Recurrent intragenic rearrangements of *EGFR* and *BRAF* in soft tissue tumors of infants

Jenny Wegert¹, Christian Vokuhl², Grace Collord³,⁴, Martin Del Castillo Velasco-Herrera³,⁴, Sarah J. Farndon³,⁵, Charlotte Guzzo³, Mette Jorgensen⁶, John Anderson⁵,⁶, Olga Slater⁶, Catriona Duncan⁶, Sabrina Bausenwein¹, Heike Streitenberger¹, Barbara Ziegler¹, Rhoikos Furtwängler⁷, Norbert Graf⁷, Michael R. Stratton³, Peter J. Campbell³, David TW Jones⁸,⁹, Christian Koelsche,¹⁰,¹¹,¹², Stefan M. Pfister⁸,⁹,¹³, William Mifsud⁶, Neil Sebire⁵,⁶, Monika Sparber-Sauer¹⁴, Ewa Koscielniak¹⁴,¹⁵, Andreas Rosenwald¹⁶,¹⁷, Manfred Gessler¹,¹⁷ & Sam Behjati³,⁴

Soft tissue tumors of infancy encompass an overlapping spectrum of diseases that pose unique diagnostic and clinical challenges. We studied genomes and transcriptomes of cryptogenic congenital mesoblastic nephroma (CMN), and extended our findings to five anatomically or histologically related soft tissue tumors: infantile fibrosarcoma (IFS), nephroblastomatosis, Wilms tumor, malignant rhabdoid tumor, and clear cell sarcoma of the kidney. A key finding is recurrent mutation of *EGFR* in CMN by internal tandem duplication of the kinase domain, thus delineating CMN from other childhood renal tumors. Furthermore, we identify *BRAF* intragenic rearrangements in CMN and IFS. Collectively these findings reveal novel diagnostic markers and therapeutic strategies and highlight a prominent role of isolated intragenic rearrangements as drivers of infant tumors.
any childhood tumors show a predilection for specific developmental stages. Tumors that predominantly occur in infancy include congenital mesoblastic nephroma (CMN), which accounts for 4% of all childhood renal malignancies and the majority of those diagnosed in children under 6 months of age. CMN is classified histologically into classical, cellular, and mixed subtypes based primarily on degree of cellularity and mitotic activity. The cellular variant is characterized by a sarcoma-like diffuse hypercellular morphology, whereas classical CMN is composed of less proliferative spindle cells. Cellular CMN is driven by rearrangements involving the tropomyosin receptor kinase (TRK) gene NTRK3, most commonly a t(12;15)(p13;q25) reciprocal translocation with the ETV6 transcription factor. Less frequent somatic aberrations include trisomies of chromosomes 8, 11, 17, and 20 and rarer TRK fusions, involving NTRK1, NTRK2, or NTRK3. By contrast, the genetic changes underpinning the classical variant, accounting for >30% of cases, are unknown. Cellular CMN shares its genetic and morphological hallmarks with infantile fibrosarcoma (IFS), a spindle cell tumor typically arising in the soft tissues of the extremities or abdomen. Standard treatment for CMN and IFS is complete surgical resection. In the case of IFS, local control frequently requires cytotoxic chemotherapy. The role for up-front chemotherapy in CMN is less clear. Recently, a phase I/II clinical trial of a selective TRK inhibitor, larotrectinib, reported high response rates in diverse tumor types harboring TRK gene fusions, including IFS and other soft tissue tumors of infancy. Morbidity and infrequent death result from tumor recurrence or from treatment-related complications.

Here, we investigated the genetic basis of CMN and IFS lacking the canonical NTRK3-ETV6 fusion gene. We identify oncogenic rearrangements in MAPK signaling genes across all cases interrogated by unbiased sequencing, notably therapeutically tractable intragenic rearrangements in EGF and BRAF.

**Results**

**Overview of the genomic landscape of CMN.** To identify the genetic basis of cryptogenic CMN, we first applied whole genome and transcriptome sequencing to a discovery cohort of ten classical CMN lacking an NTRK3 fusion (Supplementary Data 1). Somatic variants were identified by comparing tumor and matched peripheral blood sequences (see Methods). The genomic landscape was universally quiet, with a low burden of point mutations (median of 45 substitutions and 9 insertions or deletions per genome; Supplementary Data 2). The predominant mutational signatures, as defined by the trinucleotide context of substitutions, were the ubiquitous signatures 1 and 13.

---

**Fig. 1** EGF internal tandem duplication. a The genomic footprint of EGF is depicted with exons represented by gray and green vertical lines. Green exons encode the kinase domain. Blue lines superimor show the tandem duplications found in the discovery cohort of ten congenital mesoblastic nephroma of classical histology. b Schematic of the wild-type transcript. c Schematic of the fusion transcript annotated with cDNA sequence of rearrangements (sense orientation) and protein translation. d Intragenic copy number of EGF showing focal amplification over the kinase domain (x-axis: genomic coordinate; y-axis: copy number derived from coverage). e Representative phospo-ERK immunohistochemistry.
immunohistochemistry. Mixed cellularity tumors surveyed also harbored an IFS but not in CMN15. Of note, for one child with BRAF studied, with no additional driver events apparent at relapse. Of note, for one child with

**Internal tandem duplication of the EGFR kinase domain in CMN.** Annotating all cases for potential oncogenic variants revealed a single intragenic, in-frame internal tandem duplication (ITD) of the EGFR kinase domain in all ten tumors (Table 1; Fig. 1; Supplementary Data 3). The breakpoints clustered in a narrow genomic window around the kinase domain of EGFR encoded in exons 18–25 (Fig. 1a). This rearrangement is rarely observed in several other tumor types including in glioma and in lung adenocarcinoma, and confers sensitivity to a targeted EGFR inhibitor, afatinib14. We validated all rearrangements by genomic copy number analysis and reconstruction of cDNA reads spanning the breakpoint junction (Fig. 1; see Methods). Of note, the same mutant cDNA junction sequence was found in every case, irrespective of the genomic location of breakpoints. A search for additional known or novel driver variants revealed no further plausible candidates in any of the EGFR-mutant tumors. We next extended this investigation to seven non-classical CMN lacking an NTRK3 fusion, including four mixed cellularity cases and three IFS and rarely in adult cancers, but not, to our knowledge, in CMN (n = 63), IFS (n = 26), Wilms tumor (n = 208), clear cell sarcoma of the kidney without BCOR rearrangements (n = 20), malignant rhabdoid tumor (n = 3), and nephroblastomatosis.

**BRAF rearrangements in CMN and IFS.** A further striking finding was the discovery of mutations in the BRAF oncogene in 2/3 cellular histology CMNs. BRAF fusions have been implicated in a minority of IFS but not in CMN15. In both cases the BRAF rearrangement involved a compound deletion of conserved region 1 (CR1) and tandem duplication of exon 2 (Fig. 2; Table 1; Supplementary Data 3). CR1 encompasses the negative regulatory Ras-binding domain (RBD), loss of which is predicted to generate a constitutively active form of BRAF16,17. Mutated tumors displayed intense staining of phosphorylated ERK by immunohistochemistry, consistent with activated signaling downstream of BRAF (Figs. 1e and 2e). A further tumor harbored the KIAA1549-BRAF fusion, a molecular hallmark of a childhood brain tumor, pilocytic astrocytoma18,19. This fusion likewise results in loss of the N-terminal portion of the BRAF protein containing the RBD17,18.

**Other TRK fusions in CMN.** The remaining two cases of CMN interrogated by whole genome and transcriptome sequencing were accounted for by gene fusions involving NTRK1, an alternate kinase of the TRK family of protein kinases: TPR-NTRK1 and LMNA-NTRK1. Both of these fusions have been observed in IFS and rarely in adult cancers, but not, to our knowledge, in CMN20–23 (Table 1). Hence, every cryptogenic CMN interrogated by whole-genome sequencing contained an oncogenic rearrangement in BRAF, EGFR, or NTRK1, all of which encode kinases involved in MAPK signaling and are amenable to inhibition with existing drugs12,14,17,24. EGFR-ITD distinguishes CMN from other childhood renal tumors. To validate and extend our findings, we screened IFS and a range of childhood renal tumors for EGFR-ITD, BRAF-ITD, and ETV6-NTRK3 using PCR. Tumor types included additional cases of CMN (n = 63), IFS (n = 26), Wilms tumor (n = 208), clear cell sarcoma of the kidney without BCOR rearrangements (n = 20), malignant rhabdoid tumor (n = 3), and nephroblastomatosis.
is pertinent to clinical trials of TRK inhibitors in CMN and IFS.

The same clone or in independent competing clones. This raises the possibility of intrinsic resistance of some tumors to TRK inhibition, regardless of whether these mutations occur in the same clone or in independent competing clones.

In this vein a structurally similar Braf fusion transcript, albeit implicated in the same signaling mechanism of resistance to certain Braf/MEK inhibitors.

**Discussion**

In this exploration of infant tumors we identify ITD of the EGF kinase domain that delineates a genetic subgroup of CMN transcending histological subtypes. Additionally, we report a novel rearrangement of Braf present in both cellular CMN and IFS. These mutations represent diagnostic markers that can be readily integrated into routine clinical practice. Furthermore, EGF and Braf emerge as therapeutic targets, which may be exploited in certain clinical situations, e.g., large surgically intractable tumors, disease recurrence or metastases.

It is noteworthy that an oncogenic mutation was identified in every tumor that we studied by whole-genome sequencing. Of these, 78% harbored either EGF-ITD or Braf-ITD, while the remaining 22% presented with non-canonical mutations involving Braf, Ntrk1, or Ntrk3. This suggests that less recurrent rearrangement variants, albeit implicated in the same signaling circuitry, may elude detection by targeted diagnostic assays. Moreover, our results indicate that a subset of tumors harbor multiple drivers with important implications for targeted therapy efforts. The finding of co-mutation of NTRK3 and Braf in IFS raises the possibility of intrinsic resistance of some tumors to TRK inhibition, regardless of whether these mutations occur in the same clone or in independent competing clones. This finding is pertinent to clinical trials of TRK inhibitors in CMN and IFS.

In this vein a structurally similar Braf fusion transcript, albeit without duplication of exon 2, has recently been implicated as a mechanism of resistance to certain Braf/MEK inhibitors. These considerations underscore the need for adequate genomic profiling in order to match patients to the most appropriate basket studies and to enable meaningful interpretation of treatment responses. Therefore, we would advocate extending the diagnostic work-up of refractory or relapsed CMN and IFS to whole genome sequencing, particularly in the context of clinical trials.

Biologically our findings draw further parallels between CMN and IFS. We identify Braf and NTRK1 as additional cancer genes operative in both malignancies, substantiating the view that these diagnoses represent variants on the same disease spectrum converging on aberrant RAS-RAF-MEK-ERK signaling.

Furthermore, in the wider context of the childhood cancer genome, our findings add to the growing body of studies that identify short distance intragenic rearrangements as a dominant source of oncogenic mutations in otherwise quiet genomes. We note the parallel between CMN, clear cell sarcoma of the kidney and low-grade glioma that are in large part driven by ITDs often involving kinase domains, mostly as isolated driver events. Furthermore, even in acute myeloid leukemia, where Flt3-ITD is a recurrent driver event in adult disease, childhood AML demonstrates a distinct structural variant profile enriched for focal chromosomal gains and losses.

We can only speculate on the biological significance of this parallel which may allude to specific mutational mechanisms operative during discrete stages of human development.

**Methods**

**Patient samples.** All tissue samples were obtained after gaining written informed consent for tumor banking and future research from the patient (or their guardian) in accordance with the Declaration of Helsinki and appropriate national and local ethical review processes. German tissue samples were obtained from the following studies: SIOP93-01/GPOH and SIOP2001/GPOH (Ethikkommission der Ärztekammer des Saarlandes ethics reference number 105/95 and 51/2003) and the PT2.0 study (Medical Faculty Heidelberg ethics reference number S-546/2016), the CWS trials CWS-96 and CWS-2002P (Universitätsklinikum Tübingen Medizinische Fakultät ethics approval, reference numbers 105/95 and 51/2003) and the SoTisa registry (ethics approval reference 158/2009B02). UK patients were enrolled under ethics approval from National Research Ethics Service Committee East of England, Cambridge Central (reference 16/EE/0394). Use of UK archival material was approved by the National Research Ethics Service Committee London Brent (reference 16/LO/0960). Additional tissue was obtained from the UK Children’s Cancer and Leukaemia Group tissue bank.

**Sequencing.** Tumor DNA and RNA were extracted from fresh frozen tissue that had been reviewed by reference pathologists. Normal tissue DNA was derived from blood samples. Whole genome sequencing was performed by 150-bp paired-end sequencing on the Illumina HiSeq X platform. We followed the Illumina no-PCR library protocol to construct short insert libraries, prepare flowcells, and generate clusters. Coverage was at least 30×. Messenger RNA was enriched by polyA-
non-negative matrix factorization as implemented in the SomaticSignatures software. In the tumor samples and mutant allele reported in >1% of the matched normal reads, the median BWA alignment score of the reads reporting the mutant allele was ≥95%; no insertion or deletion called within a read length (150 bp) of the mutant allele, at least tenfold coverage at the mutant allele locus, minimum variant allele frequency 5%; no insert or deletion called within a read length (150 bp) of the putative substitution, no soft-clipped reads reporting the mutant allele, and a fraction 5%; no insertion or deletion called within a read length (150 bp) of the putative substitution, no soft-clipped reads reporting the mutant allele, and a median BWA alignment score of the reads reporting the mutant allele >110. The following variants were flagged for additional inspection for potential artifacts, germline contamination or index-jumping event: any mutant allele reported within 150 bp of another variant, any mutant allele with a population allele frequency >1 in 1000 according to any of five large polymorphism databases (ExAC, 1000 Genomes Project, ESP6500, CG46, Kavari), variant reported in more than 10% of the tumor samples and mutant allele reported in >1% of the matched normal reads. For indels, the inbuilt filters of the Pindel algorithm, as implemented in our pipeline, were used. In addition, recurrent indels occurring in >2 samples were flagged for additional inspection.

Mutational signatures were derived using principal component analysis and non-negative matrix factorization as implemented in the SomaticSignatures R package39.

**Variant validation.** The Cancer Genome Project (Wellcome Trust Sanger Institute) variant calling pipeline has been continually validated and bench-marked40,41. We confirmed variant calling quality through manual visualization of raw sequencing read for 8% of all variants called. All rearrangements reported were validated by recombination at base pair resolution and by cDNA reads spanning the breakpoint junction.

**Analysis of mutations in cancer genes.** We considered variants as potential drivers if they were presented in established cancer genes42. Tumor suppressor coding variants were considered if they were annotated as functionally deleterious by an in-house version of VACrENT (http://cancerit.github.io/VACrENT/)43 or were disruptive rearrangement breakpoints or local (<1 Mb) homozygous deletions. Mutations in oncogenes were considered driver events if they were located at previously reported canonical hot spots (point mutations) or amplified the intact gene. Amplifications also had to be focal (<1 Mb) and increase the copy number of oncogenes to a minimum of five copies for a diploid genome. To search for driver variants in novel cancer genes or in non-coding regions, we employed previously developed statistical methods that identify significant enrichment of mutations, taking into account various confounders such as overall mutation burden and local variation in the mutability of the genomic region44.

**Targeted mutation screening.** RNA from frozen tumors (1 µg) or corresponding DNA from formalin-fixed, paraffin-embedded (FFPE) sections was reverse transcribed using oligo-dT or random hexamer primers (RevertAid first strand cDNA synthesis kit, ThermoFisher). PCR screening was performed using primer combinations that allow amplification of candidate alterations as well as additional control fragments from the unaffected allele to assess cDNA quality. Amplified fragments were sequenced by Sanger sequencing (GATC, Konstanz, Germany) using primers detailed in Supplementary Table 1.

**Immunohistochemistry.** Immunohistochemical staining for phospho-ERK1/2 (Cell Signaling Technology, clone D13.14.E1) was performed according to standard protocol (dilution 1:800, pre-treatment with target retrieval TR6.1, Dako). Results were scored in a semi-quantitative fashion (negative, weak, moderate, strong).

**Code availability.** The algorithms used to analyze sequencing data are available at http://cancerit.github.io/.

**Data availability.** All data supporting the findings of this study are available within the article and its supplementary files or from the corresponding author on reasonable request. Sequencing data have been deposited at the European Genome-Phenome Archive (accession numbers EGAS00001002534 and EGAS00001002171).

Received: 15 December 2017 Accepted: 14 May 2018

Published online: 18 June 2018

**References.**

1. Marsden, H. B. & Lawler, W. Primary renal tumours in the first year of life. A population based review. Virchows Arch. A. Pathol. Anat. Histopathol. 399, 1–9 (1983).
2. Glick, R. D. et al. Renal tumors in infants less than 6 months of age. J. Pediatr. Surg. 39, 522–525 (2004).
3. Charles, A. K., Vujanic, G. M. & Berry, P. J. Renal tumours of childhood. Histopathology 32, 293–309 (1998).
4. Eusebi, V. et al. Congenital mesoblastic nephroma (t12;15) is associated with ETV6-NTRK3 gene fusion: cytogenetic and molecular relationship to congenital (infantile) fibrosarcoma. Am. J. Pathol. 153, 1451–1458 (1998).
5. Knezevich, S. R. et al. ETV6-NTRK3 gene fusions and trisomy 11 establish a histogenetic link between mesoblastic nephroma and congenital fibrosarcoma. Cancer Res. 58, 5046–5048 (1998).
6. Adam, L. R., Davison, E. V., Malcolm, A. J., Pearson, A. D. & Craft, A. W. Cytogenetic analysis of a congenital fibrosarcoma. Cancer Genet. Cyto. 52, 37–41 (1991).
7. Schofield, D. E., Yunis, E. J. & Fletcher, J. A. Chromosome aberrations in mesoblastic nephroma. Am. J. Pathol. 143, 714–724 (1993).
8. Church, A. J. et al. Recurrent EML4-NTRK3 fusions in infantile fibrosarcoma and congenital mesoblastic nephroma suggest a revised testing strategy. Mod. Pathol. 31, 463–473 (2018).
9. Gooskens, S. L. et al. Congenital mesoblastic nephroma 50 years after its recognition: a narrative review. Pediatr. Blood Cancer 64, e26437 (2017).
10. Orbach, D. et al. Infantile fibrosarcoma: management based on the European experience. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 28, 318–323 (2010).
11. Soule, E. H. & Pritchard, D. J. Fibrosarcoma in infants and children: a review of 110 cases. Cancer 40, 1711–1717 (1977).
12. Drilon, A. et al. Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children. N. Engl. J. Med. 378, 731–739 (2018).
13. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. Nature 500, 415–421 (2013).
14. Gallant, J. N. et al. EGFR kinase domain duplication (EGFR-KDD) is a novel oncogenic driver in lung cancer that is clinically responsive to afatinib. Cancer Discov. 5, 1155–1163 (2015).
15. Kao, Y. C. et al. Recurrent BRAF gene fusions in a subset of pediatric spindle cell sarcomas: expanding the genetic spectrum of tumors with overlapping features with infantile fibrosarcoma. Am. J. Surg. Pathol. 42, 28–38 (2018).
16. Johnson, D. B. et al. BRAF internal deletions and resistance to BRAF/MEK inhibitor therapy. Pigment Cell Melanoma Res. 31, 432–436 (2018).
17. Karoulia, Z., Gavathiotis, E. & Poulikakos, P. I. New perspectives for targeting RAF kinase in human cancer. Nat. Rev. Cancer 17, 676–691 (2017).
18. Jones, D. T. et al. Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. Cancer Res. 68, 8673–8677 (2008).
19. Ross, J. S. et al. The distribution of BRAF gene fusions in solid tumors and response to targeted therapy. Int. J. Cancer 138, 881–890 (2016).
20. Wong V. et al. Evaluation of a congenital infantile fibrosarcoma by comprehensive genomic profiling reveals an LMNA-NTRK1 gene fusion responsive to crizotinib. J. Natl Cancer Inst. 108, djv307 (2016).
21. Davis, I. L. et al. Infantile NTRK-associated Mesenchymal Tumors. Pediatr. Dev. Pathol. 21, 68–78 (2018).
22. Sartore-Bianchi, A. et al. Sensitivity to entrectinib associated with a novel LMNA-NTRK1 gene fusion in metastatic colorectal cancer. J. Natl Cancer Inst. 108, djv306 (2016).
23. Dobele, R. C. et al. An oncogenic NTRK fusion in a patient with soft-tissue sarcoma with response to the tropomyosin-related kinase inhibitor LOXO-101. Cancer Discov. 5, 1049–1057 (2015).
24. Cook, P. J. et al. Somatic chromosomal engineering identifies BCAN-NTRK1 as a potent glioma driver and therapeutic target. Nat. Commun. 8, 15987 (2017).
25. Roy, A. et al. Recurrent internal tandem duplications of BCO1 in clear cell sarcoma of the kidney. Nat. Commun. 6, 8891 (2015).
26. Zhang, J. et al. Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nat. Genet. 45, 602–612 (2013).
27. Jones, D. T. et al. Oncogenic RAF1 rearrangement and a novel BRAF mutation as alternatives to KIAA1549-BRAF fusion in activating the MAPK pathway in pilocytic astrocytoma. Oncogene 28, 2119–2123 (2009).
28. Jones, D. T. et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat. Genet. 45, 927–932 (2013).
29. Paugh, B. S. et al. Genome-wide analyses identify recurrent amplifications of receptor tyrosine kinases and cell-cycle regulatory genes in diffuse intrinsic...
pontine glioma. J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol. 29, 3999–4006 (2011).
30. Bolouri, H. et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nat. Med. 24, 103–112 (2018).
31. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics (Oxford, England) 26, 589–595 (2010).
32. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England) 29, 15–21 (2013).
33. Jones, D. et al. cgpCaVEManWrapper: simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. Curr. Protoc. Bioinform. 56, 15.10.1–15.10.18 (2016).
34. Raine, K. M. et al. cgpPindel: identifying somatically acquired insertion and deletion events from paired end sequencing. Curr. Protoc. Bioinform. 52, 15.17.11–15.17.12 (2015).
35. Raine, K. M. et al. ascatNgs: identifying somatically acquired copy-number alterations from whole-genome sequencing data. Curr. Protoc. Bioinform. 56, 15.19.11–15.19.17 (2016).
36. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England) 31, 166–169 (2015).
37. Kim, D. & Salzberg, S. L. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. 12, R72 (2011).
38. McPherson, A. et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS Comput. Biol. 7, e1001138 (2011).
39. Gehring, J. S., Fischer, B., Lawrence, M. & Huber, W. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. Bioinformatics (Oxford, England) 31, 3673–3675 (2015).
40. Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 534, 47–54 (2016).
41. Behjati, S. et al. Recurrent mutation of IGF signalling genes and distinct patterns of genomic rearrangement in osteosarcoma. Nat. Commun. 8, 15936 (2017).
42. Forbes, S. A. et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res. 45, D777–D783 (2017).
43. Menzies, A. et al. VAGrENT: Variation Annotation Generator. Curr. Protoc. Bioinform. 52, 15.8.1–15.8.11 (2015).
44. Martincorena, I. et al. Universal patterns of selection in cancer and somatic tissues. Cell 171, 1029–1041.e1021 (2017).

Acknowledgements
This work was supported by funding by the Wellcome Trust, St. Baldrick’s Foundation, the Deutsche Forschungsgemeinschaft (GE 539/13-1), the Deutsche Krebshilfe (50-2709-G2, T9/96/Trl, 50-2721-Tr2) and NIHR GOSH BRC. G.C., S.B., C.G., P.J.C., and M.R.S. received personal fellowships from the Wellcome Trust. The Cooperative Weichteilsarkom Studiengruppe (CWS) was additionally supported by the Deutsche Kinderkrebsstiftung (SoTisR, A2007/13DKS2009.08) and by the Förderkreis Krebskranke Kinder e.V. Stuttgart, Germany. The SIOP-RTSG/GPOH-nephroblastoma study group is supported by the charity “Elterninitiative krebskranker Kinder im Saarland e.V.”. We thank children and their families for participating in our research and the clinical teams involved in their care. We thank Sabine Roth and Sharna Lunn for expert technical assistance.

Author contributions
J.W., G.C., M.D.C.V.H., and C.G. analyzed sequencing data. C.V. performed histological analyses. S.Ba., H.S., and B.Z. provided technical assistance. S.J.F., M.J., I.A., O.S., C.D., R.F., N.G., D.T.W.J., C.K., S.M.P., W.M., E.K., N.S., A.R. and M.S.-S. curated and reviewed the samples, clinical data, and/or provided clinical expertise. M.R.S. and P.J.C. contributed to discussions. M.G. and S.B. directed this research and wrote the manuscript, with contributions from G.C., J.W., and M.D.C.V.H.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04650-6.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.