Whole exome sequencing of a single osteosarcoma case—integrative analysis with whole transcriptome RNA-seq data

Ene Reimann¹,²*, Sulev Kõks¹,², Xuan Dung Ho³,⁴, Katre Maasalu³,⁵ and Aare Märtson³,⁵

Abstract

Background: Osteosarcoma (OS) is a prevalent primary malignant bone tumour with unknown etiology. These highly metastasizing tumours are among the most frequent causes of cancer-related deaths. Thus, there is an urgent need for different markers, and with our study, we were aiming towards finding novel biomarkers for OS.

Methods: For that, we analysed the whole exome of the tumorous and non-tumour bone tissue from the same patient with OS applying next-generation sequencing. For data analysis, we used several softwares and combined the exome data with RNA-seq data from our previous study.

Results: In the tumour exome, we found wide genomic rearrangements, which should qualify as chromotripsis—we detected almost 3,000 somatic single nucleotide variants (SNVs) and small indels and more than 2,000 copy number variants (CNVs) in different chromosomes. Furthermore, the somatic changes seem to be associated to bone tumours, whereas germline mutations to cancer in general. We confirmed the previous findings that the most significant pathway involved in OS pathogenesis is probably the WNT/β-catenin signalling pathway. Also, the IGF1/IGF2 and IGF1R homodimer signalling and TP53 (including downstream tumour suppressor gene E24) pathways may have a role. Additionally, the mucin family genes, especially MUC4 and cell cycle controlling gene CDC27 may be considered as potential biomarkers for OS.

Conclusions: The genes, in which the mutations were detected, may be considered as targets for finding biomarkers for OS. As the study is based on a single case and only DNA and RNA analysis, further confirmative studies are required.

Keywords: Osteosarcoma, Whole exome sequencing, Integrative analysis

Introduction

Osteosarcoma (OS) is a most prevalent primary malignant bone tumour and mostly occurs in children and adolescents—75% of patients with OS are 15 to 25 years old. The etiology is unknown; however, a genetic predisposition has been suggested [1,2]. Reviewed in [3], these tumours have high potential to metastasize and are one of the most frequent causes of cancer-related deaths. The survival rate increased up to 70% after chemotherapy became available [4]. However, no further improvements have been made in the last decades in terms of survival. Thus, the survival plateau forces scientists to look for new biomarkers (diagnostic, disease monitoring, response, resistance markers, drug targets), which could lead to, i.e. applying new therapeutic agents. While OS is rare and very heterogeneous (inter-patient, inter-tumour and intra-tumour heterogeneity), the clinical study progress is slow; thus, the preclinical studies are vital. Furthermore, finding the biomarkers and detecting the potential targets for new drugs are essential to improve the present situation.

There are several next-generation sequencing (NGS) and genome-wide association studies (GWAS) about OS, which associate different genes and pathways with pathogenesis of OS [5-7]. With whole exome sequencing...
(WES) and whole genome sequencing (WGS) studies, TP53, PTEN and PRB2 are found to be mutated in significant frequency [5]. High mutation rate in TP53 has also demonstrated in OS cell lines. Additionally, deletion of CDKN2A/B locus and amplification of MDM2 were detected [8]. With GWAS studies, a single nucleotide variant (SNV) in GRM4 was detected as potential biomarker for OS [7]. Gene expression studies reveal that, i.e. WNT inhibitory factor (WIF1) has a loss of expression in OS cell lines [9]; however, we found in our previous work that the expression has increased significantly [10]. Thus, as demonstrated, the expression pattern of WNT pathway genes in different OS cases may not be similar. When correlating the expression patterns of miRNA/mRNA pairs, miRNAs regulating TGFBR2, IRS1, PTEN and PI3K have been detected [11]. In addition, several serine/threonine kinases (mechanistic target of rapamycin (mTOR)) or tyrosine kinases (SRC, IGF1R, PDGFR, KIT) are considered as targets in OS treatment [3,12,13].

When observing the related pathways, the WNT/β-catenin pathway is one of the most thoroughly studied among bone malignancies. For example, the tumour growth is regulated through this pathway and the overexpression of BMP9 suppresses its activity [14,15]. Furthermore, PI3K/AKT/mTOR signalling pathway was brought forward as a potential target for therapy, and also, pathways associated to TP53 may be altered [5]. Hypoxia-HIF-1α-CXCR4 pathway plays a crucial role during the migration of human osteosarcoma cells [16]. These are just a few examples—the network of associated genes and pathways is complex.

OS has a very unstable genome—it may contain aberrant number of chromosomes, and in most cases, these chromosomes display major structural abnormalities including amplification, deletions and translocations. For example, several studies have demonstrated the gain of chromosomal arms 6p, 8q and 17p in the case of OS [17,18]. To be more precise, i.e. VEGFA amplification and LSAMP deletion have been detected in OS [19,20]. Thus, it is suggested that genomic instability is linked to the development of this tumour [18,21-23]. Furthermore, the genomic aberrations are more frequent in metastases than in primary tumours [24]. The genes responsible for cell cycle regulation are suggested to be associated to DNA breakage and genomic instability, i.e. CDC5L overexpression and mutations in TP53 gene are correlated to the high genomic instability in OS [23,25]. Moreover, the chromothripsis event is characteristic to OS—it generates new fusion products. This may explain the sudden onset of OS and the complexity and heterogeneity of OS genome [26]. All these changes make it difficult to find biomarkers suitable for targeting OS, as there are so many different subtypes.

In the present work, we analysed the whole exome of the tumorous and non-tumour bone tissue from the same patient with osteosarcoma. We used next-generation sequencing to study how the coding region of the tumour genome has altered. Additionally, we analysed together the WES genotyping and RNA expression data (from our previous RNA-seq analysis).

Materials and methods
Subject
The protocols and informed consent form used in this study were approved by the Ethical Review Committee on Human Research of the University of Tartu. The patient signed a written informed consent, which also includes the acceptance of the report to be published. A 16-year-old Caucasian male patient with an OS diagnosis was studied. In more detail, the patient became ill with complaints of pain in the left knee area. History of trauma was missing, and GP administered painkillers and vitamins. After 6 months, the patient returned to GP with complaints of pain, swelling and dysfunction in the left distal femur and knee area. The swelling line was observed in the left femoral distal region, and the area was thicker and painful to touch. No changes in skin colour were detected. The X-ray investigation showed additional shading and structural change in the distal part of the left femur. For detailed investigation, the MRI was performed and as a result, malignant process was suspected. Patient was hospitalized, and bone biopsy was taken for histological investigation. The diagnosis of osteosarcoma was confirmed. Chemotherapy for osteosarcoma started by Scandinavian Sarcoma Group (SSG) XIV treatment protocol. The patient responded well to the therapy—the histological analysis confirmed the necrotic tissue in tumour. After 3 month of chemotherapy, surgical removal of tumour (distal part of femoral bone with knee joint) and replacement of the knee and the lower part of the femur with megaprosthesis was performed. Pathologist confirmed that resection line was without tumour cells and OS was referred as NAS (Not Further Specified). After the patient had recovered from surgery, the SSG XIV chemotherapy treatment protocol was followed. Materials for this study were collected from the surgically removed tissue.

Exome sequencing
The genomic DNA (gDNA) was extracted from two bone samples from different locations—one sample from tumour area and another sample from the uninvolved normal bone tissue as a control. For gDNA extraction, the tissue was homogenized applying liquid nitrogen and a mortar, and after that, the PureLink Genomic DNA kit (Life Technologies Corp., Carlsbad, CA, USA) was used according to manufacturer’s protocol. The Target Seq Exome Enrichment System and SOLID 5500 barcoded adaptors (Life Technologies Corp., Carlsbad, CA, USA)
were used to prepare the libraries. The SOLiD 5500xl platform and paired-end DNA sequencing chemistry (75 bp forward and 35 bp reverse direction) were applied to sequence the samples.

The data analysis
Offline cluster was used for data processing and analysis. For bioinformatic analysis, LifeScope version 2.5 was applied. LifeScope performed colour space mapping and pairing. Tertiary analysis consisted of SNV discovery (diBayes algorithm) and detection of small indels. Hg19 (GRCh37.p13) was used as a reference, and before mapping, the multifasta file was ordered in increase to order the mapping quality.

The SNVs and small indel .gff3 files were used as input in ANNOVAR software (AS; www.openbioinformatics.org/annovar/) [27] and Ingenuity Variant Analysis (IVA; http://www.ingenuity.com) QIAGEN, Redwood City, MD, USA) software. Applying refGene hg19, dbSNP135 and dbCOSMIC67 databases, AS annotated and predicted the effects of SNVs and small indels we detected in our study samples. AS also provides other prediction tools in order to get prediction scores (PolyPhen-2, SIFT, ljb2 etc.) [28-30]. Comparative distribution of SNVs and small indels between different samples was performed with Galaxy software bundle [31,32]. IVA provided tools to annotate SNVs and small indels, which may be associated to cancer. The tumour and control samples were compared, and the lists for diseases, processes and pathways related to cancer were received as output.

The .bam and .bai files were used as input in CEQer software (CS) (www.ngsbioocca.org/html/ceqer.html), which is a tool for analysing copy number variants (CNVs) and loss of heterozygosity (LOH).

About the RNA-seq data analysis, please see our previous article, where we used the bone samples from the same patient [10].

Results
For comparing the tumour tissue and non-tumour tissue (control tissue) from the same individual, different approaches were applied. After mapping the data to a reference genome, we used several tools to perform the tertiary analysis.

Sequencing statistics from LifeScope software
In the case of the tumour tissue, over 130 million (58%) mappable reads were in target and the enrichment fold was 48¥. Eighty-five percent of the detected targets were covered over 20 times, and the average coverage was 185.5. In the case of the control tissue, over 154 million (61%) mappable reads were in target and the enrichment fold was 51%. Eighty-three percent of the detected targets were covered over 20 times, and the average coverage was 157.

SNVs, small indels and CNVs
1) Results from ANNOVAR software
Using refGene hg19 database, AS was able to annotate 37,990 SNVs and 1,484 small indels. In the case of SNVs, we considered the data reliable, if the coverage was over 20; thus, 25,914 SNVs remained. In the case of SNVs, there were 23,767 germline mutations (9,067 in homozygous form and 14,700 in heterozygous form) and 2,147 somatic mutations (in the tumour tissue—116 in homozygous form and 2,031 in heterozygous form) (Table 1, Additional file 1). Furthermore, there were 896 germline small indels (278 in homozygous form and 618 in heterozygous form) and 588 somatic indels (in the tumour tissue—177 in homozygous form and 411 in heterozygous form).

Applying dbSNP135, we were able to annotate 5,281 SNVs and 239 small indels. With dbCOSMIC67, we annotated 2,569 SNVs and 59 small indels—none of these were noted to be associated to bone cancer. Applying ljb2 database, we found 469 SNVs to potentially cause a disease (average ljb2 score over 0.918), including 31 germline mutations and 4 somatic mutations (ESX1: c. A578G/p.K193R; CDC27: c.A17G/p.E6G; TMEM120B: c.G274A/p.D92N; TMEM131: c.C3947T/p.P1316L) in homozygous form in the tumour tissue.

2) Results from Ingenuity Variant Analysis software
Altogether, 207 cancer driver variants (CD-SNVs) were found in 123 genes according to IVA (Additional file 2). Fourteen CD-SNVs potentially gain and 186 lose the gene function. Only seven SNVs may have no drastic effect on gene function in the tumour tissue. Furthermore, according to IVA, none of these 207 SNVs affect the gene functionality in the control tissue. Thirteen of the CD-SNVs were homozygous in the tumour tissue (Table 2). There were no cancer-associated homozygous mutations present in the control tissue; thus, the homozygous CD-SNVs in the tumour tissue are all somatic.

According to IVA, six cancer-associated small indels were found (Table 2). Four of them are homozygous and two are heterozygous in the tumour tissue—the effect is most probably the loss of gene function. These indels are predicted to have no effect in the control tissue.

In most of the genes brought front by IVA, one CD-SNV was found in coding region in heterozygous form. However, some of the genes have more CD-SNVs in coding regions: MUC4 had even 22, ZNF717 had 8, CTBP2 had 7 and OR4C3 had 5 CD-SNVs, whereas these were not present in the control tissue (data not shown). When observing from a slightly different angle—the gene complexes, we can see that the mucin complex has the highest
significance—three genes and 27 CD-SNVs are considered (Table 3). There are also other gene complexes, which are potentially associated to cancer processes, and in different complexes, the CD-SNVs are either somatic or germline (Table 3).

In the case of cancer-associated small indels, the statistically most significant results were with complexes related to RELA gene—NFKB1-RELA and RELA-REL complexes both had $p$ value 7.56E-4.

IVA provided the first 100 cancer-associated processes and diseases related to CD-SNVs and small indels. Seventy-three genes and 135 CD-SNVs were found associated to process named as “disorder of genitourinary system” (Table 4). These findings were present in both the tumour and control tissues. There were also two processes associated to bone “myelopoiesis of bone marrow” (associated genes NPM1, RARA) and “quantity of trabecular bone” (associated genes CREBBP, SMO)—these findings were present only in the tumour tissue. In the case of small indels, all the findings were somatic and ALK and RELA genes were associated to “outgrowth of bone marrow cells” and “inflammatory response of bone marrow-derived macrophages”, respectively.

IVA found 111 genes with 202 germline CD-SNVs associated to cancer (Table 5). Fifteen genes, which had 43 somatic CD-SNVs were associated to “bone marrow cancer and tumours”. In the case of small indel, all six genes, with a finding, are associated to cancer and the found small indels are all somatic. The disease named as “tumourigenesis of bone tumour” was associated to small indel in ALK gene and was present only in the tumour tissue.

With the osteosarcoma patient’s tumour and control tissue, WES data IVA found six pathways associated to CD-SNVs and six to cancer driver small indels (Table 6). All the mutations considered here were somatic. In the case of CD-SNVs, the statistically most significant association was between tumour and WNT/β-catenin signalling pathway. In the case of small indels, associations with different cytokine pathways were found. Also, a pathway directly linked to the bone tissue—“RANK signalling in osteoclasts” was brought front.

| Table 1 The numbers of SNV and small indel findings received from data analysis with ANNOVAR software |
|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Germline mutations                  | Somatic mutations                   |                                     |
| Homozygous: non-reference           | Homozygous in tumour                | Homozygous in control               |
| Heterozygous                        | Heterozygous in tumour              | Homozygous in control               |
| Heterozygous in control             |                                     | Homozygous in control               |
|                                    |                                     | (reference)                         |
|                                    |                                     | (reference)                         |

| SNVs                                | Germline mutations                  | Somatic mutations                   |
|-------------------------------------|-------------------------------------|-------------------------------------|
| Altogether                          | 9,067                               | 14,700                              |
| Exonic (includes ncRNA)            | 5,244                               | 8,702                               |
| Nonsynonymous                       | 2,435                               | 4,035                               |
| Stopgain                            | 6                                   | 50                                  |
| Stoploss                            | 2                                   | 5                                   |
| Splicing (includes exonic)          | 11                                  | 20                                  |
| Intronic (includes ncRNA)          | 3,091                               | 4,846                               |
| 5’ UTR and 3’ UTR                   | 515                                 | 797                                 |
| Downstream and upstream             | 51                                  | 76                                  |
| Intergenic                          | 155                                 | 259                                 |

| Small indels                        | Germline mutations                  | Somatic mutations                   |
|-------------------------------------|-------------------------------------|-------------------------------------|
| Altogether                          | 278                                 | 618                                 |
| Exonic (includes ncRNA)            | 33                                  | 99                                  |
| Frameshift                          | 9                                   | 30                                  |
| Stopgain                            | 0                                   | 1                                   |
| Splicing (includes exonic)          | 4                                   | 16                                  |
| Intronic (includes ncRNA)          | 212                                 | 419                                 |
| 5’ UTR and 3’ UTR                   | 25                                  | 73                                  |
| Downstream and upstream             | 1                                   | 2                                   |
| Intergenic                          | 3                                   | 9                                   |
| Gene symbol | Chromosome number | Position | REF/ALT | Tumour zygosity | Control zygosity | Effect on function | dbSNP | SIFT function | Polyphen function | Transcript ID | Nucleotide change | Amino acid change | Gene region | Translation impact |
|-------------|------------------|----------|---------|----------------|-----------------|------------------|-------|--------------|------------------|--------------|-----------------|------------------|-------------|------------------|
| **SNVs**    |                  |          |         |                |                 |                  |       |              |                  |              |                 |                  |             |                  |
| RGPD3 (includes others) | 2   | 110586562 | A/G     | 1/1            | Loss            | 0/0              | Normal | Damaging     | Benign           | NM_001037866.1, NM_001123363.3, NM_005054.2, NM_032260.2 | c.2393A > G | p.E798G         | Exonic Missense |
| PRDM9       | 5    | 23527251 | C/T     | 1/1            | Loss            | 0/0              | Normal | Tolerated    | Probably damaging | NM_020227.2 | c.2054C > T | p.T685I          | Exonic Missense |
| FOXK1       | 7    | 4722436 | A/G     | 1/1            | Loss            | 0/0              | Normal | Damaging     | Benign           | NM_001037165.1 | c.497A > G | p.N166S         | Exonic Missense |
| CCZ1/CCZ2B  | 7    | 6841033 | T/A     | 1/1            | Loss            | 0/1              | Normal | Tolerated    | Benign           | NM_198097.3 | c.1228A > T | p.M410L         | Exonic Missense |
| PLAT4       | 8    | 4204965 | G/A     | 1/1            | Normal          | 0/0              | Normal | Tolerated    | Benign           | NM_033011.2, NM_000930.3, NM_015239.2, NM_001134707.1, NM_007101.3 | c.352C > T | p.R118W         | Exonic Missense |
| AGTPBP1*    | 9    | 88292495 | C/T     | 1/1            | Loss            | 0/1              | Normal | Tolerated    | Benign           | NM_001134707.1 | c.292G > A | p.C98R          | Exonic Missense |
| SARDH       | 9    | 136597592 | T/C    | 1/1           | Loss            | 0/0              | Normal | Tolerated    | Benign           | NM_001134707.1 | c.463A > G | p.I55SV         | Exonic Missense |
| FAH         | 15   | 80472526 | C/T     | 1/1            | Normal          | 0/1              | Normal | Damaging     | Probably damaging | NM_000137.2 | c.1021C > T | p.R341W         | Exonic Missense |
| CDC27       | 17   | 45266522 | T/C     | 1/1            | Loss            | 0/0              | Normal | Damaging     | Probably damaging | NM_001114091.1 | c.17A > G | p.E6G           | Exonic Missense |
| SBF1*       | 22   | 50893287 | T/C     | 1/1            | Loss            | 0/1              | Normal | Tolerated    | Benign           | NM_002972.2, NM_001006607.2 | c.4768A > G | p.T1590A        | Exonic Missense |
| LRRIC7A3*   | 17   | 44632540 | T/C     | 1/1            | Gain            | 0/0              | Normal | Activating   | Benign           | NM_001113738.1, NM_016632.2 | c.*2182A > G, c.259 + 15585A > G | - | 3'UTR | - | Intronic |
| ARL17A      | 17   | 44632540 | T/C     | 1/1            | Gain            | 0/0              | Normal | Activating   | Benign           | NM_001081450.1 | c.523C > G | p.R175G         | Exonic Missense |
| LILRB3      | 19   | 54725835 | G/C     | 1/1            | Gain            | 0/0              | Normal | Activating   | Benign           | NM_006864.2 |                  |                  |             |                  |
| **Small indels** | 20 | 56073500 | (N)103/T | 1/1      | Loss            | 0/0              | Normal |                  |                  | NM_001269041.1, NM_001269043.1, NM_001269040.1, NM_001269042.1, NM_001269046.1 | c.*4_105del(N)103 | 3'UTR |                  | 3'UTR |

Reimann et al. Human Genomics 2014, 8:20
http://www.humgenomics.com/content/8/1/20
Table 2 The somatic cancer driver SNVs and small indels found in data analysis with Ingenuity Variant Analysis software (Continued)

| Gene | Chromosome | Variance | Loss | Normal | NM_001134657.1 | NM_001127371.2/ | NM_001127370.2/ | NM_001127370.2/ | NM_018719.4 | NM_004304.4 | NM_014208.3 | NM_001145138.1/ | NM_001145138.1/ | NM_001243985.1 | NM_001243985.1 | NM_001243985.1 |
|------|------------|----------|------|--------|-----------------|-----------------|-----------------|-----------------|--------------|--------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| PRR23C | 3          | 138763627 | G/GCG | 1/1     | Loss 0/0       | Normal          | 63140560        |                 |              |              |                 |                  |                  |                  |                  |                  |
| CDC7L  | 7          | 21941867  | C/CTAGC | 1/1    | Loss 0/0      | Normal          |                  |                 |              |              |                 |                  |                  |                  |                  |                  |
| ALK   | 2          | 29416029  | G/GATG | 1/1     | Loss 0/0      | Normal          |                  |                 |              |              |                 |                  |                  |                  |                  |                  |
| DSPP  | 4          | 88537081  | CAG/CAGGATG | 0/1 | Loss 0/0      | Normal          |                  |                 |              |              |                 |                  |                  |                  |                  |                  |
| RELA  | 11         | 65422086  | TCT/CTG | 0/1     | Loss 0/0      | Normal          |                  |                 |              |              |                 |                  |                  |                  |                  |                  |

*The expression pattern of these genes has changed in the tumour tissue compared to that in the control tissue.*
3) Results from CEQer software

We applied CS to analyse CNVs in tumour and non-tumour tissue exomes. Compared to the control tissue, in the tumour tissue, the loss of coding sequences was found in 6 chromosomes and 183 genes and gain of coding sequences in 4 chromosomes and 65 genes (Figure 1). The loss or gain of coding sequences was altogether in 8 chromosomes, and the most altered were chromosomes 2 and 19 (193,701 bp and 115,358 bp, respectively; Figure 2). The loss of heterozygosity was detected altogether in 68 regions in 37 genes, located in 15 different chromosomes (Additional file 3).

Integrative analysis

The integrative analysis narrows down the large list of findings from NGS data. When combining the results from WES data (AS, IVA, CS) and RNA-seq data [10], we found some interesting and rather logical associations, which we would like to emphasize.

SNVs, small indels and RNA expression

To reduce down the complexity of data we received from AS, we decided to perform as follows. In the case of SNV data, we observed both somatic and germline SNVs, which are homozygous in the tumour tissue and should have an effect on translation (nonsynonymous, stopgain, stoploss findings). Thus, we got 527 homozygous germline SNVs (in 392 genes) and 8 homozygous somatic SNVs (in 7 genes), which are located in genes with altered expression in the tumour tissue compared to that in the control tissue. If also considering the Ljb2 database scores, seven homozygous SNVs with high disease-causing probability remained (Table 7).

In the case of small indels detected with AS, we observed the somatic and germline indels, which were homozygous in the tumour tissue. There was 52 germline and 26 somatic indels in introns of the genes, which expression pattern has also changed (data not shown). Furthermore, there was five germline and three somatic indels in exons of the genes with altered expression. Thus, we found altogether three frameshift small indels, which possibly have an effect on translation (frameshift insertions and deletion in exons) (Table 7).

In the case of homozygous cancer driver SNVs and small indels found with IVA (Table 2), only four genes have altered expression pattern in the tumour tissue compared to that in the control tissue. The mRNA expression was increased in the case of PLAT (log fold change (logFC) = 3.65, false discovery rate (FDR; corrected statistical significance) = 8.27E-27), AGTPBP1 (logFC = 0.91, FDR = 0.039) and LRRC37A3 (logFC = 1.14, FDR = 0.0072) and decreased in the case of SFB1 (logFC = −1.33, FDR = 0.0037).

CNVs, LOHs and RNA expression

When analysing the CNV results together with RNA expression results, we found that with gained copy numbers, there were altogether 22 genes, with altered expression profile—20 genes with increased and 2 genes with decreased mRNA expression. In the case of loss copy of number, 74 genes’ expression profile had changed—11 genes with increased and 63 genes with decreased mRNA expression. In Table 8, the genes with the lowest FDR values for gene expression results are presented. Here, we would emphasize that the INSR, which has copy number loss in area covering 174,552 bp has also a remarkable decrease in mRNA expression.
However, there are also several genes with CNVs, which could be associated to cancer.

Combining the LOH and mRNA expression data, we found that in the tumour tissue, the expression of four genes with LOH has increased significantly and expression of five genes with LOH has decreased significantly (Table 9). The rest of the genes with LOHs had no significant changes in mRNA expression level, and two genes were not detected with RNA-seq (FLJ20518, MANSC4) [10].

For additional information, please see the supplementary material as separate files for AS, IVA and CS combined with RNA-seq data.

**Discussion**

In this study, the exome profiles of the osteosarcoma patient’s tumour and normal bone tissue were compared. Additionally, the RNA-seq data from our previous work was used [10]. For WES data analysis, several softwares were applied and possibly some of them are better in detecting some mutations and not so effective in detecting others. Still, we think it is more beneficial to use different approaches and we believe it is easier to follow, if we discuss separately the results gained from each software.

The ANNOVAR software annotated a large amount of genes with SNVs and small indels, applying refGene hg19 database. Over 2,700 somatic SNVs and small indels were detected specifically in the tumour tissue, from which almost 300 are homozygous. These findings are located all over the exome. This demonstrates that the changes in OS genome are not concentrated into a single or few areas but are rather distributed.

When using ljb2 database, AS detected four homozygous somatic mutations in the tumour tissue, which could potentially cause a disease. These nonsynonymous

| Process name                                | p value       | Number of genes associated | Number of variances found | Tumour tissue | Control tissue |
|---------------------------------------------|---------------|---------------------------|---------------------------|---------------|---------------|
| CD-SNVs                                     |               |                           |                           |               |               |
| Disorder of genitourinary system            | 9.05E-14      | 73                        | 135                       | 1             | 1             |
| Cell biology                               | 4.08E-04      | 69                        | 132                       | 1             | 1             |
| Cell signalling                            | 3.83E-03      | 25                        | 31                        | 1             | 1             |
| Morphology of body region                  | 2.55E-03      | 23                        | 24                        | 1             | 1             |
| Abnormal morphology of cells               | 1.73E-03      | 18                        | 19                        | 1             | 1             |
| Abnormal morphology of body cavity         | 6.17E-04      | 17                        | 18                        | 1             | 1             |
| Morphology of body cavity                  | 1.36E-03      | 17                        | 18                        | 1             | 1             |
| Morphology of cardiovascular system        | 5.80E-04      | 13                        | 14                        | 1             | 1             |
| Abnormal morphology of cardiovascular system | 7.22E-04     | 12                        | 13                        | 1             | 1             |
| Abnormal morphology of thoracic cavity     | 1.22E-03      | 11                        | 12                        | 1             | 1             |
| **Myelopoesis of bone marrow**             | **3.64E-03**  | **2**                     | **NPM1, RARA**            | 2             | 1             | 0             |

| Quantity of trabecular bone                | 4.08E-03      | 2: CREBBP, SMO            | 2                         | 1             | 0             |

| Small indels                               |               |                           |                           |               |               |
|Tissue development                          | 1.19E-03      | 5                         | 5                         | 1             | 0             |
|Developmental process of tissue             | 1.35E-03      | 5                         | 5                         | 1             | 0             |
|Development of organ                        | 5.05E-03      | 4                         | 4                         | 1             | 0             |
|Organogenesis                               | 5.32E-03      | 4                         | 4                         | 1             | 0             |
|Coloncy formation of tumour cell lines      | 6.25E-05      | 3                         | 3                         | 1             | 0             |
|Coloncy formation of cells                  | 4.44E-04      | 3                         | 3                         | 1             | 0             |
|Coloncy formation                           | 5.46E-04      | 3                         | 3                         | 1             | 0             |
|Developmental process of tumour cells       | 3.81E-03      | 3                         | 3                         | 1             | 0             |
|Coloncy formation of carcinoma cell lines   | 5.94E-05      | 2                         | 2                         | 1             | 0             |
|Apoptosis of nervous tissue cell lines      | 2.49E-04      | 2                         | 2                         | 1             | 0             |
|**Outgrowth of bone marrow cells**          | **7.56E-04**  | **1: ALK**                | **1**                     | 1             | 0             |

| Inflammatory response of bone marrow-derived macrophages | 1.26E-03 | 1: RELA | 1 | 1 | 0 |

The sorting is performed by number of genes. The bold data reflects the processes directly associated to bone.
mutations were located in *ESX1*, *CDC27*, *TMEM120B* and *TMEM131*. Additionally, in the case of *TMEM120B* and *TMEM131*, the mRNA expression has decreased substantially in the tumour tissue compared to that in the control tissue [10]; however, further studies are needed to confirm the possible associations between found mutations and gene expression level. Available data about the possible associations between OS and these genes is very limited. In *TMEM120B*, a gene with an unclear function, the mutation COSM1599921 has been previously detected in glioma [33]. The *CDC27* is a gene possibly controlling the timing of mitosis and may have an important role in tumour cell division [34]. In addition to the somatic mutation, the *CDC27* had 33 heterozygous germline disease-causing mutations (non-synonymous) (data not shown). In the case of breast cancer, the *CDC27* has been demonstrated to be a promising biomarker in predicting the disease progression and prognostication [35]. Thus, these somatic mutations may have some effect on OS pathogenesis. Especially the abundant changes in *CDC27* may be important in terms of regulating OS tumour cell division.

In the tumour tissue, we detected homozygous somatic small indels causing the frameshift in five genes—*EI24*, *ALG1L2*, *TIGD6*, *GPATCH4* and *SSPO*. None of these genes have previously been associated to OS, and according to our RNA-seq data, only *EI24* of these five genes has altered mRNA expression—it has decreased in the tumour tissue [10], which could be due to the insertion in exon 9. The *EI24* encodes a tumour suppressor and is an immediate-early induction target of TP53-mediated apoptosis—it binds to antiapoptotic BLC2. Furthermore, the *EI24* has found to be highly mutated in the case of aggressive breast cancer and is rather

| Disease name                           | p value   | Number of genes associated | Number of variances found | Tumour tissue | Control tissue |
|----------------------------------------|-----------|----------------------------|---------------------------|---------------|----------------|
| CD-SNVs                                |           |                            |                           |               |                |
| Cancer                                 | 7.04E-23  | 111                        | 202                       | 1             | 1              |
| Tumourigenesis                         | 8.21E-16  | 111                        | 202                       | 1             | 1              |
| Cancers and tumours                    | 3.37E-15  | 111                        | 202                       | 1             | 1              |
| Organismal injury and abnormalities    | 9.45E-17  | 105                        | 194                       | 1             | 1              |
| Carcinoma                              | 3.46E-25  | 99                         | 186                       | 1             | 1              |
| Solid tumour                           | 2.64E-24  | 99                         | 186                       | 1             | 1              |
| Epithelial neoplasia                   | 3.34E-23  | 99                         | 186                       | 1             | 1              |
| Epithelioma                            | 3.34E-23  | 99                         | 186                       | 1             | 1              |
| Breast or colorectal cancer            | 5.45E-23  | 83                         | 164                       | 1             | 1              |
| Malignant neoplasm of abdomen          | 6.93E-20  | 83                         | 169                       | 1             | 1              |
| Bone marrow cancer                     | 1.69E-03  | 43                         | 1                         | 0             |                |
| Bone marrow cancer and tumours         | 1.69E-03  | 43                         | 1                         | 0             |                |

Small indels

| Disease name                           | p value   | Number of genes associated | Number of variances found | Tumour tissue | Control tissue |
|----------------------------------------|-----------|----------------------------|---------------------------|---------------|----------------|
| Cancer                                 | 9.07E-03  | 6                          | 6                         | 1             | 1^a            |
| Hematologic cancer                     | 2.36E-04  | 4                          | 4                         | 1             | 1^a            |
| Hematologic cancer and tumours         | 2.36E-04  | 4                          | 4                         | 1             | 1^a            |
| Hematological neoplasia                | 8.01E-04  | 4                          | 4                         | 1             | 1^a            |
| Lymphohematopoietic cancer             | 9.12E-04  | 4                          | 4                         | 1             | 1^a            |
| Disease of colon                       | 7.88E-03  | 4                          | 4                         | 1             | 0              |
| Hematological disease                  | 8.15E-03  | 4                          | 4                         | 1             | 1^a            |
| Immunological disease                  | 1.28E-02  | 4                          | 4                         | 1             | 1^a            |
| Gastrointestinal tract cancer          | 2.0E-02   | 4                          | 4                         | 1             | 0              |
| Gastrointestinal tract cancer and tumours | 2.0E-02 | 4                          | 4                         | 1             | 0              |
| Tumourigenesis of bone tumour          | 7.04E-03  | 1                          | 1                         | 0             |                |

^aHere, only one gene PRR23C has a small indel in heterozygous form, which most likely does not affect the gene function. See Table 2. The bold data reflects the diseases directly associated to bone.
associated to tumour invasiveness than development of the primary tumour [36-38]. In the present case, we found no mutations in TP53 nor was the expression altered [10]; thus, according to this data, we may suggest that the TP53 is functional in the tumour tissue. However, the TP53 pathway may still be suppressed due to mutated and downregulated E2F4. Moreover, the aggressive nature of OS is correlated to this finding.

Appling Ingenuity Variant Analysis software, we found over 200 cancer driver variants and 93% of these possibly
cause the loss of gene function. Thirteen homozygous somatic CD-SNVs were detected in different genes—RGPD3, PRDM9, FOXX1, CCZ1, PLAT, AGTPBP1, SARDH, FAH, CDC27, SBF1, LRRC37A3, ARL17A and LILRB3. The mRNA expression of PLAT, AGTPBP1 and LRRC37A3 has increased and of SFB1 has decreased significantly [10]. We found no previous data about the associations between OS and these genes, except SBF1. With previous OS studies, another missense mutation (p.E1539K) has detected in SBF1 [39]. SBF1 is a SET (a nuclear oncogene) binding

Table 7 The integrative analysis—genes with altered expression pattern [10] and SNVs annotated with ANNOVAR software

| Gene name          | Transcript name—exon number: nucleotide change/amino acid change | ljb2 score/indel | Chr number | Start   | End     | REF/ALT | logFC | FDR   |
|--------------------|---------------------------------------------------------------|------------------|------------|---------|---------|---------|-------|-------|
| **Germline mutations homozygous in tumour tissue** | | | | | | | | |
| STEAP4 NM_024636—exon2: c.G364A/p.A122T | | 0.647 | Chr7 | 87913221 | 87913221 | C/T | 3.015 | 1.44E-19 |
| NM_001205316—exon2: c.G364A/p.A122T | | | | | | | | |
| NM_001205315—exon3: c.G364A/p.A122T | | | | | | | | |
| DDX60L NM_001012967—exon18: c.T2491C/p.C831R | | 0.711 | Chr4 | 169341435 | 169341435 | A/G | 2.349 | 2.67E-14 |
| MT1A NM_005946—exon3: c.A152G/p.K51R | | 0.785 | Chr16 | 56673828 | 56673828 | A/G | −3.094 | 0.00795 |
| ACOX1 NM_004035—exon7: c.C936G/p.I312M | | 0.872 | Chr17 | 73949540 | 73949540 | G/C | −0.809 | 0.01538 |
| NM_001185039—exon7: c.C936G/p.I312M | | | | | | | | |
| NM_007292—exon7: c.C936G/p.I312M | | | | | | | | |
| NM_001185039—exon7: c.C822G/p.I274M | | | | | | | | |
| TMC7 NM_001160364—exon6: c.G431A/p.G144E | | 0.695 | Chr16 | 19041595 | 19041595 | G/A | 1.266 | 0.01726 |
| NM_0024847—exon6: c.G761A/p.G254E | | | | | | | | |
| MYO7A NM_001127179—exon27: c.3514_3535del/p.1172_1179del | Frameshift deletion | Chr11 | 76895771 | 76895792 | GGAGGC/GGGGC | 1.541 | 0.03810 |
| ATRNL1 NM_001276282—exon8: c.1399_1400insTT/p.L467fs | Frameshift insertion | Chr10 | 116931101 | 116931101 | /TT | 2.321 | 0.04535 |
| TMEM120B NM_001080825—exon3: c.G274A/p.D92N | | 0.981 | Chr12 | 122186317 | 122186317 | G/A | −1.548 | 0.00064 |
| →X→COSM15999921 | | | | | | | | |
| TMEM131 NM_015348—exon31: c.C3947T/p.H3792L | | 0.945 | Chr2 | 98409046 | 98409046 | G/A | −0.799 | 0.01371 |
| EL24 NM_001007277—exon9: c.733dupC/p.R244fs | Frameshift insertion | Chr11 | 125452300 | 125452300 | /C | −0.815 | 0.01569 |

These germline or somatic SNVs are all nonsynonymous and homozygous in the tumour tissue and according to ljb2 database have a disease-causing effect.
Table 8 The integrative analysis—CNVs and RNA expression data [10] is observed together

| Gene name | Chr number | Start | End | Area length | CNV p value | Copy number fold change | logFC | FDR       |
|-----------|-------------|-------|-----|-------------|-------------|-------------------------|-------|-----------|
| **Loss**  |             |       |     |             |             |                         |       |           |
| INSR      | Chr19       | 7119459 | 7294011 | 174,552     | 3.18E-11    | 9.32E-03                | 9.67E-31 |
| NFIX      | Chr19       | 13065583 | 13201204 | 94,621      | 0           | 8.44E-12                | 1.63E-17 |
| FARA      | Chr19       | 13034964 | 13044558 | 9,594       | 0           | 8.12E-11                | 1.96E-16 |
| RADD3A    | Chr19       | 13056627 | 13063667 | 7,040       | 0           | 8.12E-11                | 4.46E-16 |
| GINS4     | Chr8        | 41386724 | 41399418 | 12,694      | 0           | 8.12E-11                | 2.79E-15 |
| GADD45G1  | Chr19       | 13064971 | 13068050 | 3,079       | 0           | 8.12E-11                | 3.67E-15 |
| IFI1      | Chr2        | 163123358 | 163175218 | 51,630     | 0           | 8.12E-11                | 3.69E-14 |
| RPL31     | Chr2        | 101618690 | 101622885 | 4,195       | 0           | 8.12E-11                | 5.08E-13 |
| PLEKHB4   | Chr5        | 156185  | 181790 | 25,605      | 0           | 8.12E-11                | 1.58E-12 |
| ZNF358    | Chr19       | 7581003 | 7581135 | 132         | 3.18E-11    | 2.25E-12                | 2.56E-11 |
| ARHGEF18  | Chr19       | 7459998 | 7532004 | 72,006      | 0           | 8.12E-11                | 1.44E-10 |
| STX10     | Chr19       | 13255223 | 13260987 | 57,64       | 0           | 8.12E-11                | 1.96E-10 |
| COLS5A    | Chr19       | 10102679 | 10123117 | 18,468      | 0           | 8.12E-11                | 5.45E-10 |
| MGAT4A    | Chr2        | 99242185 | 99347589 | 105,404     | 0           | 8.12E-11                | 9.95E-10 |
| **Gain**  |             |       |     |             |             |                         |       |           |
| SLC40A1   | Chr2        | 190428309 | 190428951 | 642         | 1.83E-05    | 4.28                    | 1.05E-14 |
| KIT       | Chr4        | 55524094 | 55603446 | 79,352      | 1.69E-06    | 4.79                    | 1.17E-13 |
| PTPLAD2   | Chr9        | 21008019 | 21031635 | 23,616      | 7.73E-14    | 7.48                    | 4.49E-13 |
| ATPB41    | Chr4        | 42571177 | 42629126 | 75,949      | 1.83E-07    | 5.22                    | 4.54E-10 |
| FOCA1     | Chr9        | 20658308 | 20993327 | 335,019     | 7.73E-14    | 7.48                    | 9.12E-08 |
| FAM200B   | Chr4        | 15683351 | 15692070 | 8,719       | 3.54E-05    | 4.14                    | 8.59E-07 |
| SLIT2     | Chr4        | 20255234 | 20512189 | 256,955     | 4.16E-05    | 4.10                    | 1.73E-05 |
| MLLT3     | Chr9        | 20353522 | 20622514 | 268,992     | 7.73E-14    | 7.48                    | 2.19E-05 |
| LCORL     | Chr4        | 17887690 | 18023483 | 135,793     | 4.16E-05    | 4.10                    | 5.04E-05 |

Only the genes with lowest FDR value are presented.

Table 9 The integrative analysis - loss of heterozygosity and RNA expression data observed together

| Gene name | Chr number | LOHs | LOH position | Alleles | LOH | LOH p value | logFC | FDR       |
|-----------|-------------|------|--------------|---------|-----|-------------|-------|-----------|
| MS4A14    | Chr11       | 60165358–60165379 | G/C   | CopyNeutralLOH | 0.025 | 2.46 | 3.20E-08 |
| DSC2      | Chr18       | 28665654–28665556 | A/C   | CopyNeutralLOH | 0.025 | 1.87 | 3.82E-07 |
| RPS4X     | ChrX        | 71495409–71495414 | G/C   | CopyNeutralLOH | 0.01  | -1.44 | 7.25E-07 |
| RPS23     | Chr5        | 81571874           | A/C   | CopyNeutralLOH | 0.005 | -1.43 | 1.04E06  |
| IL7R      | Chr5        | 35874575           | C/T   | 1AlleleGain    | 0.025 | 1.59 | 6.69E-06 |
| PCNXL2    | Chr1        | 233398713          | C/T   | CopyNeutralLOH | 0.01  | 1.20 | 0.00027  |
| HILPDA (C7orf68) | Chr7 | 128098270    | T/G   | CopyNeutralLOH | 0.0001 | -1.12 | 0.00094  |
| HRNR      | Chr1        | 152188041          | C/T   | Allele(s)Loss  | 0.025 | -3.09 | 0.00796  |
| MUC4      | Chr3        | 195515594, 195516630 | C/G   | CopyNeutralLOH | 0.025 | -2.22 | 0.01230  |

Only the genes with significant mRNA expression changes in the tumour tissue compared to that in the control tissue are presented.
factor 1 and may inhibit the cell division [40]. The decreased expression in the tumour tissue may be responsible for the increased cell proliferation. Some other associations, which might be interesting—PLAT gene is important for cell migration and tissue remodelling and the overexpression might cause hyperfibrinolysis [41], which has not previously described in the case of OS. Two mutations in ARL17A have detected in chondrosarcoma cells [42]. In the case of CDC27, the same mutation (p.E6G) was also brought front by AS as potentially disease causing, which is discussed above. Thus, it is highly likely that at least some of these genes participate in some level of OS pathogenesis.

Additionally, with IVA four homozygous somatic small indels were detected in the tumour tissue. These were in noncoding regions of genes CTCFL, PRR23C, CDCA7L and ALK; thus, the effect might be post-transcriptional. CTCFL is a genetic paralog of CTCF; latter is an important methylation pattern regulator. In the case of CTCF, it has previously demonstrated that in the OS tissue, the changes in its methylation pattern may also cause loss of imprinting of IGF2 and H19 genes, which further alters their expression pattern [43]. In our OS patient's tumour tissue, the mRNA expression of both IGF2 and H19 has increased significantly (FDR = 3.46E-15 and FDR = 0.0015, respectively) [10]. Thus, the association may be valid here also. In PRR23C, one nonsense mutation (p.R190W) has detected previously in the OS tissue [42]. ALK encodes a receptor tyrosine kinase and is rearranged, mutated or amplified in several tumours. However, in the case of OS, there are only few reports about ALK [44,45]. In addition, two heterozygous somatic small indels were detected in DSPP and RELA exons; however, we found no previous data about these findings and associations to OS. The small indels might have an effect on the expression of these genes both pre- and post-transcriptional level; however, these suggestions need to be further studied.

According to IVA, there were several genes with more than one mutation—in MUC4, there were even 22 somatic mutations in exons and 44 in introns, although they all were heterozygous. Thus, we found MUC4 locus to be the most altered in the tumour tissue compared to that in the control tissue. This might explain why its mRNA expression in the tumour tissue has decreased (FDR = 0.012) [10]. Mucin 4 is among major constituents of mucus, and it has demonstrated that primary bone tumours rarely express MUC4 protein [46], which correlates to our finding. Furthermore, with IVA, we found mucin complex (MUC2, MUC4, MUC6) to have a highest significance in OS among others. However, there are also other mucin genes (MUC16, MUC17, MUC20) with somatic heterozygous CD-SNVs. The expression pattern of all other detected mucins has not changed significantly. Thus, mucins may have a role in OS pathogenesis, but we dear not to make any further conclusions.

With IVA, there was four bone-related processes brought front only in the case of the tumour tissue—“myelopoesis of bone marrow” (NPM1, RARA), “quantity of trabecular bone” (CREBBP, SMO), “outgrowth of bone marrow cells” (ALK) and “inflammatory response of bone marrow-derived macrophages” (REL). Furthermore, in disease list, 16 genes with over 40 somatic variations were associated to “bone marrow cancer” and “bone tumour”; however, there were also over 200 germline CD-SNVs associated to cancer. Thus, here, we would like to emphasize that in the case of both cancer-associated processes and diseases, the ones associated with bone are somatic mutations; however, the findings possibly promoting cancer are germline mutations. This is one of the phenomena, which we would like to observe in our future studies.

The most significant pathway found with IVA was “WNT/β-catenin signalling pathway” (altered genes: CREBBP, RARA, SMO, SOX10, SOX7, TCF3). Reviewed in [15], the pathway is required for bone development and has demonstrated to be altered in pathogenesis of OS—overexpression of numerous WNT pathway components including WNT ligands, FZDs and LRP receptors and epigenetic silencing of the pathway inhibiting genes, i.e. WIF1. However, in our previous study, we found WNT7B and WNT11 to be downregulated and WNT2B and WNT5B upregulated; FZD4 and FZD8 upregulated and LRP8 and LRP12 downregulated and DVL3 downregulated and WIF1 and SOST upregulated. Additionally, genes with CD-SNVs—RARA, SMO and SOX7 were upregulated [10]. Thus, our results are rather controversial to several previous studies demonstrating the WNT/β-catenin pathway to be upregulated [47-49]. However, there are also studies correlating to our findings [50,51]. As our study is based on a single case, we dear not to conclude, why the WNT/β-catenin pathway is rather downregulated here, but we suggest the controversial results may occur due to major heterogeneity of OS. Nevertheless, the present study demonstrates that in addition to altered expression pattern, the genes involved in WNT/β-catenin signalling pathway carry the CD-SNVs.

In the case of small indels, the IVA brought front the pathways associated to RELA and these are mostly cytokine signalling pathways (Table 6). Previously, it has demonstrated that interaction of IL17A and IL17AR promotes metastasis in OS cells. Furthermore, IL17 stimulates osteoclast resorption [52]. In our previous study, we found IL17AR to be significantly upregulated [10]. Osteoclasts are important in pathogenesis of OS—the more active they are, the more aggressive the tumour is [53]. RELA is demonstrated to enhance the osteoclast differentiation [54]. As IVA predicts the loss of RELA functionality (at least partially, as the small indel is heterozygous), in the present case, the OS might not have been as aggressive as it usually would.
Previously, it has demonstrated that chromothripsis event is common to early stage of OS—hundreds of genomic rearrangements will appear in a single instability event [26]. In the present case, the CEQer software detected nearly 2,400 gain and loss events in 8 chromosomes involved, which should qualify as the chromothripsis. However, the initiating cause of this massive rearrangement is unknown, as there were no traumas or other environmental causes we are aware of.

In general, the gain of copy number should increase the mRNA expression and loss of copy number should decrease the expression [6]. In present work, this pattern was valid in the case of 86.5% of the genes with CNVs and altered expression. One of the strongest findings here was the amount of CNVs in INSR, which expression has decreased remarkably (Table 8). The main physiological role of the insulin receptor appears to be metabolic regulation [55]. However, together with IGFI1R it forms a hybrid receptor for IGFI1, latter together with IGF2 is thought to have a key role in driving the proliferation and survival of sarcoma cells [56]. Furthermore, the growth hormone and IGFI1 axis controls the growth and bone modelling/remodelling [57]. Additionally, the IRS1, which is phosphorylated by the INSR, is important for both metabolic and mitogenic pathways [58]. In the present case, the mRNA expression of both IGFI1 and IGF2 has increased (FDR = 4.65E-35 and FDR = 3.46E-15, respectively); however, the expression of IGFI1R remained the same in the tumour tissue compared to that in the control tissue [10]. Furthermore, in IGFI1R we found a heterozygous germline nonsynonymous mutation (p.G1117R) with AS, which according to ljb2 database is a disease causing (data not shown). Similarly to INSR, the mRNA expression of IRS1 is decreased in the tumour tissue compared to that in the control tissue (FDR = 2.62E-10) [10]. Thus, in the present case it seems, the proliferation of tumour cells might be rather supported by increased effect of IGFI1, IGF2 and IGFI1R homodimer associations, than IGFI1, IGF2 and INSR-IGFI1R heterodimer associations or INSR effects on IRS1.

The loss of heterozygosity has been reported to be extensive in OS exomes [39]. In the present case, we did not detect whole chromosome or gene region loss; however, we did detect the loss of heterozygosity in smaller regions. The genes with LOH findings and increased mRNA expression—MS4A14, DSC2, IL7R and PCNXL2 have not associated to OS previously. However, in the case of DSC2, the overexpression has demonstrated to be inversely correlated to bone metastasis-free survival [59]. The mutations in IL7R exon 6 have been demonstrated to be present in leukaemia patients’ bone marrow samples but not associated to other solid tumours [60]. The five genes with LOHs and decreased mRNA expression—RPS4X, RPS23, HILPDA (C7orf68), HRNR and MUC4 also do not have previous information associated to OS. Nonetheless, also the LOH analysis brought forward different genes in mucin family. In addition to MUC4, there were also other genes with LOHs but with insignificant mRNA expression changes in the tumour tissues—MUC2, MUC6 and MUC17. Thus, these results also support the idea that mucins might have a role in pathogenesis of osteosarcoma.

In summary, the present case has several characteristics previously demonstrated in OS. The wide genomic arrangements have appeared—SNVs and small indels all over the genome and CNVs in some chromosomes; and in several cases, these rearrangements may have an effect on gene expression. Furthermore, the germline mutations seem to be associated to cancer in general and somatic mutations to bone tumours. The most significant pathway was the one probably most thoroughly studied in the case of OS—the WNT/β-catenin signalling pathway. We found several genes in this pathway carrying the cancer driver variances. Additionally, the IGFI1/IGF2 and IGFI1R homodimer signalling might have an essential effect on OS pathogenesis. Which also needs to be emphasized is that according to our data (based on DNA and RNA studies), there is no evidence of a non-functional TP53; however, the TP53 pathway might be suppressed in further levels—the downregulation of EI24. In addition, with this study, we found associations between different genes and OS pathogenesis, which have not demonstrated before in earlier studies. We found the MUC4 locus to be the most altered in the tumour tissue compared to that in the control tissue; furthermore, several other mucin genes are also possibly associated to OS. The somatic mutation in CDC27 was brought front by two different data analysis softwares and might have a role in OS pathogenesis.

Conclusions
All genes, in which the mutations were detected, may be considered as potential targets for additional studies (i.e. functional, histopathological, clinical studies) for finding OS biomarkers. The present study brought front the WNT pathway genes, IGFI1/IGF2 and IGFI1R homodimer signalling pathway genes, TP53 together with EI24, MUC4 together with other mucin genes and CDC27 as potential biomarkers for OS. Finally, as this study is based on a single case and only DNA and RNA analysis, these data may not be taken as conclusive evidence and further studies are needed to confirm the present findings.

Additional files

Additional file 1: ANNOVAR software. The file contains the list of SNVs (coverage at least 20 times) and small indels detected from WES study.
Osteosarcoma (osteogenic sarcoma).
The etiology of osteosarcoma.

Additional file 2: Ingenuity Variant Analysis software. The file contains the list of SNVs, which according to IVA are associated to cancer. Additionally, the dbSNP135, SIFT and POLYPHEN functions and RNA-seq information is added if available.

Additional file 3: CEGer software. The file contains the list of genes where CNVs and LOHs were detected. Additionally, the RNA-seq information is added if available.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
ER participated in the research concept and design, data analysis and interpretation, writing of the article and critical revision of the article. KM participated in the research concept and design, collection and/or assembly of data, data analysis and interpretation and critical revision of the article. AM involved in the research concept and design, collection and/or assembly of data, data analysis and interpretation and critical revision of the article. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by the institutional research funding IUT (IUT20-46) of the Estonian Ministry of Education and Research, by the Centre of Tranlational Genomics of University of Tartu (SP1i2GARENG) and by the European Regional Development Fund (Centre of Translational Medicine, University of Tartu).

Author details
1Department of Pathophysiology, University of Tartu, 19 Ravila Street Tartu 50411, Estonia.
2Department of Reproductive Biology, Estonian University of Life Sciences, 64 Kreutzwaldi Street, Tartu, Estonia.
3Department of Oncology, Hue University of Medicine and Pharmacy, 6 Ngo Quyen Street, Hue, Vietnam.
4Traumatology and Orthopaedics Clinic, Tartu University Hospital, 8 Puusepa Street, Tartu, Estonia.
5Department of Reproductive Biology, Estonian University of Life Sciences, 64 Kreutzwaldi Street, Tartu, Estonia.
6Department of Physiology, University of Tartu, 8 Puusepa Street, Tartu, Estonia.

Received: 30 July 2014 Accepted: 10 November 2014 Published online: 11 December 2014

References
1. Picci P: Osteosarcoma (osteogenic sarcoma). Orphanet J Rare Dis 2007, 2:6.
2. Ottaviani G, Jaffe N: The etiology of osteosarcoma. Cancer Treat Res 2009, 152:15–32.
3. Botter SM, Neri D, Fuchs B: Recent advances in osteosarcoma. Curr Opin Pharmacol 2014, 14C:15–23.
4. Allison DC, Canney SC, Ahlmann ER, Hendifar A, Chavlia S, Fedenko A, Angeles C, Menendez LR: A meta-analysis of osteosarcoma outcomes in the modern medical era. Sarcoma 2012, 2012:704872. doi:10.1155/2012/704872.
5. Kiezun A, Janeway K, Toraci P, Mora J, Aguilar S, Mercado G, Melendez J, Garraway L, Rodriguez-Galindo C, Orkin S, Golub T, Getz G, Yen CC, Chen WM, Chen TH, Chen WY, Chen PC, Chiou HJ, Hung GY, Wu HT, Wei CJ, Shiau CY, Wu YC, Chao TC, Tzeng CH, Chen PM, Lin CH, Chen YJ, Fletcher JA: Identification of chromosome aberrations associated with disease progression and a novel 3q13.31 deletion involving LSAMP gene in osteosarcoma. Int J Cancer 2009, 53(4):775–788.
6. Märtson A, Köks S, Reimann E, Prans E, Erm T, Maasalu K: Transcriptome association study identifies two susceptibility loci for osteosarcoma. Nat Genet 2013, 45(7):799–803.
7. Ottaviani L, Schaefer KL, Gajewski M, Huckenbeck W, Baldus S, Rogul U, Mackintosh C, de Alava E, Myklebost O, Kresse SH, Meza-Zepeda LA, Serra M, Clenet-Jansen AM, Hogendoorn PC, Buerger H, Aigner T, Gabbert HE, Poremba C: Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoneffet effort. Genes Chromosomes Cancer 2010, 49(1):40–51.
8. Kansara M, Tsang M, Kodjabachian L, Sims NA, Trivett MK, Ehrlich M, Dobrovic A, Slavin J, Choong PF, Simmons PJ, Dawid IB, Thomas DM: Wnt inhibitory factor 1 is epigenetically silenced in human osteosarcoma, and targeted disruption accelerates osteosarcomagenesis in mice. J Clin Invest 2009, 119(6):837–851.
9. Märtson A, Köks S, Reimann E, Prans E, Erm T, Maasalu K: Transcriptome analysis of osteosarcoma identifies suppression of wnt pathway and up-regulation of adipopectin as potential biomarker. Genomic Medicine 2013, 1:1–9.
10. Namlos HM, Meza-Zepeda LA, Barot Y, Ostensen IH, Kresse SH, Kuijjer ML, Serra M, Burger H, Clenet-Jansen AM, Myklebost O: Modulation of the osteosarcoma expression phenotype by microRNAs. PLoS One 2012, 7(10):e48086.
11. Hingorani P, Zhang W, Gorlick R, Kolb EA: Inhibition of Src phosphorylation alters metastatic potential of osteosarcoma in vitro but not in vivo. Clin Cancer Res 2009, 15(10):3416–3422.
12. Yap TA, Arkenau HT, Camidge DR, George S, Sarkova N, Gywether SJ, Spratlin JL, Lai R, Spicer J, Desouza NM, Leach CH, Mick J, Fazzari M, Boinpally R, Gedrich R, Brock K, Stephens A, Eckhardt SG, Kaye SB, Demetri G, Scir M: First-in-human phase I trial of two schedules of OSI-930, a novel multikinase inhibitor, incorporating translational proof-of-concept studies. Clin Cancer Res 2013, 19(4):399–419.
13. Li Z, Wang C, Yuan T, Liu Y, Song T, Liu Y, Chen C, Yang M, Tang Z, Shi Q, Weng Y: Bone morphogenetic protein 9 regulates tumor growth of osteosarcoma cells through the Wnt/beta-catenin pathway. Oncol Rep 2014, 31(2):899–904.
14. Cai Y, Cai T, Chen Y: Wnt pathway in osteosarcoma, from oncogenic to therapeutic. J Cell Biochem 2014, 115(6):625–631.
15. Guo M, Cai C, Zhao G, Qiu X, Zhao H, Ma Q, Tian L, Li X, Hu Y, Liao B, Ma B, Fan H: Hypoxia promotes migration and induces CXCR4 expression via HIF-1alpha activation in human osteosarcoma. PLoS One 2014, 9(9):e105518.
16. Raymond A, Ayala A, Knuutila S: Conventional osteosarcoma, genetics. Lyon: IARC Press; 2002.
17. Lai CC, Harris CP, Lu XY, Perlaky L, Gogineni S, Chintagumpala M, Hicks J, Johnson ME, Davino NA, Huvos AG, Meyers PA, Healy JH, Gorlick R, Rao PH: Frequent amplification and rearrangement of chromosomal bands 6p12-p21 and 17p11.2 in osteosarcoma. Genes Chromosomes Cancer 2004, 39(1):11–21.
18. Kresse Sh, Ohnstad HO, Paulsen EB, Bjerkenehn B, Szuha K, Serra M, Schaefer KL, Myklebost O, Meza-Zepeda LA: LSAMP, a novel candidate tumor suppressor gene in human osteosarcomas, identified by array comparative genomic hybridization. Genes Chromosomes Cancer 2003, 48(5):679–693.
19. Park S, Shin A, Durbin AD, Stavropoulos DJ, Baskin B, Ray PN, Novokmet A, Mallin D: 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 Res 2010, 70(1):160–171.
20. Tarkkanen M, Karhu R, Kallioniemi A, Elomaa I, Kivioja M, Kuijjer ML, Bohling T, Karaharju E, Hyytiainen J, Knuutila S: Gains and losses of DNA sequences in osteosarcomas by comparative genomic hybridization. Cancer Res 1995, 55(6):1334–1338.
21. Tarkkanen M, Elomaa I, Blomqvist L, Kivioja AH, Lellouche-Lehtinen P, Bohling T, Valle J, Knuutila S: DNA sequence copy number increase at 8q: a potential new prognostic marker in high-grade osteosarcoma. Int J Cancer 1999, 84(2):114–120.
22. Ovehtolerz Z, Ostman Z, Favis R, Elowicz MB, Barany F, Ladanyi M, Gorlick R, Levine AJ: The presence of p53 mutations in human osteosarcomas correlates with high levels of genomic instability. Proc Natl Acad Sci U S A 2003, 100(20):11547–11552.
23. Yen CC, Chen WM, Chen TH, Chen WY, Chen PC, Chiou HJ, Hung GY, Wu HT, Wei CJ, Shiao CY, Wu YC, Chao TC, Tzeng CH, Chen PM, Lin CH, Chen YJ, Fletcher JA: Identification of chromosome aberrations associated with disease progression and a novel 3q13.31 deletion involving LSAMP gene in osteosarcoma. Int J Cancer 2009, 53(4):775–788.
25. Lu XY, Lu Y, Zhao YJ, Jiao YQ, Kang K, Xiao-Nan L, Ge G, Meyer R, Perilksy L, Hicks J, Chinagumpala M, Cai WW, Ladanyi M, Gorlick R, Lau CC, Pati D, Sheldon M, Rao PH. Cell cycle regulator gene CDCL5: a potential target for 6p12-p21 amplicon in osteosarcoma. MCR 2008, 6(6):937–946.

26. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Bearde D, Stebbings LA, McLaren S, Lin ML, McBride DJ, Varela I, NK-Zainal S, Leroy C, J A, Meneses A, Butler AP, Teague JW, Quail MA, Burton J, Swedlow H, Carter NP, Mosriberger LA, Iacobuzio-Donahue C, Follows GA, Green AR, Flanagan AM, Stratton MR, et al: Massive genomic rearrangement in acquired in a single catastrophic event during cancer development. Cell 2011, 144(1):277–283.

27. Wang K, Li M, Hakonson H: ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010, 38(16):e164.

28. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Zimin AV, Sunyaev SR: A method and server for predicting damaging missense mutations. Nat Methods 2010, 7(4):248–249.

29. Gonzalez-Perez A, Lopez-Bigas N: Improving the assessment of the outcome of nonsynchronous SNVs with a consensus deleteriousness score. Condis Am J Human Genetics 2011, 88(4):469–469.

30. Kumar P, Henkoff S Jr, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009, 4(7):1073–1081.

31. Blankenberg D, Von Kuster G, Conaco N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J, Galaxy: a web-based genome analysis tool for experimentalists. Current protocols in molecular biology edited by Frederick M. Ausubel et al. 2010, Chapter 19:Unit 19.10.1:1–21.

32. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, separat mais A, McLysaght J, Rosenbloom K, et al: Genome: a platform for interactive large-scale genome analysis. Genome Res 2005, 15(10):1451–1457.

33. Yost SE, Pastelino S, Rozerhalah S, Smith EN, Jiang P, Kesari S, Pastelino S, Rozerhalah S, Smith EN, Jiang P, Kesari S, Kumar P, Harismendy O: Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score. Nat Methods 2010, 7(4):248–249.

34. Ma Y, Ren Y, Han EQ, Li H, Chen D, Jacobs JJ, Garetto SM, Pittman MA, Garetto SM, Pittman MA, Strome SE, Lowik CW: Low cdc27 and high securin expression predict breast cancer invasiveness. Cancer Res 2011, 71(23):7479–7489.

35. Talvinen K, Karra H, Pitkanen R, Ahonen I, Nykanen M, Lintunen M, Soderstrom M, Sheldon M, Rao PH: low cdc27 and high securin expression predict breast cancer invasiveness. Cancer Res 2011, 71(23):7479–7489.

36. Zhao YG, Zhao H, Miao L, Wang L, Sun F, Zhang H: Immunohistochemical expression of corticotropin-releasing hormone (CRH) and its receptors in human pituitary adenoma. Endocr Regul 2001, 35(3):193–202.

37. Gu Z, Flemington C, Chittenden T, Zambetti GP: Interaction of Src homology-2 domain with the C-terminal region of cyclin H. J Biol Chem 2010, 285(28):22815–22822.

38. Burton J, Swerdlow H, Carter NP, Morsberger LA, Iacobuzio-Donahue C, Egeler RM, Gelderblom H, Taminiau AH, Hopko AM: High-resolution mutational profiling suggests a systems biology approach reveals common metastatic pathways in osteosarcoma. BMC Syst Biol 2011, 5(2):303–307.

39. Choy E, Hornicek F, MacConaill L, Harmon D, Tariq Z, Garraway LA, Lando D: High-throughput genotyping in osteosarcoma identifies multiple mutations in phosphoinositide-3-kinase and other oncogenes. Cancer 2012, 118(11):2905–2914.

40. Tirabosco R, Berisha F, Ye H, Halai D, Amay F, Flanagan AM: Assessment of MUC4 expression in primary bone tumors. Nippon 2013, 63(1):142–143.

41. Flores RJ, Li Y, Yu A, Shen J, Rao PH, Lau SS, Vannucci M, Lau CC, Man TK: A systems biology approach reveals common metastatic pathways in osteosarcoma. BMC Syst Biol 2012, 6:50.

42. Ma Y, Ren Y, Han EQ, Li H, Chen D, Jacobs JJ, Garetto SM, Pittman MA, Garetto SM, Pittman MA, Strome SE, Lowik CW: Low cdc27 and high securin expression predict breast cancer invasiveness. Cancer Res 2011, 71(23):7479–7489.

43. Ulaner GA, Vu TH, Li T, Hu JF, Yao XM, Yang Y, Gorlick R, Meyers P, Healey J, Ladanyi M: A major cartilage tumor suppressor gene is inactivated in osteosarcoma. Cancer Res 2003, 63(16):4576–4584.

44. Reimann et al. Human Genomics 2014, 8:20

http://www.humgenomics.com/content/8/1/20

Page 16 of 16

Cite this article as: Reimann et al.: Whole exome sequencing of a single osteosarcoma case—integrative analysis with whole transcriptome RNA-seq data. Human Genomics 2014 8:20.