Title: Deep sequencing in conjunction with expression and functional analyses reveals activation of FGFR1 in Ewing sarcoma

Authors: Konstantin Agelopoulos†1, Günther HS Richter†*2, Eva Schmidt†1, Uta Dirksen3, Kristina von Heyking2, Benjamin Moser3, Hans-Ulrich Klein4, Udo Kontrny5, Martin Dugas4, Kathrin Poos6, Eberhard Korschling6, Thorsten Buch7, Matthias Weckesser3, Isabell Schulze1,14, Regina Besoke9, Anika Witten10, Monika Stol10, Gabrielle Köhler11, Wolfgang Hartmann12, Eva Wardelmann12, Claudia Rossig3, Daniel Baumhoer13, Heribert Jürgens3, Stefan Burdach2, Wolfgang E. Berdel#1, Carsten Müller-Tidow#*1,14

Affiliations: †Department of Medicine A, Hematology, and Oncology, University Hospital of Muenster, 48149 Muenster, Germany, 2Children's Cancer Research Center and Department of Pediatrics, Roman Herzog Comprehensive Cancer Research Center and Klinikum rechts der Isar, Technische Universität München, 81664 Munich, together with the German Cancer Consortium (DKTK), Germany, 3Department of Pediatric Oncology and Hematology, University Children's Hospital Muenster, 48149 Muenster, Germany, 4Institute of Medical Informatics, University of Muenster, 48149 Muenster, Germany, 5Department of Pediatrics, Medical Faculty, RWTH Aachen University, 52074 Aachen, Germany 6Institute of Bioinformatics, University Hospital of Muenster, 48149 Muenster, Germany 7Institute for Medical Microbiology, Immunology and Hygiene, Technische Universität München, 81675 Munich, Germany, Institute of Laboratory Animal Science, University of Zurich, 8091 Zurich, Switzerland 8Department of Nuclear Medicine, University of Muenster, 48149 Muenster, Germany, 9Institute of Human Genetics, University of Muenster, 48149 Muenster, Germany, 10Institute of Human Genetics, Genetic Epidemiology, University of Muenster, 48149 Muenster, Germany, 11Institute of Pathology, Klinikum Fulda, Fulda, Germany, 12Department of Pathology, University of Muenster, 48149 Muenster, Germany, 13Bone Tumor Reference Center at the Institute of Pathology, University Hospital Basel, 4031 Basel, Switzerland, 14Department of Medicine IV, Hematology and Oncology, State Center for Cell and Gene Therapy, University Hospital Halle, Halle (Saale), Germany

†KA, GR and ES are joint first authors, # WEB and CMT are joint senior authors

Corresponding authors: Carsten Müller-Tidow and Günther Richter

Prof. Dr. Carsten Müller-Tidow, University Hospital Halle, Department of Medicine IV, Hematology and Oncology, Ernst-Grube-Str. 40, 06120 Halle (Saale), Germany
Phone: 0049-345-557-4972; Fax: 0049-345-557-2950, e-mail: carsten.mueller-tidow@uk-halle.de

PD Dr. Günther Richter, Technische Universität München, Children's Cancer Research Center and Department of Pediatrics, Koelner Platz 1, 80804 Munich
Tel: 0049-89-3068 3235, Fax: 0049-89-3068 3791, e-mail: guen-ther.richter@lrz.tum.de

Key words: Ewing Sarcoma, next generation sequencing, FGFR1
Abstract

**Purpose:** A low mutation rate seems a general feature of pediatric cancers, in particular in oncofusion gene driven tumors. Genetically, Ewing sarcoma (ES) is defined by balanced chromosomal EWS/ETS translocations, which give rise to oncogenic chimeric proteins (EWS-ETS). Other contributing somatic mutations involved in disease development have only been observed at low frequency.

**Experimental design:** Tumor samples of 116 Ewing sarcoma (ES) patients were analyzed here. Whole Genome Sequencing was performed on two patients with normal, primary and relapsed tissue. Whole Exome Sequencing was performed on 50 ES and 22 matched normal tissues. A discovery data set of 14 of these tumor/normal pairs identified 232 somatic mutations. Recurrent non-synonymous mutations were validated in the 36 remaining exomes. Transcriptome analysis was performed in a subset of 14/50 ES and DNA copy number gain and expression of FGFR1 in 63/116 ES.

**Results:** Relapsed tumors consistently showed a 2- to 3-fold increased number of mutations. We identified several recurrently mutated genes at low frequency (ANKRD30A, CCDC19, KIAA0319, KIAA1522, LAMB4, SLFN11, STAG2, TP53, UNC80, ZNF98). An oncogenic fibroblast growth factor receptor 1 (FGFR1) mutation (N546K) was detected, and the FGFR1 locus frequently showed copy number gain (31.7%) in primary tumors. Furthermore, high-level FGFR1 expression was noted as a characteristic feature of ES. RNA-Interference of FGFR1 expression in ES lines blocked proliferation and completely suppressed xenograft tumor growth. FGFR1 TKI inhibitor therapy in a patient with ES relapse significantly reduced 18-FDG-PET activity.

**Conclusions:** FGFR1 may constitute a promising target for novel therapeutic approaches in ES.
Translational relevance: Pediatric cancers including Ewing sarcoma are characterized by a low mutation count with difficult to target driver lesions. Our investigation provides evidence that integration of sequencing data with gene amplification, expression and functional analysis are crucial to identify oncogenic drivers in cancers with low rates of recurrent mutations. By this approach aberrant FGFR1 activity emerged as a novel target for Ewing sarcoma. FGFR1 inhibitors are in clinical trials for other cancers and could be tested in Ewing sarcoma with different ways of FGFR1 activation.
Introduction

Ewing sarcoma (ES) is a highly malignant bone and soft tissue neoplasia of still enigmatic histogenesis with a prominent stemness phenotype (1, 2). Histogenesis may be endothelial, neuroectodermal (3-5) or osteochondrogenic (6, 7). Genetically, ES is defined by specific balanced chromosomal EWS/ETS translocations which give rise to oncogenic chimeric proteins (EWS-ETS), the most common being EWS-FLI1 as a consequence of the t(11;22)(q24;q12) translocation (8-11).

EWS-ETS chimeric proteins promote cell transformation through abnormal regulation of specific target genes involved in the control of a variety of cellular processes (12). However, little is known about cooperating events that affect oncogenesis and/or progression of ES. TP53 mutations and CDKN2A deletions are known to confer poor prognosis but are infrequently observed for this sarcoma (approx. 13% of cases) (13). Furthermore, despite a relatively stable genome, copy number alterations (CNA) are recurrently observed in ES (14), but only 1q gains as detected in 31% of tumor samples seem to be associated with poor prognosis (15). In genome-wide association studies (GWAS), few genomic loci were associated with susceptibility (16). Cancer arises due to genetic mutations that alter the function of a wide array of genes (17). Cancer genomes are often aneuploid, contain amplifications and deletions, and typically comprise hundreds to thousands of DNA point mutations. However, not all cancers are necessarily that complex. Large-scale sequencing projects have revealed that mutation rates display 10- to 100-fold differences among cancer types and even among different cancers of the same type (18-20). A low mutation rate appears to be a general feature of pediatric cancers (21), challenging the view that genomic instability is critical for tumor progression. A low mutation frequency so far has been observed in medulloblastoma (22), rhabdoid cancers (23), glioblastoma (24), retinoblastoma (25), and seems to be a feature of ES (20). Hence, the pathomechanisms of ES and suitable molecular therapeutic targets have mostly remained elusive. However, even in translocation-driven tumors, multiple clones with varying genetic aberrations and driving capacity for the disease may occur (26). ES has been reported to contain a low number of mutations in coding regions (20). Given the still dismal prognosis of advanced ES (27, 28), we performed whole genome sequencing of diagnosis and relapse samples to identify potential drivers and relapse associated genes. Further, we performed exome sequencing to expand the findings and to identify novel therapeutic targets.
Materials and Methods

Patient material
Patient material was obtained from clinical studies of the Cooperative Ewing Sarcoma Study Group (CESS) in Europe. All patients provided informed consent.

Nucleic acid extraction
We extracted nucleic acids from peripheral blood leukocytes, tumor material and cell cultures by use of either the DNAeasy or the RNAeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) according to the suppliers recommendations.

Whole genome sequencing (WGS)
WGS was performed at Complete Genomics using 10 µg of high molecular weight DNA. Sequences were mapped to the National Center for Biotechnology Information (NCBI) reference genome build 37. Unless explicitly stated otherwise, all computations were performed using the statistical environment R (http://www.r-project.org).

Exome sequencing
Exomes were enriched in solution with SureSelect XT Human All Exon 50 Mb kits (Agilent Technologies, Santa Clara, CA, USA) and sequenced as 100 bp paired-end runs using TruSeq SBS chemistry v3 (Illumina, San Diego, CA, USA) generating 5-12 Gb of sequence and an average read depth between 66 and 166 on target regions. More than 90 % of the target regions were covered 20 times or more. Burrows-Wheeler Aligner (BWA v 0.5.9) with standard parameters was used for read alignment against the human genome assembly hg19 (GRCh37). We performed single-nucleotide variant and small insertion and deletion (indel) calling for the regions targeted by the exome enrichment kit using SAMtools (v 0.1.18). Large indels were called with Pindel (v 0.2.4t) and copy number variations (CNV) were determined using the R package ExomeDepth (v 0.9.7).

To discover putative somatic variants, we retrieved only those variants of a tumor that were not found in the corresponding control tissue. To reduce the number of false positives, we filtered out variants that were already present in 3,600 “in house” control exomes (patients with unrelated diseases and healthy controls from other projects) or had variant quality of less than 30. Furthermore, the variants were filtered according to several quality criteria using the SAMtools varFilter script. We used default parameters, with the exception of the maximum read depth (-D) and the minimum P-value for base quality bias (-2), which we set to 9999 and 1e-400,
respectively. Moreover, we applied a custom script that marked all variants where the median base quality of adjacent bases was low, because these variants are often sequencing artifacts. We then manually investigated the raw read data of the remaining variants using the Integrative Genomics Viewer (IGV v2.3.19).

**FGFR1 copy number**

*Real time PCR.* FGFR1 gene copy number was determined using TaqMan copy number assays (Life Technologies) according to the supplier’s protocol. In brief each 10 ng of genomic DNA were used for duplexed real time PCRs targeting human FGFR1 and RNase P gene. Reactions were carried out in duplicates. Gene copy number was calculated using Copy Caller v2.0 software.

*Interphase fluorescent in situ hybridization (FISH).* Cells were grown under standard culture conditions. Standard cytogenetic techniques were used for harvesting, isolation of interphase nuclei and slide preparation. FISH was carried out using a dual labeled break apart probe (Kreatech) targeting FGFR1 and surrounding sequences of 1.1 MB length and a centromere 8 probe. Nuclei were counterstained with DAPI.

**FGFR1 expression**

*Immunohistochemistry.* Slides of the 41 Ewing sarcoma samples were deparaffinized and rehydrated. Primary antibody targeting FGFR1 and HRP coupled secondary antibody was used for antigen detection. Staining intensities were graded semi-quantitatively based on the percentage of positive cells (0: negative, 1: <25%, 2: <50%, 3: >50%) using positive controls recommended by the antibody supplier.

*Real-Time RT PCR.* Reverse transcribed mRNA of 61 ES was used for quantification of FGFR1 expression by means of SYBR green real-time PCR. GAPDH was used as housekeeping gene and FGFR1 expression was calculated relative to the expression in the ES cell line VH-64 using the 2^−delta CT method.
In vitro studies

**FGFR1 knock-down and proliferation.** Cells were lentivirally transduced with the vector pLKO-GFP containing 2 different shRNAs against human FGFR1. Scrambled shRNA sequence cloned into pLKO-GFP served as control. Cells were FACS sorted for GFP expression after 48 hours. For adherent cell lines (A673, SK-NMC, A549) 50,000 cells per well were seeded into 24 wells in a 12-well plate format whereas 5,000 cells were used for suspension cell lines (KASUMI, HL60). Cells were cultured under standard conditions in RPMI-Medium containing 10% FCS. Cell numbers were determined every day for 5 consecutive days. Biological replicates were performed.

**FGFR1 mutagenesis and proliferation.** FGFR1 cDNA (ThermoFisher ClonelID: 3911101) was mutated (N546K) using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technology) following the suppliers instructions. Wildtype and mutated cDNA was cloned into MSCV2.2- GFP and transduced into NIH-3T3 cells using standard techniques. Effect of N546K on proliferation was analyzed as described above for adherent cell lines.

**FGFR1 mutagenesis and colony forming.** Soft agar assays were used to determine the effect of N546K on colony forming capacity. Briefly, 10,000 cells per well were transduced with either control, wildtype or mutated FGFR1. After mixing with agarose and plating they were allowed to form colonies for 7 days.

Animal model

Immune deficient Rag2⁻/⁻γc⁻/⁻ mice on a BALB/c background were obtained from the Central Institute for Experimental Animals (Kawasaki, Japan) and maintained in our animal facility in Munich under pathogen-free conditions in accordance with the institutional guidelines and approval by local authorities. Experiments were performed in 5 mice/group at the age of 6-10 weeks.

In vivo experiments

For the analysis of in vivo tumor growth cells were harvested by trypsinization, washed with Dulbecco's phosphate buffered saline, and injected in a volume of 0.2 ml into immune deficient Rag2⁻/⁻γc⁻/⁻ mice. To investigate local tumor growth, 2 x 10⁶ ES cells and derivatives were injected subcutaneously into the groin using a 26-gauge needle attached to a 1 ml syringe. Tumor size was determined as described
(1). Mice bearing a tumor > 10 mm in diameter were considered as positive and sacrificed.

**Additional statistical analysis**

Descriptive statistics is used to determine parameters like mean, standard deviation and standard error of the mean (SEM) from more than two independent experiments. Differences were analyzed by unpaired two-tailed student’s t-test as indicated using Excel (Microsoft) or Prism 5 (GraphPad Software); p values < 0.05 were considered statistically significant (*p < 0.05; **p < 0.005; ***p < 0.0005).
Results

Specimens from 116 Ewing sarcoma (ES) patients were investigated in this project. Matched normal, primary and relapse specimens from two patients were analyzed by whole genome sequencing (Fig.1; green labeled probes). For exome sequencing and expression analysis we utilized tumor samples of 51 patients (Fig.1; ES001-ES051). For 24 of these patients normal tissue controls were available. We performed whole exome sequencing (WES) in a discovery set of 14 of the matched tumor/normal sample pairs that was selected based on average coverage and high tumor content (for 7 primary only, for 6 relapsed only and for one both, (Fig.1, orange labeled probes). As a validation cohort we performed exome sequencing on further 36 ES patients (17 primary with three matched normal tissues, 18 relapses with five matched normal tissues and one with primary and relapse) for the detection of recurrent mutations. Expression data were obtained from tumor material from 14/53 patients, which were analyzed by WES and compared to a normal body map of 21 different healthy tissues (NBA, see Supplementary Materials and Methods). Finally, the remaining 63/116 ES patient samples were utilized for analysis of FGFR1 DNA copy number gain, mRNA profiling and immunohistochemistry. Combined patient characteristics and subjected methods are shown in Fig. 1 and Supplementary Fig. S1.

Whole genome sequencing (WGS) in patients with relapses of Ewing sarcoma (ES). Patient #1 was a 36-year old male with first diagnosis of ES in the left tibia. One year after multimodal therapy the patient relapsed locally. A second local relapse occurred along with lung metastases one year later. Ultimately, the patient died of the disease 8 months later. WGS was performed using specimens obtained from the first and second relapse and peripheral blood as germline control.

The female patient #2 was 12 years old at diagnosis of the primary tumor in the left pubic bone and the left acetabulum and lung metastases. She was treated according to the EURO-Ewing-99-protocol. The first relapse, lung metastases and local recurrence, occurred 2 years after diagnosis. Three years later the second pulmonary relapse was diagnosed. Samples for WGS were taken from peripheral blood, the primary tumor and two consecutive pulmonary relapses.

In both patients, the EWS-FLI1 mutation was present in all diagnostic and relapse samples.

Randomly selected somatic mutations obtained via whole genome sequencing were verified by Sanger sequencing to develop receiver-operating curves (ROC) that ena-
bled accurate quantitation of mutation load per mega base (Supplementary Table S1 and Fig. S2). In the female patient, the total number of somatic mutations was 1325 and doubled with each relapse (1<sup>st</sup> relapse = 3229 mutations; 2<sup>nd</sup> relapse = 6443 mutations), indicating that recurrent disease appeared to consistently double the number of somatic mutations (Fig. 2A). Similar findings were obtained in the male patient with 2506 mutations at first sample and 6219 mutations in the second sample. Interestingly, the 2- to 3-fold increase in mutation rate for each relapse was consistent among different types of mutations. For single nucleotide variations there was a 1.9- to 2.5-fold increase in mutation frequency. Deletions occurred in relapsed samples at a 2.7- to 3.2-fold increase compared to the previous specimen (Supplementary Fig. S2B). These findings suggest, that despite different treatments, the increasing genomic instability of ES is evenly distributed among different mutation types. Nonetheless, deletions occurred up to two times more frequently than insertions. This trend was observed in all direct comparisons of recurrence with previous tumor stage. Between 100 and 500 mutations per specimen affected exonic regions. Of these, only a minority between 27 and 115 were found as tier 1 mutations with expected functional consequences for coding proteins (Fig. 2B). In line, a strong increase was observed from the time of diagnosis to the consecutive relapsed situations (patient #1: 1<sup>st</sup> relapse = 61 mutations; 2<sup>nd</sup> relapse = 109 mutations; patient #2: primary = 27 mutations; 1<sup>st</sup> relapse = 62 mutations; 2<sup>nd</sup> relapse = 115 mutations). The increase in mutation frequency was about 2-fold (Supplementary Fig. S2C). Mutations were evenly distributed throughout the genome in the male patient (Supplementary Fig. S3A left) as well as in the female patient with primary tumor, first and second relapse (Supplementary Fig. S3A right). A further analysis of the mutation types for single nucleotide variations revealed, that Cytosine conversions into Thymidine were most frequent followed by Adenosine (A) to Guanosine (G) transversions (Fig. 2C) which was similarly observed after exome sequencing (Fig. 3D). C to T conversions are frequently observed in damaged long single-strand DNA-regions due to the activity of APOBEC enzymes (29, 30). Of note, in the second relapse of the female patient the A to G mutation was most frequently observed.

In the female patient, two major clones were observed at the time of diagnosis (Fig. 2D). At the time of first relapse and second relapse only one predominant clone was observed with indications for smaller subclones. In the male patient only one major clone was observed and one predominant clone was present also at the time of sec-
ond relapse. These findings indicate that clonal heterogeneity does exist in ES. Also, similar to other cancers it is likely, that clonal selections occur via chemotherapy and are associated with fewer subclones in relapses.

**Whole exome sequencing of ES patient samples.** In addition to whole genome sequencing in primary and relapse patients we sought to further identify the spectrum of mutations in ES. In a discovery data set of 14 individual patients (7 primary, 6 relapsed and one patient where both, primary and relapse were available) we analyzed matched tumor/normal pairs. High tumor cell content, the availability of matched normal tissue and the high quality and amount of DNA enabled exome sequencing analysis with especially high sensitivity and specificity. As a validation cohort we performed exome sequencing of further 36 ES (17 primary, 18 relapses and one with both) for the detection of recurrent mutations; part of them (8 samples) matched to normal controls. The ES samples in the discovery cohort in their majority exhibited EWS/FLI1 translocations (Supplementary Table S2). To reduce the number of false positives, we filtered out variants that were already present in 3,600 “in house” control exomes (see Materials and Methods). The matched tumor/normal pairs were used to identify the rate of somatic mutations per megabase exomes. We identified 232 somatic mutations in 14 ES, of which 166 were damaging mutations leading to loss of protein function. On average 0.38 somatic mutations per megabase were observed in ES. Most of these mutations were non-synonymous mutations (Fig. 3A). Pediatric tumors have been reported to harbor a low number of mutations (21), a finding that was recapitulated in our study. However, ES also occurs in adult patients. Importantly, in our investigation the mutation rate did not increase significantly with age of the patient (Fig. 3B). This indicates that age is not the determining factor for the number of mutations in ES. Also, higher mutation load and increased genomic instability might not be the reason for the less favorable outcome of ES in adult patients. Primary tumors harbored on average 9 mutations per exome whereas around 18 mutations per exome were found in relapsed tumors (Fig. 3C). This again is in line with the finding in the whole genome sequencing that relapses are associated with a 2- to 3-fold increased number of somatic mutations. Examples comprising mutation patterns in particular patients are given in Supplementary Fig. S3B.

For subsequent analysis of mutation frequency, 36 additional ES samples were investigated by whole exome sequencing. This approach identified a set of recurrently mutated genes. The further examination was performed according to presumed vari-
ations in patient-specific mutation frequency and gene-specific background mutations by use of the MutSig algorithm (http://www.broadinstitute.org/cancer/cga/mutsig). The MutSig analysis led to a list of genes with non-synonymous mutation frequencies ranging from 5.7 % to 11.5 % (Table 1, Supplementary Table S4). Mutations in e.g. ankyrin repeat domain-containing protein 30A (ANKRD30A; at 9.6 %), coiled-coil domain containing protein 19 (CCDC19, at 11.5 %), laminin, beta 4 (LAMB4; at 9.6 %), schlafen family member 11 (SLFN11; at 7.7 %), unc-80 homolog (UNC80; at 7.7 %), and zinc finger protein 98 (ZNF98; at 7.7 %) may involve different aspects of tumor pathology presumably important but not essential for disease development.

However, such recurrent mutation as filtered by the MutSig algorithm where of low statistical significance (with no FDR value below 0.1; table 1). Silencing mutations of stromal antigen 2 (STAG2) in ES were recently described (31-34) and were confirmed by Sanger sequencing (Supplementary Fig. S4A). The presumed aneuploidy in cases with STAG2 mutations as described by others (31) was not confirmed in our study similar to other published observations (35).

Comparison with expression data interestingly demonstrated an up-regulation of SLFN11 (fold change (FC) = 10.9), STAG2 (FC = 2.1), and TP53 (FC = 1.79) in ES when compared to the mean of a normal body map of healthy tissue (NBA, see Supplementary Materials and Methods) (Table 1, Supplementary Table S4, S5). Others, like ANKRD30A, CCDC19, LAMB4, or ZNF98 were ubiquitously expressed while some were up in brain or fetal brain tissue such as KIAA0319, KIAA1522, or UNC80, which is also true for STAG2.

Furthermore, analysis of overlapping somatic mutations in three different patient triplets of primary and relapsed tumor samples (Pat.1 and 2, and ES014) revealed recurrent mutations for each patient (Supplementary Table S5). Overall a large amount of mutations were shared in consecutive sarcomas (Pat. 1: 61 mutations in relapse 1, of these 42 shared with relapse 2; Pat. 2: 27 mutations in primary, of these 8 shared with relapse 1 and/or relapse 2; ES014: 13 mutations in primary, of these 9 shared with metastasis).

Overall, the observed low frequency of recurrent non-synonymous mutations of which only STAG2 seemed to be statistically significant suggests that collaborative events in ES do not require single nucleotide variation in addition to EWS/ETS translocations.
Deregulation of FGFR1 in ES. Exome sequencing was performed to identify novel driver genes in Ewing sarcoma. Interestingly, one somatic mutation in the male patient with ES that underwent WGS was a known oncogenic mutation in FGFR1. This mutation (N546K) occurred in the tyrosine kinase domain (Fig. 4A) and was independently discovered as an activating mutation in pilocytic astrocytoma and glioblastoma (36, 37). This mutation was confirmed in all tumor specimens from this patient (Supplementary Fig. S4B). Sanger sequencing of the FGFR1 kinase domain in 30 additional Ewing sarcoma specimens did not reveal additional mutations (data not shown). As a complementary approach for the identification of driver events, we utilized exome data to identify copy number variations. Deletions involving different exons of cyclin-dependent kinase inhibitor 2a (CDKN2A) (15) were observed in 4 of 14 tumors, but gain of chromosome 8, that was found in 5 of the 14 tumors, was among the most frequent alterations (Fig. 3E). This is in line with previous findings (15), (38). Trisomy 8 as well as amplifications of chromosome 8 are the most frequent cytogenetic findings in ES (14). The relevant gene or the relevant genes have not been identified so far. Of note, all amplifications involve the FGFR1 gene. FGFR1 is located on chromosome 8p, and we next analyzed the copy number status of FGFR1 in ES. A real time PCR-based method identified a gain of the FGFR1 locus in 13 of 41 patients (31.7 %). Of these, 12 patients had gained one copy and one patient harbored four gene copies (Fig. 4C). Our analyses included copy number gains due to chr8 amplification as well as direct gene copy number gain. FGFR1 copy numbers closely correlated with FGFR1 gene expression (Fig 4D). Gains of FGFR1 copy number did not only occur in primary tumors, but were consistently observed in ES cell lines A673, CADO, SK-NMC, SK-ES, and VH-64. Copy number gains of FGFR1 occur due to trisomy 8 in the majority of patients but a proportion of patients and cell lines also showed amplification beyond three copies of the entire chromosome. (Fig. 4B)

In line, FGFR1 was highly expressed in ES on the mRNA level (Supplementary Fig. S5A and S5B) and on the protein level as indicated by immunohistochemistry (Fig. 4E and F).

Functional analyses of FGFR1 expression in ES. Activating mutations, copy number gains and other alterations as well as increased expression of FGFR1 occur in ES (e.g. compared to normal tissue, Supplementary Fig. 5B) with a high frequency. We therefore further analyzed the potential role of FGFR1 in ES. First, we retrovirally
transduced FGFR1 wild type or the FGFR1 N546K mutant into NIH3T3 cells. Both, overexpressed FGFR1 wild type and the N546K mutation enhanced proliferation (Fig. 5A). In colony formation assays, FGFR1 N546K mutation more than quadrupled the number of colonies obtained after empty vector transduction, whereas the FGFR1 wild type doubled colony formation (Fig. 5B). The transformation of NIH3T3 cells by FGFR1 wild type and FGFR1 N546K mutation is also indicated in Fig. 5C by the formation of three-dimensional clusters. Alterations of FGFR1, especially the gain of the locus as well as expression of the protein, were associated with a trend towards inferior survival in patients with ES (Supplementary Fig. S6).

To analyze whether FGFR1 signaling was required for ES growth, we performed shRNA-based knock-down after transduction with lentiviral constructs encoding scrambled (pLKO_shcontrol) or FGFR1 specific shRNA (pLKO_shFGFR1), respectively. Specific shRNA knock-down suppressed FGFR1 mRNA expression by 70% or more in ES cell lines and was confirmed on the protein level, too (Supplementary Fig. S5C and S5D). Knock-down of FGFR1 inhibited growth of ES and other cancer cells such as the A549 lung carcinoma cell line or the leukemia cell lines HL60 and Kasumi. Of note, almost complete growth suppression was observed only in the ES cell lines (Fig. 5D).

To analyze whether FGFR1 inhibition was also relevant for ES growth in vivo, lentivirally transduced A673 ES cells were injected into Rag2-/-γc-/- mice (6). Viability at the time of injection was 100% for both scrambled (pLKO_shcontrol) and FGFR1 shRNA (pLKO_shFGFR1) transduced cell lines. However, mainly scrambled transduced A673 cells formed tumors in mice (100%), whereas only in one mouse a tumor was formed by FGFR1 shRNA (pLKO_shFGFR1) transduced cells. Tumor growth and survival curves are shown in Fig. 5F and 5G. These findings suggested that FGFR1 is a crucial signaling pathway for Ewing sarcoma.

Ponatinib is a tyrosine kinase inhibitor that among others, effectively targets FGFR1 at low nanomolar concentrations (39). Accordingly, proliferation of all tested ES cell lines was inhibited by Ponatinib treatment in the nanomolar range of concentrations (Fig. S6C). One patient with ES, who had exhausted all other treatment options, was offered Ponatinib treatment. FGFR1 inhibition by Ponatinib elicited a response by significantly decreasing the glucose uptake of the tumor (Supplementary Fig. S6D, Supporting information). the treatment effect was not maintained and the patient eventually died with clinical progress of CNS metastases (brain and meninges).
DISCUSSION

Pediatric malignancies, and especially ES, have been reported to contain few if any recurrent mutations (21, 33, 34). Our study leads to important novel insights into the pathogenesis of ES. First, the mutation load in ES remains very low on whole genome as well as whole exome sequencing level. No frequent genetic driver mutation besides the well described EWS/ETS translocation was identified (40). Mutation load did not depend on age at diagnosis which suggests that mutation load does not play an important factor for the worse prognosis observed in adult patients. It appears that the number of observed mutations in ES is disease specific and was not closely associated with patient age at the time of diagnosis. This is in contrast to age-associated cancers such as AML, which show increased mutation load in older patients (41). Despite the unifying events of the balanced EWS/FLI1 translocation, clonal subpopulations exist. These contain different mutations, which are only partially overlapping. Signs of clonal selection are present in relapse specimens and each relapse was associated with a 2- to 3-fold increase in somatic mutation load compared to the primary tumor or the previous relapse. So far, no targetable mutations had been reported in ES. In our study, we identified alterations of the FGFR1 gene as a highly frequent event in ES pathogenesis. We initially discovered a somatic point mutation in the tyrosine kinase domain with activating activity in two subsequent relapse specimen indicating stability of the mutation. The structural analysis of this mutation (37) as well as functional analyses provide evidence that this N546K mutation induces enhanced FGFR1 activity and increased proliferation. In addition, wildtype FGFR1 is overexpressed in many ES tumors. Forced expression of FGFR1 increased proliferation and clonogenic growth in fibroblasts. Intriguingly, copy number gain of FGFR1 was detected in a substantial fraction of ES either due to gene amplification or Trisomy 8 cytogenetic aberration. Trisomy 8 is the most frequent cytogenetic aberration that is found in ES but no involvement of a specific gene locus was identified so far (14), (38). Accordingly, FGFR1 was highly expressed on the protein level in almost all analyzed ES. Gene amplification and overexpression are crucial pathogenetic mechanisms for growth factor receptors activation in multiple cancers (42, 43). For example, EGFR amplification and overexpression with subsequent auto-activation has been described in breast cancer and lung cancer. Along the same line, Her2/NEU/ERBB2 amplification and expression is of pathogenetic relevance in breast cancer and associated with a poor prognosis (44). The potential relevance of
FGFR1 for ES was suggested in our data by the association of high FGFR1 expression levels with increased relapse rates and more importantly by the fact, that suppression of FGFR1 abolished the growth of ES cell lines. This phenomenon was evident for all analyzed ES cell lines and was observed in vitro as well as in vivo. The efficacy of shRNA based FGFR1 knock-down was far more prominent in ES than in several other cancer types, such as lung cancer, that had previously been described to be associated with FGFR1 activity (42). Ponatinib as a FGFR1 inhibitor (39) showed significant activity towards proliferation of ES cell lines in our experiments. Of note, proliferation of all ES cell lines was inhibited by Ponatinib (Supp. Fig. 6C).

Ewing sarcoma cells appear to be far more addicted to FGFR1 signaling than carcinoma cells which implies that FGFR1 inhibitors should be tested in ES. In addition to the pre-clinical data for Ponatinib, we also treated one patient with ES, who had exhausted all available therapy options, with Ponatinib monotherapy. Tumor activity was assessed by positron emission tomography scanning. After 3 weeks of therapy, 18F-FDG-glucose uptake was widely reduced. This may be due to FGFR1 inhibition, but we do not rule out at all that other TKs may be involved. Taken together, these data suggest, that FGFR1 might be a suitable therapeutic target in ES. Two main conclusions that can be drawn from our data concern the role of tyrosine kinase alterations in ES and the approaches to identify driver mutations in cancers with a low frequency of recurrent mutations. Mutations, that are not found to be recurrent in a significant fraction of cancers, are usually considered to be bystander or passenger mutations. In many studies, these mutations are neglected, but the direct association between non-recurrent and non-functional, e.g. as passenger mutation, might be misleading. An important role for FGFR1 in ES pathogenesis is evident not only by the oncogenic point mutation but also by gene copy gain and knock-down studies. Thus, DNA-sequencing alone is obviously not sufficient to assign relevant mutations and driver events due to the low frequency of these events in ES. In contrast, the combination of protein expression, genomics and functional analysis provided evidence for the involvement of FGFR1 in ES pathogenesis. Given the high frequency of trisomy 8, it can be envisioned, that about half of all ES patients show copy number gain of FGFR1 on the genetic level. Increased protein expression, that might be even more widespread, can also be influenced and induced by multiple other mechanisms, many of these obviously need to be defined for future studies.
Cancers with balanced translocations provide a unique opportunity to identify cooperating events. In some leukemias, a two hit model was proposed that depends on a balanced translocation predominantly affecting self-renewal and differentiation (45). An increased tyrosine kinase activity is a cooperating event that was thought to be responsible preferentially for increased proliferative and anti-apoptotic activity. It is obvious that this kind of model is overly simplistic for ES and FGFR1 alterations are just one event that cooperates with EWS/ETS translocations.

In summary, our data provide evidence, that integration of genomic data with gene copy number gain, expression and functional analysis are crucial to identify oncogenic drivers in cancers with low rates of recurrent mutations. FGFR1 may constitute a novel target for therapeutic approaches in Ewing sarcoma.
ACKNOWLEDGMENTS

Exome sequencing was performed with help of Tim M Strom and Thomas Schwarz-mayr from the Institute of Human Genetics, Helmholtz-Zentrum München, 85764 Neuherberg, Germany. We are grateful to Julia Humberg for assistance in sequencing and functional studies and to all C.MT lab members for constant help and discussions. We are grateful to the team of the Cooperative Ewing Sarcoma study group (CESS) of the Society for Pediatric Oncology and Hematology for the correlation of our findings with the clinical data.

**Funding:** This work is part of the *Translational Sarcoma Research Network* (TransSaRNet; 01GM0870), “Rare Diseases” and *Prospective Validation of Biomarkers in Ewing Sarcoma for Personalised Translational Medicine* (PROVABES; 01KT1311), Funding Programs of the Federal Ministry of Education and Research (BMBF), Germany. This work was supported by the German Cancer Consortium (DKTK) with help of the German Cancer Research Center (DKFZ), Heidelberg, Germany, and grants from the *Else-Kröner-Fresenius Stiftung* (2013_A49), the *Wilhelm-Sander Stiftung* (2006.109.2), and by the Deutsche Forschungsgemeinschaft, DFG EXC 1003 Cells in Motion – Cluster of Excellence, Münster, Germany and the Deutsche Krebshilfe (DKH-108128).
References


**Author contribution:**
KA, GR, GK, CR, HJ, UD, SB, WEB and CMT worked on conception and design. Experiments were performed as follows: NGS data acquisition was done by KA, GR, ES, KH, IS, AW and MS; FGFR-1 Expression was profiled by GK, WH and EW (Immunohistochemistry), by GR, TB and KH (Microarray analysis) and by KA and ES (RT-PCR); FGFR-1 copy number was analysed by KA (real time PCR) and RB (FISH); in vitro studies were conducted by KA, ES and BM; in vivo experiments were performed by GR and KH; patients were recruited and samples prepared by UD, UK, GK, EW, SB, and DB. Analysis and interpretation of NGS data was done by KA, GR, ES, HUK, MD, KP, EK, TB, CR and CMT. Imaging of patients was done by MW. KA, GR, ES, KP, UD, WEB and CMT drafted the manuscript. All authors revised and finally approved the manuscript.

**Competing Interests:** No potential conflicts of interest were disclosed by the authors
FIGURE LEGENDS

**Figure 1. Combined patient characteristics and subjected methods.** For Whole Genome (WGS, green) and Whole Exome Sequencing (WES) tumor samples derived from 52 Ewing sarcoma patients were used. A discovery data set of 14 matched tumor-normal pairs (orange) allowed identification of somatic mutations and analysis of chr8 amplification. FGFR1 copy number (CNV) was analyzed by real time PCR. Expression of FGFR1 was determined either on RNA level by microarray technology or real time PCR or on protein level by immunohistochemistry (IHC). Additional sarcomas were analyzed for FGFR1 expression or CNV only (Supplementary Fig. S1)

**Figure 2. Whole genome sequencing analysis of Ewing sarcomas. (A)** Frequency of somatic mutations per megabase of genomic DNA in primary and relapsed samples. A two to three-fold increase was shown for each relapse. (B) Classification of somatic mutations using a tier-4 scheme (tier 1 = within coding regions, tier 2 = within conserved regions, tier 3 = non-repetitive regions, and tier 4 = the remaining ones). (C) Percentage of the observed mutations within the six different possible base-pair substitutions. (D) Clonality analysis based on mutant allele frequency in WGS data. Only SNV for sequenced alleles with >50x coverage were used. Allele frequency was not corrected for DNA contribution of non-tumor cells.

**Figure 3. Whole exome sequencing analysis of Ewing sarcoma; analysis of somatic mutations in 14 tumor-normal pairs** (7 primary, 6 relapsed and for one both; male:female 1:1; median age at sampling 18.4 years, range 5-47 y). (A) Frequency of protein-coding (non-syn) and silent somatic mutations (syn) per megabase of genomic DNA. (B) Number of observed somatic mutations per age. Each dot represents an individual patient. (C) Quantity of somatic mutations in newly diagnosed patients (primary) and patient samples with relapsed disease. (D) Percentage of the observed mutations within the six different possible base-pair substitutions. (E) Observed copy number variations across autosomes in individual tumor-normal pairs of the exome discovery data set.

**Figure 4. FGFR1 in Ewing sarcomas. (A)** Structure, functional domains and known and identified (p.N546K) recurrent somatic mutations of FGFR1 (Fig. modified ac-
cording to NCBI. (B) FGFR1 targeting interphase FISH in cell lines of the Ewing sarcoma family using a FGFR1 locus specific break apart probe spanning 1.1 Mbp (red/green) and a centromere 8 probe (red). (C) FGFR1 copy number in Ewing sarcoma determined by TaqMan based copy number assays. Duplexed real time PCRs targeting human FGFR1 and RNase P were performed to calculate copy number using Copy Caller 2 software (Life Technologies) (D) Correlation of FGFR1 copy number and gene expression. Copy number (CN) was determined as explained above and RNA expression by means of SYBR green real-time PCR. GAPDH was used as housekeeping gene and FGFR1 expression was calculated relative to the expression in the ES cell line VH-64 using the 2^(-delta CT) method. Expression correlated significantly with copy number as calculated with the Wilcoxon signed-rank test (CN2, n=16; CN3, n=10; CN4, n=1). (E) FGFR1 expression (0=neg, 1=weak, 2=mod., 3=strong) in Ewing sarcoma by means of immunohistochemistry (IHC). (F) IHC of FGFR1 of representative Ewing sarcoma samples (IHC intensities: ES068=0, ES074=1, ES026=2, ES080=3).

Figure 5. Functional analyses of FGFR1. (A) Proliferation of NIH-3T3 cells expressing a control (empty vector), wildtype (WT) or mutant (MUT) FGFR1. (B) Colony forming capacity of NIH-3T3 cells according to wildtype and mutant FGFR1 expression. (C) Representative colonies of NIH-3T3 cells expressing a control (EV), wildtype (WT) or mutant (MUT) FGFR1. Only FGFR1 expressing cells (WT and MUT) showed three-dimensional clusters indicating cellular transformation. (D) Diagram of the generation of A673 FGFR1 knock-down cells by lentiviral transduction and their use for in vitro and in vivo analyses. (E) Proliferation of cell lines after lentiviral infection with scrambled (pLKO_shcontrol) or FGFR1 specific shRNA constructs (pLKO_shFGFR1) (Ewing sarcoma family: A673 and SK-NMC; leukemia: HL60 and KASUMI, lung carcinoma: A549). (F) Mice were injected with A673 cells infected with scrambled shRNA (pLKO_shcontrol) or FGFR1 shRNA (pLKO_shFGFR1). Tumor growth until day 50 is shown. (G) Survival curve of Rag2^-/-γc^-/- mice after s.c. injection of A673 FGFR1 shRNA (pLKO_shFGFR1) and scrambled shRNA transduced (pLKO_shcontrol) cells, respectively as shown in 5F. Only one mouse of the group of mice injected with A673 FGFR1 shRNA infectants (pLKO_shFGFR1) developed a tumor at day 67 after inoculation.
Table 1: Genes with significant frequencies of recurrent mutations across 50 Ewing sarcoma.

<table>
<thead>
<tr>
<th>name</th>
<th>gene description</th>
<th>recurrence</th>
<th>frequency (%)</th>
<th>p-value*</th>
<th>q-value*</th>
<th>expression (FC)**</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANKRD30A</td>
<td>ankyrin repeat domain 30A</td>
<td>5</td>
<td>9.6</td>
<td>4.11E-05</td>
<td>1</td>
<td>1.02</td>
<td>ubiquitous</td>
</tr>
<tr>
<td></td>
<td>coiled-coil domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>containing 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCDC19</td>
<td>coiled-coil domain</td>
<td>6</td>
<td>11.5</td>
<td>4.39E-04</td>
<td>1</td>
<td>0.92</td>
<td>ubiquitous</td>
</tr>
<tr>
<td></td>
<td>containing 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIAA0319</td>
<td>---</td>
<td>4</td>
<td>7.7</td>
<td>1.20E-02</td>
<td>1</td>
<td>0.57</td>
<td>increased in fetal brain</td>
</tr>
<tr>
<td>KIAA1522</td>
<td>---</td>
<td>4</td>
<td>7.7</td>
<td>1.60E-02</td>
<td>1</td>
<td>0.69</td>
<td>increased in brain, fetal brain</td>
</tr>
<tr>
<td>LAMB4</td>
<td>laminin, beta 4</td>
<td>5</td>
<td>9.6</td>
<td>3.57E-02</td>
<td>1</td>
<td>1.03</td>
<td>ubiquitous</td>
</tr>
<tr>
<td>SLFN11</td>
<td>schlafen family member 11</td>
<td>4</td>
<td>7.7</td>
<td>1.60E-02</td>
<td>1</td>
<td>10.91</td>
<td>increase in ES, also in cerebellum</td>
</tr>
<tr>
<td>STAG2</td>
<td>stromal antigen 2</td>
<td>6</td>
<td>11.5</td>
<td>4.74E-05</td>
<td>4.47E-01</td>
<td>2.10</td>
<td>increase in ES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP53</td>
<td>tumor protein p53</td>
<td>3</td>
<td>5.7</td>
<td>4.46E-04</td>
<td>1</td>
<td>1.79</td>
<td>increase in ES</td>
</tr>
<tr>
<td>UNC80</td>
<td>unc-80 homolog</td>
<td>4</td>
<td>7.7</td>
<td>4.32E-02</td>
<td>1</td>
<td>0.36</td>
<td>increased in brain</td>
</tr>
<tr>
<td>ZNF98</td>
<td>zinc finger protein 98</td>
<td>4</td>
<td>7.7</td>
<td>1.08E-02</td>
<td>1</td>
<td>1.34</td>
<td>ubiquitous</td>
</tr>
</tbody>
</table>

*based on MutSigCV algorithm ([http://www.broadinstitute.org/cancer/cga/mutsig](http://www.broadinstitute.org/cancer/cga/mutsig))

**FC: fold change of expression of ES samples compared to a normal body map (GSE45544)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
</tr>
<tr>
<td>tissue</td>
<td>primary</td>
<td>metastasis/relapse</td>
<td>normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA / RNA profiling</td>
<td>WGS</td>
<td>WES</td>
<td>RNA array</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGFR1 gene dosage</td>
<td>Chr.8 amplification</td>
<td>FGFR1 CNV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGFR1 expr.</td>
<td>RNA</td>
<td>IHC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2

A

![Bar chart showing mutations per megabase for patients 1 and 2 across primary, 1st relapse, and 2nd relapse stages.]

B

![Bar chart showing mutations per megabase for primary, 1st relapse, 2nd relapse, 1st relapse tier 1, 1st relapse tier 2, 1st relapse tier 3, 1st relapse tier 4, 2nd relapse tier 1, 2nd relapse tier 2, 2nd relapse tier 3, and 2nd relapse tier 4.]

C


D

![Density plots showing variant frequency for primary, 1st relapse, and 2nd relapse stages.]

Tumor Variant Allele Frequency

Downloaded from clincancerres.aacrjournals.org on April 18, 2021. © 2015 American Association for Cancer Research.
Deep sequencing in conjunction with expression and functional analyses reveals activation of FGFR1 in Ewing sarcoma.

Konstantin Agelopoulos, Gunther HS Richter, Eva Schmidt, et al.

*Clin Cancer Res* Published OnlineFirst July 15, 2015.