012;7(10):e48086

- Mohseny AB, Cai Y, Kuijjer ML, Xiao W, van den Akker B, de Andrea CE, Jacobs R, ten Dijke P, Hogendoorn PCW, Cleton-Jansen AM. The activities of Smad and Gli mediated signalling pathways in high-grade conventional osteosarcoma. *European Journal of Cancer.* 2012 Dec;48(18):3429–3438

- Lenos K, Grawenda AM, Lodder K, Kuijjer ML, Teunisse AF, Repapi E, Grochola LF, Bartel F, Hogendoorn PCW, Wuerl P, Taubert H, Cleton-Jansen AM, Bond GL, Jochemsen AG. Alternate splicing of the p53 inhibitor HDMX offers a superior prognostic biomarker than p53 mutation in human cancer. *Cancer Research.* 2012 Aug 15;72(16):4074–84

- Kuijjer ML, Rydbeck H, Kresse SH, Buddingh EP, Lid AB, Roelofs H, Bürger H, Myklebost O, Hogendoorn PCW, Meza-Zepeda LA, Cleton-Jansen AM. Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes, Chromosomes and Cancer.* 2012 Jul;51(7):696–706

- Pansuriya TC, van Eijk R, d'Adamo P, van Ruler MA, Kuijjer ML, Oosting J, Cleton-Jansen AM, van Oosterwijk JG, Verbeke SL, Meijer D, van Wezel T, Nord KH, Sangiorgi L, Toker B, Liegl-Atzwanger B, San-Julian M, Sciot R, Limaye N, Kindblom LG, Daugaard S, Godfraind C, Boon LM, Vikkula M, Kurek KC, Szuhai K, French PJ, Bovée JVGM. Somatic mosaic IDH1 and IDH2 mutations are associated with enchondroma and spindle cell hemangioma in Ollier disease and Maffucci syndrome. *Nature Genetics.* 2011 Nov 6;43(12):1256–1261

- Kuijjer ML, Namløs HM, Hauben EI, Machado I, Kresse SH, Serra M, Llombart-Bosch A, Hogendoorn PCW, Meza-Zepeda LA, Myklebost O, Cleton-Jansen AM. mRNA expression profiles of primary high-grade central osteosarcoma are preserved in cell lines and xenografts. *BMC Medical Genomics.* 2011 Sep 20;4:66

- Buddingh EP†, Kuijjer ML†, Duim RA, Bürger H, Agelopoulos K, Myklebost O, Serra M, Mertens F, Hogendoorn PCW, Lankester AC, Cleton-Jansen AM. Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. *Clinical Cancer Research.* 2011 Apr 15;17(8):2110–2119 †Shared first authorship

Acknowledgments

It is a great pleasure to give respect to those who made this thesis possible. I owe sincere gratitude to my promotor, prof. dr. P.C.W. Hogendoorn, and to my copromotor, dr. A.M. Cleton-Jansen. Pancras, I enjoyed working in your group, which has been an excellent environment to start my scientific career. Anne-Marie, your enthusiasm and knowledge were always stimulating, and I thank you for giving me freedom to find my own research direction. All colleagues from the department of Pathology are acknowledged, in particular the technicians, who provided me with excellent technical support. I am indebted to my roommates for their patience and support and especially to Cathelijn, Jolieke, Sara, and Wei for their friendship. I dearly thank Lisanne Vijfhuizen, Frauke Liebelt, and Elleke Peterse; I enjoyed to work with you and wish you all the best with your careers. Special thanks to all EuroBoNeT partners and to PamGene for the good collaborations. An honorable mention goes to my family and friends for their understanding and support. Finally, I thank Alessandro Marin for supporting me in every possible way, and for believing in me. *Ti amo.*