Genomic alterations and allelic imbalances are strong prognostic predictors in osteosarcoma

Jan Smida\textsuperscript{1,2§}, Daniel Baumhoer\textsuperscript{1,3,4§}, Michael Rosemann\textsuperscript{5,6}, Axel Walch\textsuperscript{5}, Stefan Bielack\textsuperscript{7}, Christopher Poremba\textsuperscript{8}, Klaus Remberger\textsuperscript{9}, Eberhard Korsching\textsuperscript{10}, Wolfram Scheurlen\textsuperscript{11}, Christian Dierkes\textsuperscript{12}, Stefan Burdach\textsuperscript{2}, Gernot Jundt\textsuperscript{3,4}, Michael J. Atkinson\textsuperscript{5,6}, and Michaela Nathrath\textsuperscript{1,2}

\textsuperscript{1}Clinical Cooperation Group Osteosarcoma, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Germany, \textsuperscript{2}Department of Pediatrics, Technische Universitaet Muenchen and Pediatric Oncology Center, Germany, \textsuperscript{3}Institute of Pathology and \textsuperscript{4}Bone Tumor Reference Center at the Institute of Pathology, University Hospital Basel, Switzerland, \textsuperscript{5}Institute of Pathology and \textsuperscript{6}Institute of Radiation Biology, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Germany, \textsuperscript{7}Klinikum Stuttgart Olgahospital, Pediatrics 5 - Oncology, Hematology, Immunology, Germany, \textsuperscript{8}Institute of Pathology, Heinrich-Heine-University, Duesseldorf, Germany and Center of Histopathology, Cytology and Molecular Diagnostics (CHCMD), Research Park Trier, Germany, \textsuperscript{9}Institute of Pathology, University of the Saarland, Homburg-Saar, Germany, \textsuperscript{10}Institute of Bioinformatics, University of Muenster, Germany, \textsuperscript{11}Cnopfsche Kinderklinik Nuremberg Children’s Hospital, Germany, and \textsuperscript{12}Institute of Pathology, Justus-Liebig-University, Giessen, Germany

\textsuperscript{§} The authors contributed equally to this work

Corresponding author / adress for reprint requests: Michaela Nathrath, Clinical Cooperation Group Osteosarcoma, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany, Phone: +49-89-3187-3797, Fax: +49-89-3187-3381, EMail: michaela.nathrath@lrz.tu-muenchen.de

Grant support: Jan Smida, Eberhard Korsching and Michaela Nathrath are members of the Translational Sarcoma Research Network supported by the BMBF. Michaela Nathrath is supported by the Wilhelm Sander-Stiftung and the Elterninitiative krebskranke Kinder München e.V. Daniel Baumhoer was supported by the Novartis Foundation, formerly Ciba-Geigy-Jubilee-Foundation.

Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Copyright © 2010 American Association for Cancer Research
Running title: Prognostic impact of genomic alterations in osteosarcoma

Key words: osteosarcoma, prognosis, SNP array, CDK4, MYC

Disclosure / Conflict of Interest: We declare that we have no conflict of interest.
Statement of Translational Relevance

Osteosarcoma is a highly aggressive neoplasm with 30-40% of patients dying due to progressive disease, despite of intense multidisciplinary treatment protocols. Up to now, the clinical course of patients can not be anticipated reliably at the time of initial diagnosis and generally the histologically assessed response to neoadjuvant chemotherapy after definitive surgery has to be awaited. Since up to 50% of osteosarcoma patients only poorly respond to neoadjuvant treatment, molecular predictors might help to stratify patients more accurately in distinct prognostic and therapeutic subgroups. In the present study, we propose that genomic alterations, determined by using a chromosomal alteration staging (CAS) system, are strong prognostic predictors in osteosarcoma that can be verified at initial biopsy and are superior to the histologic regression grading.
ABSTRACT

Purpose: Osteosarcoma, the most common primary malignant tumor of bone, is characterized by complex karyotypes with numerous structural and numerical alterations. Despite attempts to establish molecular prognostic markers at the time of diagnosis, the most accepted predictive factor remains the histological evaluation of necrosis after neoadjuvant chemotherapy. The present approach was carried out to search for genome-wide recurrent loss of heterozygosity (LOH) and copy number variations (CNV) that could have prognostic and therapeutic impact for osteosarcoma patients. Experimental Design: Pretherapeutic biopsy samples of 45 osteosarcoma patients were analyzed using Affymetrix 10K2 high-density SNP arrays. Numerical aberrations and allelic imbalances were correlated with the histologically assessed response to therapy and clinical follow-up. Results: The most frequent genomic alterations included amplifications of chromosome 6p21 (15.6%), 8q24 (15.6%, harboring MYC) and 12q14 (11.1%, harboring CDK4) as well as LOH of 10q21.1 (44.4%). All these aberrations and the total degree of heterozygosity of each tumor were significantly associated with an adverse outcome of patients and used to define a chromosomal alteration staging (CAS) system with a superior predictive potential compared to the histologic regression grading. Conclusions: Structural chromosomal alterations detected by SNP analysis provide a simple but robust parameter to anticipate response to chemotherapy. The proposed CAS system might therefore help to better predict the clinical course of osteosarcoma patients at the time of initial diagnosis and to adept neoadjuvant treatment in patients resistant to the current protocols.
INTRODUCTION

Due to advances in therapy over the last decades, long-term survival in osteosarcoma patients with localized disease has improved and now reaches about 65% with multimodal therapy. Nevertheless, 30-40% of the patients die because of tumor progression or relapse (1). Individualized intervention schemes based on the biological characteristics of each tumor have been shown to be successful in other pediatric cancer types e.g. in acute lymphatic leukaemia and in neuroblastoma. Also concerning osteosarcomas, several studies have been carried out to combine common histological parameters with genetic and molecular findings to predict the clinical outcome of patients (2-4). However, the histological response to neoadjuvant chemotherapy (“regression grade”) assessed after definitive surgery is still the gold standard concerning prognostic prediction (5, 6).

Osteosarcomas are characterized by highly complex karyotypes and a high frequency of chromosomal copy number changes (3, 7-12). It has already been shown in other tumors, such as gastrointestinal tract carcinomas, that a high complexity of chromosomal instability is correlated with an unfavorable outcome (13). In this study, we describe the use of Affymetrix single nucleotide polymorphism (SNP) arrays in a genome-wide high-resolution approach. Both, loss of heterozygosity (LOH) and variations in DNA copy numbers (CNV), were assayed in order to identify possible genomic fingerprints for the prediction of response to chemotherapy and prognosis. Furthermore, the osteosarcoma genomes were investigated for chromosomal gains and losses to identify candidate genes as potential therapeutic targets.
MATERIAL AND METHODS

Tissue samples and patients characteristics
Our series of pretherapeutic fresh frozen biopsy samples included 79 osteosarcomas which were numbered consecutively (OS1-OS79). Due to insufficient size or inadequate amount of vital tumor only 45 samples were selected for the present study including osteosarcoma of the extremities (n = 43) and the pelvis (n = 2). In nine cases, tumor and blood samples from the same patient were available. There were 25 males and 20 females with a mean age of 16.5 years (median 14 years; range 4–51 years); 10 patients had lung metastases at the time of initial diagnosis. All patients were treated between 1993 and 2007. Preoperative and postoperative chemotherapy was given according to the protocols of the Cooperative German-Austria-Swiss Osteosarcoma Study Group (reviewed and approved by the appropriate ethics committee) after informed consent. Response to chemotherapy was assessed according to the Salzer-Kuntschik (S-K) histological six-graded scale. SK-grades 1 to 3 (≤ 10% viable tumor cells) were classified as good responders and SK-grades 4 to 6 (> 10% viable tumor cells) as poor responders. Information on histological response was available for 44 patients. Follow-up data (0.6-12.7 years, mean 3.84 years) were available for all patients (Table 1). Additionally, blood samples from 5 healthy donors were used as hybridization quality controls.

Affymetrix 10K2 high-density SNP arrays
Genomic DNA was extracted using the QIAamp-Mini DNA extraction kit (Quiagen, Hilden, Germany) from frozen tissue and blood samples.
DNA samples were processed according to the standard GeneChip Mapping 10K (V2.0) Xba Assay protocol (Affymetrix Inc., Santa Clara, CA) as suggested by the manufacturer [http://www.affymetrix.com/products/arrays/index.affx]. In brief, 250 ng of DNA was digested with XbaI and ligated to the XbaI adaptor prior to PCR amplification using AmpliTaq Gold (Applied Biosystems, Foster City, CA) and primers that recognize the adapter sequence. PCR-amplified DNA was fragmented using DNAse I, end-labeled with a fluorescent tag and hybridized to the array. Hybridized arrays were processed with an Affymetrix Fluidics Station 450 and fluorescence signals were detected using the Affymetrix GeneChip Scanner 3000.
The primary hybridization data were processed using Affymetrix GCOS and GTYPE software, yielding the genotype and hybridization intensity for each individual SNP marker. In a second step the Affymetrix® GeneChip® Chromosome Copy Number Analysis Tool (CNAT and GTYPE Version 4.1) was applied to visualize the data corresponding to their chromosomal location and to convert the hybridization intensities into DNA copy numbers (CN). Copy number variations (CNV) with a score below 0.75 were classified as losses and with a score >3 as gains.

CNAT also detects genomic regions with unusual or long stretches of contiguous homozygote SNPs and assigns them a score for the likelihood of being caused by a tumor specific loss of heterozygosity (LOH). To distinguish tumor-specific somatic LOH from potential germline homozygosity, a cut-off LOH value of 5 was used. The value varies between 0.01 (no LOH = heterozygous SNP) and 75 (no heterozygous SNP on the entire chromosome arm) and has been explicitly discussed in Affymetrix White Papers. We validated this threshold by comparing regions that showed allelic losses in a small subset of tumor and blood samples. The cut-off LOH value of 5 had the best fit with the LOH data of matched samples and is equivalent to stretches of homozygosity greater than 4 Mbp. A homozygous interval of this length was not observed either in the investigated blood samples or in the genome of more than 100 healthy donors (control data provided by Affymetrix). All physical positions were determined using the NetAffx™ Analysis Center's UCSC Genomic Query Tool provided by the manufacturer and correspond to the NCBI sequencing database (version GRCh37, hg19).

All statistical analyses were done with STATISTICA 4.5 (StatSoft, Tulsa, OK, USA). The correlation analysis performed in this study used Spearman's rank order correlation coefficient. Differences between good and poor responders to chemotherapy were considered significant at $P < 0.05$ ($\alpha$-2 sides) using the Wilcoxon-Mann-Whitney rank test.
RESULTS

We collected frozen pretherapeutic biopsy specimens of osteosarcomas from 45 patients, all but one with a record of the subsequent response to chemotherapy. We used Affymetrix 10K SNP microarrays to assay DNA copy-number changes and LOH. A mean call rate of 88.5% (range 82.71% to 96.2%) resulted in a minimum of 8.200 SNP genotypes per case.

Heterozygosity and LOH

In all osteosarcoma samples we observed multiple regions of contiguous SNP homozygosity, ranging in length from 4.7 Mb to 156.6 Mb. These regions were classified by Affymetrix analysis software as areas with LOH (LOH-likelihood score 5 to 70, determined by Affymetrix CNT tool).

An appropriate method to estimate the whole-genome LOH frequency is to compare the frequency of SNP genotypes in a tumor with the population average. All SNPs used in the array are known to be heterozygous in at least 38% of the reference population (data by Affymetrix). The effective heterozygosity in the investigated blood samples as well as in control DNA provided by Affymetrix was on average 34.4%. The percentage of heterozygote markers in the osteosarcoma samples, however, dropped to an average of 27% in good responders and even to 24% in poor responders (Fig. 1A). Applying the LOH scoring algorithm as described above this converts into approximately 100, 1200 and 2000 out of 9769 SNPs being homozygote due to a putative allelic loss in blood samples, good-responding tumors and poor-responding tumors, respectively (Fig. 1B).

Any LOH in a tumor sample is evident from a reduction of the average heterozygosity over contiguous stretches of a chromosome. LOH accompanied by a copy number loss along an entire chromosome, were found in 22 out of 45 tumors. In all investigated osteosarcomas LOH affected at least 6 chromosomes with a non-randomly distribution throughout the genome. Preferentially affected were chromosomes 2, 3, 5, 6 10 and chromosome 13, which together harbored 51% of all SNPs scored as LOH.

LOH score and LOH profile as putative prediction factors

Response to chemotherapy and metastatic disease are well known prognostic factors that were also shown in our cohort of patients: the response to chemotherapy
significantly reflected a favorable outcome (event-free survival in 73% of the good responders compared with 38% of the poor responders, \( P = 0.005 \), data not shown) while the presence of metastases represented the most decisively negative prognostic factor (Fig. 1C-D). However, the frequency of LOH was also strongly associated with the patients’ outcome (Fig. 1E). The most striking finding was the unequal distribution of LOH with respect to the histologically evaluated response to chemotherapy: patients with a high LOH-score (more SNPs scored as LOH than average of 1500) significantly more often had a poor response to chemotherapy than patients with a low LOH-score (Wilcoxon-Mann-Whitney rank test, data not shown). The prognostic value of heterozygosity depended on the chromosomal region of the affected SNP. The percental distribution of the LOH values along the genome of all investigated osteosarcomas, stratified by their response to chemotherapy, is shown for each autosomal chromosome in Figure 2A.

**LOH regions in relation to response to therapy and relapse**

The most common LOH locus (in 22/45 cases; 48.9%) was found on a 4.2 Mb region of chromosome 13 (13q14.13-13q14.2; minimal deleted region in positions 45.553.450-49.731.408; 24 SNPs) harboring among other genes the osteosarcoma associated tumor suppressor gene \textit{RB1}. LOH on 13q did not show any correlation to the response to chemotherapy or event-free survival (Fig. 2B and 2E).

The region 11p15.1-11p15.4 on chromosome 11 exhibited LOH in 5 separate SNPs-stretches (in 12/44 cases; 27%), the most prominent measuring 3.6 Mb (positions 8.460.319-11.973.737; 16 SNPs) and harboring among others the genes \textit{WEE1}, \textit{ST5} and \textit{LMO1}. LOHs of 11p were found significantly (\( P < 0.001 \)) more often in poor responders than in good responders. However, there was no significant correlation with event-free survival (Fig. 2C and 2F).

The region 10q21.1 on chromosome 10 (2.5 Mb; positions 55.666.767–58.131.081; 17 SNPs), harboring among others \textit{PCDH15} and \textit{ZWINT1}, showed a very high LOH frequency (21/45 cases; 46.7%) and was significantly more often found in patients with poor response to chemotherapy (\( P < 0.005 \)). LOH in this region was significantly associated with event-free survival, independent of the absence (\( P = 0.0108 \)) or presence (0.0089) of primary lung metastases (Fig. 2D and 2G).
Clusters of Copy Number Variations (CNV)

LOH can be associated with DNA-Copy number variations (CNV). Signal intensity data for the quantification of allelic copy numbers changes indicated that 41% of the LOHs found in osteosarcoma genomes were due to allelic losses and gains, including events of duplication or chromosomal loss. The remaining 59% of SNPs showed LOH without CNV.

Long stretches of homozygosity were frequently associated with CNVs. To predefined groups, the osteosarcomas were clustered according to their CNV-similarity using the processed CNV score-profiles of all SNPs. It was found that there were three different groups of CNV-similarity on individual SNPs (Fig. 3A; squares A, B and C).

The three defined clusters illustrate the expected heterogeneity of osteosarcomas. Medial cluster B conjoins osteosarcomas (22 out of 45) of heterogenic and variable characteristics (short branches indicate low distinguishing potential) and with no obvious relation to response and prognosis; Kaplan-Meier analysis of event-free survival reveals no significant correlation with the clustering. However, the clusters A and C, which represent the both extreme characters of the dendrogram (longest branches indicate the highest dissimilarity with other subclusters), include those osteosarcomas that exhibit in a great majority either uniformity in response to chemotherapy (Fig. 3A, cluster C: poor response) or in frequency of LOH per SNP (Fig. 3B, cluster A: on average 10% of SNPs affected).

Regions of high amplifications

CNV analysis enabled a construction of CGH-like profiles (including aneuploidy and common variations on chromosome arms) and the detection of amplified “hot spot” regions. In our cohort, 27 out of 45 osteosarcomas exhibited gains of more than two-fold which affected at least one chromosomal region (more than 5 contiguous SNPs), but only 13 of these osteosarcomas showed high level amplifications (> 3-fold copy number gains). Amplifications were found to be most frequent (10 of 45 osteosarcomas) on 6p12-p21 (Fig. 4A), where three separate hot spot regions could be identified. In one case, an additional amplified region of 0.4 Mb was identified at 6p22 harboring the E2F3 gene. Six SNPs in a row displayed a signal that was on average 16 times higher than the surrounding signals (Fig. 4A).
Amplification in one or more of the three 6p-regions was found in 10/45 osteosarcomas (22%); the most frequently amplified region (in 6/45 osteosarcomas = 14%) was approximately 4.7 Mb in size, localized between physical positions 40,910,958 and 45,607,978, and included the potential candidate genes CCND3, PTK7 and RUNX2. Amplifications on 6p12-p21 were found in both response groups, namely in 4/23 patients with good and 6/22 patients with poor response. However, event-free survival was significantly (P = 0.018) less common in patients, whose osteosarcomas showed high amplifications in this region (Fig. 4B).

The MYC-containing region 8q24.21, which is 0.8 Mbp in size, was found to be amplified in 7 cases, including 5 highly amplified cases. Amplifications were significantly associated with an adverse overall survival (P = 0.01) and event-free survival (P = 0.0003). Three out of 7 patients exhibiting an 8q-amplification had primary pulmonary metastases (ppm). However, even after excluding these 3 patients with ppm from the collective the remaining 4 had an invariable significant association with event-free survival (P = 0.004) and overall survival (P = 0.01) (Fig. 4C).

The CDK4-harboring region 12q14, which is 4 Mbp in size, was found to be amplified in 5 cases, including 3 highly amplified cases. All 5 amplifications were associated with primary lung metastases (Fig. 4D) and with relapse (P = 0.07). Additionally, 10 cases had LOH stretches at the same region that were significantly associated with poor event-free survival in the whole collective (P = 0.002) and also in patients without primary metastases (P = 0.011).

Chromosomal alteration staging as predictive factor

The association between chromosomal alterations and clinical outcome of patients was investigated using Fisher’s exact test. Therefore three distinct constellations of genomic events were defined and subsequently compared with the Salzer-Kuntschik regression grading system concerning their prognostic impact: CAS1 included tumors with either an above average LOH score (more SNPs scored as LOH than average of 1500) or amplification of 6p, 8p or 12q, CAS2 comprised osteosarcomas with either LOH on 10q or amplification of 6p, 8p or 12q and CAS3 comprehended cases with a combination of 2 out of the following 4 events: LOH on 10q, amplification of 6p, 8q or 12q.

Table 2 compares the Salzer-Kuntschik regression grading and the proposed CAS constellations regarding the number of correctly predicted cases (nC). The occurrence of
relapse was used to differentiate between cases with good and poor prognosis. No difference was found between the group of good responders (S-K I-III) and CAS3 (18 correctly predicted cases with no relapse in both groups), but in all other subgroups CAS1-3 were superior in predicting the correct clinical course of patients compared with the S-K score. Since the cut-off of 10% viable tumor cells defines good and poor responders we firstly focussed on all cases showing a S-K grade III or IV, comprising exactly the cases that were close to that cut-off score. Among these cases, CAS3 predicted the clinical course accurately in 80.8% of patients compared with only 65.4% using the S-K regression grading. Also in the group of S-K poor responders (IV-VI), CAS3 predicted more than 25% more patients correctly than the S-K score (i.e. S-K IV-VI \( n_C = 61.9\% \) vs \( n_C = 90.5\% \) in CAS3). In Figure 4E-G the proposed CAS categories were evaluated concerning event-free survival and compared with the Salzer-Kuntschik regression grading (Kaplan-Meier analysis). CAS3 unequivocally displays a better prediction of the clinical outcome of patients compared with the histologically assessed response to therapy.

Since metastases at the time of initial diagnosis are the most decisive prognostic factor in osteosarcoma patients (Fig. 1C-D) we omitted these cases (n = 10) in a separate calculation. In Figure 4H-K the proposed CAS categories were again compared with the Salzer-Kuntschik score and still show a highly significant predictive impact. Nevertheless, there were two cases with primary metastases that were CAS1-3 negative and clinically showed a complete remission following therapy. One case is free of disease since 4.77 years, the other case only since 0.76 years. Although nine month of follow-up is definitely not long enough to consider a patient cured, there is at least one patient with metastatic disease at the time of diagnosis that was correctly classified as CAS negative. The remaining eight cases with primary metastases were all CAS positive and relapsed (n = 1) or died of disease (n = 7) subsequently. Chromosomal alteration staging therefore seems to be meaningful also in patients with primary metastases.

We did not find statistically significant correlations between osteosarcoma subtypes and genomic alterations, but at least one interesting trend: 11/45 osteosarcomas demonstrated at least partial chondroblastic differentiation and 8/11 were classified as poor responders according to the Salzer-Kuntschik grade. However, 4/8 poor responders are currently free of disease (range of follow-up 1.56-5.47 years, mean 3.20
years) and, thus, were prognostically misclassified. Interestingly, all these cases were CAS negative contrary to the remaining 4/8 poor responders that indeed revealed a poor clinical outcome and all turned out to be CAS positive.


**DISCUSSION**

Although the understanding of the molecular pathogenesis of osteosarcoma has advanced in the last two decades, risk assessment continues to be based mainly on clinical and histopathological parameters (14-16). Many studies have been conducted in recent years in order to establish molecular markers that might help to identify patients for whom maximal therapy is necessary and others for whom therapy of reduced intensity is sufficient to achieve long-term survival. So far, such molecular markers have not been identified. The present study was undertaken to search for genes or chromosomal locations with potential pathogenetic, prognostic and therapeutic impact in a genome-wide approach.

**LOH patterns in osteosarcoma**

In the vast majority of the genomes of the investigated osteosarcomas we observed LOHs and CNVs indicating a high degree of chromosomal aberrations and allelic imbalances. Significantly more chromosomal variations, especially LOHs, were observed in tumors with a poor response compared to tumors with a good response to preoperative chemotherapy. Furthermore, patients with tumors showing above average LOH-scores developed recurrent disease more often than patients with tumors showing low LOH-scores. A correlation between chromosomal imbalances and the prognosis of patients has recently been shown not only in osteosarcomas, but also in other tumors, including gastric carcinomas, however the biological basis of these observations is poorly understood (13, 17, 18). The high LOH scores in osteosarcomas with a poor response to chemotherapy found in our study might represent an imbalanced loss of various genes involved in cell cycle and apoptosis regulation and therefore might reflect a more aggressive phenotype with a potential survival advantage against chemotherapy.

The locus most frequently affected by LOH in our study was found on a 4.2 Mb region of chromosome 13, harboring the tumor suppressor gene \( RB1 \). The \( RB1\)-LOH rate of 43% in our study was comparable to the rate of 37.2% to 39% published in the literature (19, 20). Allelic imbalances of the \( RB1 \) locus have previously been reported in osteosarcoma but their predictive impact remains controversial. Earlier studies revealed high LOH rates for osteosarcoma at the 13q14 locus and the authors found them to be associated
with a poor prognosis (20, 21). More recently, Heinsohn et al. were not able to confirm this prognostic impact, which is in accordance with our findings (19). Remarkably, the LOH region on chromosome 10q21.1, harboring PCDH15 and ZWINT1, was more important in respect of prognostic discrimination (Fig. 2D and 2G). The observed LOH frequency was equivalent to those of the RB1 locus, but occurrence of LOH in this region was adversely correlated with event-free and overall survival. Another region with a recurring LOH pattern in 22% of the investigated cases was found on chromosome 12q13-14 harboring CDK4, which was described in other studies before (16, 22). In our study, we remarkably did not only find amplifications of this region but also high score LOH-stretches which were significantly correlated with relapse.

Only recently, the deletion of the 9p region containing the CDKN2A gene has been reported to represent an early event in mouse models for osteosarcoma development and to constitute an independent factor for adverse clinical outcome of osteosarcoma patients (23, 24). Due to limited resolution of the 10K2 high-density SNP arrays used, our data did not allow to determine deletions of the CDKN2A gene but detected LOHs of the respective locus in 5/45 cases (11%), which is in agreement with the literature. However, we found no correlation to clinico-pathological parameters including the clinical outcome of patients. Freeman and colleagues reported the common event of copy number gains in the EGFR and copy number losses in the PTEN gene and therefore suggested a dysregulated PI3K-AKT/mTOR pathway to play a role in the development of osteosarcoma (25). In our study we identified LOHs in the region of the EGFR gene in 13/45 cases (29%) and of the PTEN gene in 6/45 cases (13%). In agreement with the study of Freeman et al. we did not find correlations to clinico-pathological parameters. Also the region 3q13.31, harboring the LSAMP gene, has recently been reported to be commonly affected by chromosomal alterations in osteosarcoma. These findings, predominantly representing deletions, were correlated with disease progression and poor survival (26, 27). Due to sufficient resolution of the respective region, we were able to detect deletions in 15/45 cases (33%) in our series, 5 of which additionally showed LOHs in the flanking DNA. Additionally, we found four cases with LOH that did not show coincidental deletion of 3q13.31. However, although not statistically significant, the deletion of 3q13.31 showed a trend towards adverse outcome in our series, which is in agreement with the literature (26, 27).
CNV and Aneuploidy

To estimate the differentiation power of CNV values determined in our study we constructed clusters of similarity (Fig. 3). The resulting tree revealed three major similarity groups and one outlier (OS49). Two groups (denoted A and C) covered osteosarcomas that exhibited a good correlation between LOH and the response to chemotherapy. The remaining cases (group B, harboring the majority of patients with primary metastases) represented another subgroup of osteosarcomas which lacked this concordance. Thus, a definite classification of osteosarcomas is not feasible on the basis of the CNV pattern, most probably due to the high heterogeneity of these tumors. Suggested proclivity for genomic instability in cancer cells is reflected by the degree of aneuploidy (28). It has been suggested that imbalances affecting thousands of genes as caused by long-range chromosomal gains or losses may be an independent contributor to carcinogenesis (29, 30). Frequent amplification of centromeres or abnormalities in spindle apparatus development leading to missegregation of chromosomes has been observed in numerous types of malignant tumors and is considered as the major contributing factors for chromosome instability in cancer cells (31). Interpretation and combination of our LOH- and CNV-data allow an overview of frequencies of which chromosomes or chromosomal parts are affected by allelic duplications or losses (virtual karyotyping). Aneuploidy is very common in osteosarcomas, only 2 samples within our collective showed a normal karyotype without aneuploidy. On average 5.5 chromosomes were affected in every sample (median 6), with a maximum of 16 chromosomes in the karyotype of one patient. Aneuploidy is a common characteristic of tumors and has been proposed to drive tumor progression (32-34). Our results show that high level aneuploidy in osteosarcomas is related to poor response to chemotherapy, probably reflecting the aggressiveness of these tumors.

Amplification hotspots

Chromosome 6p is often affected by copy number gains in several types of cancer, as revealed by numerous CGH studies (35, 36). The most commonly amplified genomic interval of 6p21–p23 harbors among others several candidate genes: angiogenesis-associated VEGFA, cell cycle-associated E2F3, CCND3, RUNX2 and carcinoma-associated PTK7. These genes are directly or indirectly involved in the pathogenesis and pathways of malignant tumors and amplifications have also been described in a
high percentage of osteosarcomas (4, 11, 12, 17, 37-39). In the present study, we found high amplification spots in 24% of the investigated tumors (Fig. 4B) that were significantly correlated with a poor event-free survival. Our data imply a complex pattern of imbalances; there is not a single amplicon covering the whole central region of 6p, instead we found a few smaller hot spot regions harboring CCND3, RUNX2 and VEGFA. These genes have already been found to be overexpressed in osteosarcomas (38) and high expression of VEGFA has been correlated with a poor prognosis (40, 41). Interestingly, the cell-cycle-associated E2F3, a downstream target of RB1, which has been found amplified and overexpressed in other tumors e.g. in urothelial cancer (42), was the only gene located within a small amplified spot suggesting an alternative way to interfere with the RB1-pathway.

Amplifications found on 8q24.21 (MYC) were significantly correlated with poor event-free survival, independent of the presence of primary metastases. Only one patient of our study harboring this amplification is still event-free (short follow-up of only 1.2 years up to now). Prior investigations on mRNA expression of c-MYC in osteosarcomas showed a correlation between c-myc overexpression and event-free survival (43), supporting the suggested unfavorable prognostic value of the MYC gene overexpression. The authors found c-MYC overexpression in 42% of osteosarcomas, more than two times more frequent than in our group of tumors (15.6%). These findings implicate that DNA copy-number gains are not the only possible mechanisms to regulate c-MYC expression. MYC amplification is already known as a poor prognostic marker in other tumors e.g. neuroblastomas and lead to therapy stratification. The distinct prognostic value of MYC amplification in our study should be verified prospectively.

Amplification of 12q14 has already been shown in osteosarcomas (44). This region harbors CDK4, an important co-factor in mechanisms regulating cell-cycle progression, and MDM2, often reported to be co-amplified with CDK4. Amplification of the genomic interval containing MDM2 was not found in our study, a strong hint in favor of two independent amplification events (45, 46). Amplifications on 12q14 were found in 11% of the investigated tumors, but exclusively in patients with primary pulmonary metastases, suggesting a potential role of CDK4 in developing a metastatic phenotype at least in a subset of osteosarcomas.
Chromosomal alteration staging system as predictive factors

Our study reports one of the first uses of a SNP array for genome-wide screening in osteosarcoma, a tumor type that is known for its complex genomic alterations. This complexity hampers the identification of genes in all genome-wide approaches, but confers a clear advantage for studies that combine LOH and CNV data concurrently. With the powerful method of SNP array analysis, we identified a robust and simple marker, the overall genomic instability (the overall LOH score), to predict response to chemotherapy at the time of diagnosis. Additionally, we have shown that this tool is able to find relevant genes for prognosis (MYC, CDK4) and to identify new potential target genes (E2F3, ZWINT-1), which will be subject to ongoing studies.

Based on the detailed analysis of chromosomal alterations we were able to define a staging system (CAS) that suggests high superiority in predicting the clinical outcome of patients compared with the histologically assessed regression grading. Especially the combination of two different events of chromosomal instability (LOH on 10q, amplification of 6p, 8q or 12q), designated CAS3, was significantly more reliable in differentiating patients with good and poor prognosis, even in the S-K groups near the cut-off of 10% viable tumor cells (S-K groups III and IV, defining good and poor responders). Notably, our proposed CAS system is able to predict the prognosis of osteosarcoma patients at the time of initial diagnosis, whereas the histologic assessment is possible not before the end of neoadjuvant treatment. Since the currently used regimens include high-dose chemotherapy for ten weeks prior to definitive surgery, knowledge of the achievable effect of this treatment at initial diagnosis might lead to refinement and modification of the therapy-schedule. Furthermore, patients who do not benefit from the current chemotherapy protocols could be spared its associated short-term and long-term side effects. Importantly, the proposed CAS system is applicable for both osteosarcoma patients with and without primary metastases as was shown by omitting the patients with metastatic disease at the time of diagnosis (Fig. 4 H-K). Even in the subset of cases with initial metastases the CAS system was able to detect CAS negative cases that are currently free of disease.

The practicability of CAS in the routine diagnostic procedure has to be proven. At present, we are developing a simplified test based on PCR-SNP assays, whose analysis is technologically feasible in a modern diagnostic laboratory. It will be important to validate the findings reported here in prospective clinical trials and to evaluate the
usefulness of the chromosomal markers identified here for strategic treatment decisions in osteosarcoma.

**Acknowledgements**

We thank Nina Weber for excellent technical assistance.
LITERATURE


FIGURE LEGENDS

Figure 1: 
Affymetrix 10K2 based DrawBoxPlot-analysis of 44 osteosarcoma patients grouped according to their Salzer-Kuntschik (S-K) score in good and poor responders (24 patients with S-K score I-III vs 21 patients with S-K score IV-VI), A: heterozygosity in % SNPs; B: amount of SNPs scored as LOH (9.769 SNPs analyzed); Kaplan-Meier analyses comparing patients with and without primary pulmonary metastases (ppm) show highly significant differences concerning overall survival (C, p=0.00003) and event-free survival (D, p=0.0002); E: Kaplan-Meier analysis comparing patients with and without high LOH scores demonstrate a trend towards adverse outcome for patients with high LOH scores (patients with ppm were excluded in dotted and included in solid lines).

Figure 2: 
A: The chromosomal pattern of genome-wide LOH probability derived by SNP arrays in 44 osteosarcomas (proportional sized probe sets for each chromosome). Each chromosome illustration shows the average percentage of LOH affected SNPs, calculated for the good (left greenish column) and the poor (right reddish column) response to chemotherapy patient-group, respectively. In the heatmap-diagram, cold colors represent low, hot colors high frequencies of LOH; B-D: LOH distribution in osteosarcomas on chromosomal regions. Average percentage of LOH affected SNPs was calculated separately for patients with poor (black color) and good (grey color) response to chemotherapy. B: chromosome 13 contained a region on 13q14.2 – harboring the RB1 gene – of high LOH frequency that affected 22/44 osteosarcomas (50%); C: 5 distinct regions (marked by red colored bars) on chromosome 11p15.1-4 (marked by red dots) found in 12/44 osteosarcomas (27%) strongly discriminated between poor and good response to chemotherapy; D: chromosome 10q21.1 contained a region with LOH in 21/44 osteosarcomas (47%) that also decisively distinguished between good and poor responders; E-G: Kaplan-Meier analysis for event-free survival.
(the patients with ppm were excluded in dotted and included in solid lines): no significant correlation was found between LOH on 13q14.2 and relapse (E, p=0.3) or between LOH on 11p15.1-4 and relapse (F, p=0.8); between LOH on 10q21.1 and relapse a highly significant correlation was detected (G), independent of including (p=0.0089) or excluding (p=0.0108) ppm patients.

Figure 3: CNV-clusters display the similarity among the entire copy number profile of the osteosarcomas. Numbers indicate individual osteosarcomas. Grey and black bars indicate good and poor response to chemotherapy, respectively. Bars with a red margin indicate patients with primary lung metastases. Three different similarity branches are discriminable (clusters A-C); the only non-clustering sample is osteosarcoma 49. A shows the distribution of the prognostic characteristics defined by histological response and primary lung metastases displayed at the CNV cluster tree. The similarity cluster C includes a high percentage of poor responders (7/8, 87%), small branches in A and B include exclusively osteosarcomas with good response; B shows the distribution of defined LOH at the same CNV-cluster. Remarkably, many osteosarcomas grouped in cluster A (12/16 = 75%) display a low frequency of LOH affected SNPs (<10%).

Figure 4: A: Chromosomal region 6p12-p21 spans a stretch of nearly 25 Mb between the positions 54741994 (6p12) and 30200689 (6p21). High amplification patterns in osteosarcomas found on chromosome 6 showed three distinct loci of amplification with a mean peak of frequency in a 5 Mb stretch between positions 40910958 and 45607978. The additional single event of high-amplification found on 6p22 (at positions 19921598-20345022) harbors the gene E2F3; B-D: high-amplification regions in osteosarcomas – Kaplan-Meier analysis for event-free survival (the patients with ppm were excluded in dotted and included in solid lines); B: 6p12-p21: amplification was significantly associated with poor event-free survival (p=0.018); C: 8q24.21: all patients affected by
amplification suffered from relapse within 2 years after initial diagnosis, irrespective of including (p=0.004) or excluding (p=0.0003) ppm patients; D: 12q14: all five patients affected by amplification showed primary metastases; there was a trend towards event-free survival (p=0.07); E-G: Kaplan-Meier analysis for event-free survival comparing the Salzer-Kuntschik regression grading (dotted lines) and the proposed CAS systems (solid lines): E: CAS1, F: CAS2 and G: CAS3; Kaplan-Meier analysis for event-free survival comparing the Salzer-Kuntschik regression grading (dotted lines) and the proposed CAS systems (solid lines) only for patients without primary metastases: H: CAS1, I: CAS2 and K: CAS3.
A

\[\text{%SNP scored as LOH}\]

\[\begin{array}{c}
\text{13} & \text{14} & \text{15} & \text{16} & \text{17} & \text{18} & \text{19} & \text{20} & \text{21} & \text{22} \\
\text{23 OS good response} & \text{21 OS poor response} \\
\text{0}\% & \text{25}\% & \text{50}\% \\
\end{array}\]

B

\[\text{RB1}\]

\text{Chromosome 13q}

\[\begin{array}{c}
\text{q14.2} \\
\text{q21.1} \\
\text{q31.1} \\
\text{q34} \\
\end{array}\]

C

\[\text{p15.4 - p15.1}\]

\text{Chromosome 11p}

\[\begin{array}{c}
\text{p15.4} \\
\text{p13} \\
\text{p12} \\
\end{array}\]

D

\[\text{q21.1}\]

\text{Chromosome 10}

\[\begin{array}{c}
\text{p14} \\
\text{p13} \\
\text{p14} \\
\text{q21.1} \\
\text{q21.3} \\
\text{q25.1} \\
\end{array}\]

E

\[\text{LOH-score in 13q14.13-14.2}\]

\[\text{EFS} \quad \text{n} = 26(-5\text{ ppm})\]

\[\text{relapse} \quad \text{n} = 20(-5\text{ ppm})\]

F

\[\text{LOH-score in 11p15.1-15.4}\]

\[\text{EFS} \quad \text{n} = 31(-6\text{ ppm})\]

\[\text{relapse} \quad \text{n} = 15(-4\text{ ppm})\]

G

\[\text{LOH-score in 10q21.1}\]

\[\text{EFS} \quad \text{n} = 25(-5\text{ ppm})\]

\[\text{relapse} \quad \text{n} = 21(-5\text{ ppm})\]
A 6p Amplifications

<table>
<thead>
<tr>
<th>NoProbe</th>
<th>Physical size</th>
<th>genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
<td>Position</td>
<td></td>
</tr>
<tr>
<td>SNP_A-1516084</td>
<td>19921598 bp</td>
<td>E2F3</td>
</tr>
<tr>
<td>SNP_A-1509639</td>
<td>20345022 bp</td>
<td></td>
</tr>
<tr>
<td>SNP_A-1511850</td>
<td>22502049 bp</td>
<td>HDGFL1, GMNN</td>
</tr>
<tr>
<td>SNP_A-1518505</td>
<td>25382624 bp</td>
<td></td>
</tr>
<tr>
<td>SNP_A-1510234</td>
<td>30200686 bp</td>
<td></td>
</tr>
</tbody>
</table>

SNP_A-1514645 | 36901313 bp | CCND3, RUNX2 |
| SNP_A-1513942 | 40910958 bp |       |
| SNP_A-1518706 | 45607978 bp |       |
| SNP_A-1516564 | 46437363 bp |       |
| SNP_A-1510836 | 53540169 bp |       |
| SNP_A-1510328 | 54741994 bp | KLHL31, ZNF45 |
| SNP_A-1517757 | 58724802 bp |       |

B Amplification on 6p12-21

EFS

relapse n = 10 (-5 ppm)
n = 36 (-5 ppm)

C Amplification on 8q24.21

EFS

relapse n = 7 (-3 ppm)
n = 39 (-7 ppm)

D Amplification on 12q14

EFS

relapse n = 5

E CAS1 staging

n = 29

P=0.02

F CAS2 staging

n = 26

P=0.006

G CAS3 staging

n = 17

P=6.3x10^{-6}

H CAS1 staging

n = 21

P=0.06

I CAS2 staging

n = 18

P=0.04

J CAS3 staging

n = 9

P=0.0002
Table 1: Patient characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Number of patients (n = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at diagnosis (years)</strong></td>
<td></td>
</tr>
<tr>
<td>range</td>
<td>4-51</td>
</tr>
<tr>
<td>mean</td>
<td>16.5</td>
</tr>
<tr>
<td>median</td>
<td>14</td>
</tr>
<tr>
<td><strong>Gender (n)</strong></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>25</td>
</tr>
<tr>
<td>female</td>
<td>20</td>
</tr>
<tr>
<td><strong>Blood samples available (n)</strong></td>
<td></td>
</tr>
<tr>
<td>number</td>
<td>9</td>
</tr>
<tr>
<td><strong>Site of primary tumor (n)</strong></td>
<td></td>
</tr>
<tr>
<td>femur</td>
<td>27</td>
</tr>
<tr>
<td>tibia</td>
<td>8</td>
</tr>
<tr>
<td>humerus</td>
<td>6</td>
</tr>
<tr>
<td>pelvis</td>
<td>2</td>
</tr>
<tr>
<td>fibula</td>
<td>1</td>
</tr>
<tr>
<td>unknown</td>
<td>1</td>
</tr>
<tr>
<td><strong>Histologic differentiation (n)</strong></td>
<td></td>
</tr>
<tr>
<td>osteoblastic</td>
<td>28</td>
</tr>
<tr>
<td>chondroblastic</td>
<td>11</td>
</tr>
<tr>
<td>fibroblastic</td>
<td>4</td>
</tr>
<tr>
<td>teleangiectatic</td>
<td>2</td>
</tr>
<tr>
<td><strong>Metastasis (n)</strong></td>
<td></td>
</tr>
<tr>
<td>lung</td>
<td>10</td>
</tr>
<tr>
<td><strong>Follow up (years)</strong></td>
<td></td>
</tr>
<tr>
<td>range</td>
<td>0.6-12.7</td>
</tr>
<tr>
<td>mean</td>
<td>3.84</td>
</tr>
</tbody>
</table>
### Clinical outcome (n)

<table>
<thead>
<tr>
<th>Response</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>27</td>
</tr>
<tr>
<td>CR-lost</td>
<td>2</td>
</tr>
<tr>
<td>AWD-lost</td>
<td>3</td>
</tr>
<tr>
<td>DOD</td>
<td>13</td>
</tr>
</tbody>
</table>

### Response to chemotherapy (n)

<table>
<thead>
<tr>
<th>S-K grade</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>10</td>
</tr>
<tr>
<td>III</td>
<td>12</td>
</tr>
<tr>
<td><strong>good</strong></td>
<td><strong>23</strong></td>
</tr>
<tr>
<td>IV</td>
<td>14</td>
</tr>
<tr>
<td>V</td>
<td>6</td>
</tr>
<tr>
<td>VI</td>
<td>1</td>
</tr>
<tr>
<td><strong>poor</strong></td>
<td><strong>21</strong></td>
</tr>
<tr>
<td>no S-K available</td>
<td>1</td>
</tr>
</tbody>
</table>

**Abbreviations:** CR = complete remission, lost = lost of follow up, AWD = alive with disease, DOD = dead of disease; S-K = Salzer-Kuntschik, grade I = no residual viable tumor, II = solitary viable tumor cells, III = < 10% viable tumor, IV = 10-50% viable tumor, V = 50-80% viable tumor, VI = > 80% viable tumor.
Table 2: Correlation between the Salzer-Kuntschik regression grading and the proposed chromosomal alteration staging systems

<table>
<thead>
<tr>
<th></th>
<th>nC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>above average</td>
</tr>
<tr>
<td>All cases</td>
<td></td>
</tr>
<tr>
<td>n = 45</td>
<td></td>
</tr>
<tr>
<td>nR = 18 (40%)</td>
<td>28 (62,2%)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S-K I-III</td>
<td></td>
</tr>
<tr>
<td>n = 23</td>
<td></td>
</tr>
<tr>
<td>nR = 5 (21,7%)</td>
<td>11 (47,8%)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S-K III-IV</td>
<td></td>
</tr>
<tr>
<td>n = 26</td>
<td></td>
</tr>
<tr>
<td>nR = 9 (34,6%)</td>
<td>15 (57,7%)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S-K IV-VI</td>
<td></td>
</tr>
<tr>
<td>n = 21</td>
<td></td>
</tr>
<tr>
<td>nR = nC = 13 (61,9%)</td>
<td>16 (76,2%)</td>
</tr>
<tr>
<td>Abbreviations: ( n_C ) = number of cases with correct prediction of the clinical course, ( n_R ) = number of cases with relapse, ns = not significant, * = insufficient case numbers for statistics.</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>S-K V-VI</td>
<td>( n = 7 )</td>
</tr>
<tr>
<td></td>
<td>6 (85.7%)*</td>
</tr>
</tbody>
</table>
Genomic Alterations and Allelic Imbalances are Strong Prognostic Predictors in Osteosarcoma

Jan Smida, Daniel Baumhoer, Michael Rosemann, et al.

Clin Cancer Res  Published OnlineFirst July 7, 2010.

Updated version  Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-10-0284

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.