Novel Fusion Genes and Chimeric Transcripts in Ependymal Tumors

Thale Kristin Olsen ¹-³, Ioannis Panagopoulos ¹,², Ludmila Gorunova ¹,², Francesca Micci ¹,², Kristin Andersen ¹, Hege Kilen Andersen ¹, Torstein R. Meling ⁴, Bernt Due-Tønnessen ⁴, David Scheie ⁵,⁶, Sverre Heim ¹-³, and Petter Brandal ¹,²,⁷

¹ Section for Cancer Cytogenetics, Institute for Cancer Genetics and Informatics, Oslo University Hospital, Norway

² Centre for Cancer Biomedicine, Faculty of Medicine, University of Oslo, Norway

³ Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway

⁴ Department of Neurosurgery, Oslo University Hospital, Norway

⁵ Department of Pathology, Rigshospitalet, Copenhagen, Denmark

⁶ Department of Pathology, Oslo University Hospital, Norway

⁷ Department of Oncology, Oslo University Hospital – The Norwegian Radium Hospital, Norway

* Correspondence to: Thale Kristin Olsen, Section for Cancer Cytogenetics, Institute for Cancer Genetics and Informatics, Oslo University Hospital, P.o box 4953 Nydalen, 0424 Oslo, Norway. Email: thaleko@gmail.com

Abbreviated title: Novel chimeric transcripts in ependymal tumors

Supported by: The Norwegian Cancer Society, the Norwegian Children’s Cancer Society, the Radium Hospital Foundation, and the Research Council of Norway through its Centers of Excellence funding scheme, project number 179571.

RNA sequencing was performed by the Norwegian Sequencing Centre, a national technology platform hosted by the University of Oslo/Oslo University Hospital and funded by the Functional Genomics and Infrastructure programs of the Research Council of Norway and the South-Eastern Regional Health authorities.
We have previously identified two ALK rearrangements in a subset of ependymal tumors using a combination of cytogenetic data and RNA sequencing. The aim of the present study was to perform an unbiased search for fusion transcripts in our entire series of ependymal tumors. Fusion analysis was performed using the FusionCatcher algorithm on 12 RNA-sequenced ependymal tumors. Candidate transcripts were prioritized based on the software’s filtering and manual visualization using the BLAST (Basic Local Alignment Search Tool) and BLAT (BLAST-like alignment tool) tools. Genomic and reverse transcriptase PCR with subsequent Sanger sequencing was used to validate the potential fusions. Fluorescent in situ hybridization (FISH) using locus-specific probes was also performed. A total of 841 candidate chimeric transcripts were identified in the 12 tumors, with an average of 49 unique candidate fusions per tumor. After algorithmic and manual filtering, the final list consisted of 24 potential fusion events. Raw RNA-seq read sequences and PCR validation supports two novel fusion genes: a reciprocal fusion gene involving UQCR10 and C1orf194 in an adult spinal ependymoma and a TSPAN4-CD151 fusion gene in a pediatric, infratentorial anaplastic ependymoma. Our previously reported ALK rearrangements and the RELA and YAP1 fusions found in supratentorial ependymomas were until now the only known fusion genes present in ependymal tumors. The chimeric transcripts presented here are the first to be reported in infratentorial or spinal ependymomas. Further studies are required to characterize the genomic rearrangements causing these fusion genes, as well as the frequency and functional importance of the fusions.
INTRODUCTION

Fusion genes are known to be of pathogenetic importance in several human neoplastic entities. They are often a result of chromosomal rearrangements such as translocations, inversions, or deletions. In addition to the pathogenetic, prognostic, and diagnostic information they convey, some fusion gene products may also represent promising targets for novel therapies (Mertens et al., 2015; Shaw et al., 2013).

Traditionally, fusion genes have been identified by cytogenetic methods followed by molecular characterization of the chromosomal breakpoints involved. During the last decade, high-throughput RNA sequencing has become a widely applied, efficient approach for the detection of gene fusions in cancer, thereby in part omitting the first cytogenetic step on the way towards the detection of cancer-specific molecular-level genomic changes (Maher et al., 2009). Several different bioinformatic tools are currently available for this purpose (Wang et al., 2013).

Ependymal tumors constitute 2% of all primary CNS neoplasms. The annual incidence is about two to four cases per million inhabitants (Ostrom et al., 2014). Overall 5-year relative survival is around 70% (Crocetti et al., 2012). According to the new WHO classification, ependymal tumors are subdivided into subependymomas (WHO grade I), myxopapillary ependymomas (WHO grade I), ependymomas (WHO grade II), anaplastic ependymomas (WHO grade III), and RELA fusion-positive ependymomas (Louis et al., 2016).

It has long been known that spinal ependymomas harbor recurrent NF2 mutations (Ebert et al., 1999). Studies published during the last decade have shed considerable light on ependymal tumor biology. The knowledge resulting from this work suggests that ependymomas located in different anatomical compartments represent distinct molecular entities (Johnson et al., 2010; Taylor et al., 2005; Pajtler et al., 2015). In the case of infratentorial ependymoma, two clinically different subgroups have been identified (Mack et al., 2014; Witt et al., 2011). It was recently found that the majority of supratentorial ependymomas harbor RELA or YAP1 fusions (Pajtler et al., 2015; Parker et al., 2014; Pietsch et al., 2014).

We have previously described the chromosomal aberrations detected by karyotyping and comparative genomic hybridization (CGH) in a series of Norwegian ependymal tumor patients (Olsen et al., 2014). Using these chromosomal data combined with high-throughput RNA sequencing, we demonstrated the presence of ALK tyrosine kinase receptor rearrangements in two supratentorial ependymal tumors with considerable astrocytic differentiation (Olsen et al., 2015). The aim of the present study was to extend our search for fusion genes, now independent of low-resolution genomic data, to the remaining ependymal tumors without ALK rearrangement.
MATERIALS AND METHODS

Patients and Tumor Samples

The present analysis was performed on a series of 12 samples from 11 patients. These were collected between January 2005 and December 2012 at the Department of Neurosurgery, Oslo University Hospital, Rikshospitalet. The clinical and pathological details of these tumors are described in Table 1.

DNA and RNA Extraction, High-Throughput Sequencing, and Bioinformatic Analyses

DNA was extracted from frozen tumor tissue (Olsen et al., 2014) and RNA extraction and paired-end high-throughput sequencing (Olsen et al., 2015) was performed as previously described.

Three samples of total human brain RNA (Life Technologies, Carlsbad, CA) were used as normal controls. One sample (Normal3; catalog no. AM6050) consisted of RNA pooled from 23 adult individuals, whereas the remaining two samples (Normal1 and Normal2; catalog no. AM7962) consisted of RNA from two different adult individuals. Sequencing was performed by the Norwegian Sequencing Centre (http://www.sequencing.uio.no/) using the HiSeq 2000 platform (Illumina, San Diego, CA) with a strand-specific, paired-end TruSeq library protocol. Read length for the control samples was 125 base pairs.

The presence of fusion gene transcripts was analyzed using the algorithm FusionCatcher (Nicorici et al., 2014) with Ensembl release 75 as reference. RNA-seq reads were also aligned using TopHat version 2.0.9, as described by Trapnell et al. (2012). Gene expression in ependymal and normal samples was quantified using htseq-count (Anders et al., 2015) and DESeq2 (Anders and Huber, 2010). Candidate fusion transcripts from FusionCatcher were manually assessed using the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST; http://www.ncbi.nlm.nih.gov/blast) and the University of California Santa Cruz (UCSC) BLAT-like alignment tool (BLAT; http://genome.ucsc.edu/cgi-bin-hgBlat).

Fusion Validation by PCR and Direct Sanger Sequencing

cDNA was synthesized from extracted RNA by reverse transcription using the iScript Advanced cDNA Synthesis Kit for RT-qPCR (Bio-Rad, Hercules, USA). PCR primers were designed using Primer3 (http://primer3.ut.ee/) and synthesized by Life Technologies. All primer sequences used are listed in Supporting Information Table 1. The 25 µl PCR mix consisted of ExTaq premix (TaKaRa, Otsu-shi, Japan), 0.4 µM forward and reverse primer, and 2 µl cDNA. PCR was run on the C1000 Thermal Cycler (Bio-Rad) with an initial denaturation at 94 °C for 30 sec, followed by 35 cycles (7 sec at 98 °C, 30 sec at 58 °C, and 2 min at 72 °C) and a final extension of 5 min at 72 °C before cooling. Long-range PCR was performed using the LA PCR kit v2.1 (TaKaRa) with thermal cycles as recommended by the manufacturer.
Three µl of the PCR products were stained with GelRed (Biotium, Hayward, CA) and run on a 1 % agarose gel electrophoresis. PCR products were purified using the PCR cleanup kit (Macherey-Nagel, Düren, Germany) and direct Sanger sequencing of purified PCR products was performed via the GATC Biotech Light Run Sequencing service (http://www.gatc-biotech.com/lightrun). The Sanger sequences and chromatograms were then analyzed using the NCBI BLAST and UCSC BLAT tools.

**Fluorescence In Situ Hybridization**

Bacterial artificial chromosome (BAC) probes for the *UQCR10* (clone RP11-772E21), *C1orf194* (clone RP11-695F11), *DPYSL2* (clone RP11-1G11), and *GFAP* (clone RP11-643P22) genes were retrieved from the Human 32K BAC Re-Array library (BacPac Resource Center, Oakland, CA) (Osoegawa, 2001). Two probes overlapping the *CLU* locus (clones RP11-810P7 and RP11-16P20; BacPac Resource Center, https://bacpac.chori.org) were used. The clones were selected based on their mapping to the chromosomal sites of the putative gene fusions. Clones were grown in selective media according to the manufacturer’s protocol and BAC DNA was extracted using the High Pure Plasmid isolation kit (Roche, Basel, Switzerland). The DNA probes were labelled by nick translation using fluorescein isothiocyanate (FITC) and Texas Red (Perkin Elmer, Waltham, MA) as described previously (Nyquist et al., 2011). Interphase FISH was performed on fresh slides using frozen tumor cell suspension and the slides were counterstained using 4′,6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Burlingame, USA). CytoVision software (Leica biosystems, Wetzlar, Germany) was used to analyze the FISH images.

**Ethics Statement**

This study has been approved by the regional health research ethics committee (project number S-06046). Written informed consent was obtained from all patients or from their parents. One patient was included postmortem following particular permission from the Norwegian Directorate of Health.

**RESULTS**

On average, 96 million read pairs were obtained per tumor sample (range: 83 – 114 million). Among the normal control samples, 234 million read pairs were obtained from the pooled RNA, whereas 122 million read pairs were obtained from each of the two individual RNA samples.

Fusion analysis of the 12 tumor samples yielded a total of 841 candidate fusion transcripts. There was on average 49 unique candidate fusions per tumor sample (range: 34 – 74). 120 fusion transcripts were detected in the three control samples. Thus, there was a total of 961 fusion candidates present in the 15 samples analyzed.
Putative fusion transcripts that were tagged by FusionCatcher as “no protein product” (tumor samples: n = 305; control samples: n = 36) or “banned” (because of their known presence in healthy individuals; tumor samples: n = 39; control samples: n = 2) were excluded. The remaining 579 putative fusion transcripts were then manually visualized using the BLAST and BLAT tools with the putative fusion sequence provided by FusionCatcher.

In the cases where BLAST/BLAT reported the putative fusion sequence as one continuous transcript, the fusion candidate was considered a readthrough and disregarded (tumor samples: n = 401; control samples: n = 71). Fusion sequences that mapped to not only one or two, but several genomic regions (unspecific sequences; tumor samples: n = 48; control samples: n = 2), and masked repeat sequences (tumor samples: n = 19; control samples: n = 4) were also removed. The final list of candidate fusions consisted of 30 unique fusion gene candidates (tumor samples: n = 26; control samples: n = 4; Supporting Information Table 2). Two of these were the ALK rearrangements described in our previous paper (Olsen et al., 2015) and thus removed from the final candidate list.

In this final list (Supplementary Table 2), candidate fusions were found in 8 of 12 tumor samples (not in samples 6a, 6b, 8, and 10). Twenty-six different genes were involved in the 24 putative fusions seen in the tumor samples. Notably, the GFAP gene was involved in 11 different candidate fusion events (46 %) and CLU in eight (33 %). CHI3L1, CST3, and AGT were each involved in two different putative fusions. In one case, two reciprocal transcript variants were found (UQCR10-C1orf194 and C1orf194-UQCR10).

Six of the 24 candidate tumor fusions were validated by RT-PCR and direct Sanger sequencing of the PCR products (samples 4, 5, 7, and 11; Table 2) involving a total of 10 different genes. One of the fusions found in the normal controls, CLOCK-NMU, was also validated (N04; Table 2). In all remaining cases, PCR and Sanger sequencing results were negative.

Two of these chimeric transcripts (F12 and F19; Table 2) were found in subependymomas and involved exon 5 of the CLU gene. There was also a TSPAN4-CD151 fusion in a pediatric infratentorial ependymoma (fusion no. F05), a reciprocal fusion involving UQCR10 and C1orf194 in an adult spinal ependymoma (F02 and F02r), and a DPYSL2-GFAP fusion in a subependymoma (F15). In the case of F12, FusionCatcher originally reported this as a fusion involving CAPS and CLU. However, PCR and Sanger sequencing revealed the fusion to be complex, involving mainly material from the DIS3L and CLU genes, separated by a short sequence which was found to belong to the CAPS gene.

In order to examine how the reported chimeric transcripts related to the expression levels of the genes involved, we quantified gene expression for all the genes involved in a PCR-verified fusion (Table 2), using all reads which were properly aligned to the reference by the TopHat algorithm.
Results are shown in Supporting Material. There were no cases of obvious over- or underexpression in any of the fusion-positive samples.

We also performed FISH using custom-made BAC probes to look for genomic rearrangements that could explain the presence of the chimeric transcripts. Cell suspension from short-term tumor cultures was available for FISH analysis in cases 4, 5, and 7. Sample 4 was tested for the UQCR10-C1orf194 fusion, sample 5 was tested for the DPYSL2-GFAP fusion and splitting of the CLU gene (in two separate experiments), whereas sample 7 was tested for splitting of the CLU gene. Although CLU fusions had not been detected in sample 5, this tumor was also analyzed using FISH because this, too, was a subependymoma. The two other subependymomas of the series displayed CLU fusions and we wanted to investigate whether this aberration was present in all subependymomas of our series.

The number of nuclei available for analysis ranged from 50 to 74. In sample 4, 23 of 53 nuclei (43 %) displayed three green (C1orf194) and two red (UQCR10) signals. The remaining 30 nuclei displayed two green and two red signals. However, no fusion (i.e., yellow) signals were seen. In sample 5, 74 of 74 nuclei displayed two yellow (CLU) signals and 52 of 52 nuclei displayed two red (DPYSL2) and two green (GFAP) signals. In sample 7, 50 of 50 nuclei displayed two yellow (CLU) signals. Thus, we were not able to detect any genomic rearrangement that could explain the presence of the chimeric transcripts. Due to lack of material, we were not able to examine by FISH for the C1orf194-UQCR10 or AGT-CLU fusions.

We then looked for raw RNA-seq reads covering the fusion breakpoints by searching the raw .fastq files using the “grep” UNIX command and a sequence of 20 nucleotides; 10 from each fusion partner. This was found for fusions F02, F02r, and F05 (Table 2; reads shown in Fig. 1).

We performed fusion-specific PCR on genomic tumor DNA for fusions F02, F02r, and F05. For fusion F02, we were able to amplify a fragment containing nucleotides 109650635-109650756 from chromosome 1 and 30163373-30163426 from chromosome 22 (Fig. 2). This corresponds to a fusion between intron 1 of C1orf194 (breakpoint in chr1:109650756) and the 5’ UTR of UQCR10 (breakpoint in chr2:30163373); as originally reported by the FusionCatcher algorithm (Supporting Information Table 2).

For fusion F02r, we were able to amplify a fragment of approximately 3 kb (Fig. 3). The 5’ end of this fragment contained a 200 bp sequence from the 3’ UTR of UQCR10, whereas the 3’ end contained a 757 bp segment from intron 2 and exon 3 of C1orf194. We performed Sanger sequencing using primers located closer to the putative breakpoint (Supporting Information Table 1); however, we were not able to characterize the genomic breakpoint further.

For fusion F05, we were able to amplify a fragment of approximately 10 kb (Fig. 3) using long-range PCR. The 5’ end of this fragment contained an approximately 1 kb long sequence from
intron 1 of TSPAN4. The 3’ end contained an approximately 350 bp long sequence from intron 2 of CD151 (Fig. 3). As for the UQCR10-C1orf194 fusion, subsequent rounds of Sanger sequencing with primers closer to the putative breakpoint were not successful, as we did not obtain chromatograms of sufficient quality for sequence analysis.

**DISCUSSION**

We performed high-throughput RNA sequencing to detect novel fusion genes in a small, heterogeneous series of ependymal tumors. Six of 24 (25%) tumor-specific candidate fusion transcripts detected were also detected by reverse transcriptase PCR. In three of these transcripts, we found RNA-seq reads containing the breakpoint; two reciprocal fusions involving C1orf194 and UQCR10 in an adult ependymoma of the spinal cord, and a TSPAN4-CD151 fusion in a pediatric infratentorial ependymoma. Using genomic PCR, we were able to validate the C1orf194-UQCR10 fusion on the DNA level, with a breakpoint identical to the one provided by FusionCatcher. We were not able to characterize its corresponding reciprocal variant, UQCR10-C1orf194, precisely. However, our genomic PCR results indicate that this fusion also is present at the DNA level.

This reciprocal fusion is most likely caused by a genomic rearrangement such as a t(1;22) translocation. UQCR10-C1orf194 and C1orf194-UQCR10 do not have identical breakpoints; however, seemingly balanced translocations may be unbalanced when investigated at the nucleotide level (Mertens et al., 2015). The tumor in question, a grade II spinal ependymoma, has been investigated by G-banding (case 7 in our previous paper (Olsen et al., 2014)) but did not display any structural aberrations corresponding to these gene fusions. The karyotype was near-triploid, with numerical aberrations only. Metaphase CGH analysis of the tumor did not reveal any focal gains or losses in the UQCR10 and C1orf194 loci. Still, a substantial number of RNA-seq reads spanned the fusion breakpoints (Fig. 1). This discrepancy might be due to representative tumor cells not being able to divide in culture prior to G-banding analysis. It is also possible that the genomic rearrangement causing the fusion was too small to be detected by the low resolution level of karyotyping.

The genomic PCR results also support TSPAN4-CD151 as a fusion gene caused by a genomic rearrangement. TSPAN4 and CD151 were previously reported to be involved in cancer-related gene fusions (Asmann et al., 2011; Lapuk et al., 2012; Seo et al., 2012; Yoshihara et al., 2014). Lapuk et al. (2012) detected a WDTC1-CD151 fusion in a prostate adenocarcinoma in which exon 2 of CD151 constituted the 3’ moiety of the chimera. This fusion was validated using PCR analysis of tumor cDNA followed by Sanger sequencing. In sample 11 of our series, a pediatric infratentorial anaplastic ependymoma, exon 3 of TSPAN4 (5’) was fused to exon 3 (3’), the first coding exon of CD151. The putative fusion transcript involves the first coding exon of TSPAN4 fused to the entire coding sequence of CD151.
TSPAN4 and CD151 are neighboring genes only 4 kb apart on the plus strand of chromosome 11. They both encode members of the tetraspanin protein superfamily. TSPAN4 is normally located downstream of CD151, but in our case, TSPAN4 was the 5’ partner of the chimeric transcript. Such an arrangement could potentially be caused by a tandem duplication. This has previously been described in other tumors of the central nervous system. For example, the FGFR3-TACC3 fusion in glioblastoma is caused by a 70 kb tandem duplication on 4p16.3 (Parker et al., 2013). Tetraspanins are evolutionarily conserved, widely expressed transmembrane proteins. Interestingly, CD151 is reported to contribute to several cellular processes implicated in cancer, such as tumor initiation and progression, angiogenesis, and metastasis (Hemler, 2014; Sadej et al., 2014). To investigate further if the chimeric TSPAN4-CD151 transcript is caused by a tandem duplication, one might analyze the tumor using array CGH (aCGH). This approach can reveal copy number gains of the duplicated region as demonstrated in the FGFR3-TACC3 study by Parker et al. (2013). Another approach could be to perform whole-genome sequencing (WGS) and subsequent structural variation analysis (Nakagawa et al., 2015).

Several of the genes involved in the remaining RT-PCR validated chimeric transcripts (F12, F15, and F19; Table 2) involved highly expressed genes, such as GFAP and CLU (Supporting Material). Such highly expressed genes may generate artefactual chimeric transcripts during RNA-seq sample preparation (Sboner et al., 2010). This, and the fact that no raw RNA-seq reads contained the putative breakpoints of fusions F12, F15, and F19, lead us to believe that these transcripts may represent false positives. It is known that other artefacts may occur during cDNA synthesis. For example, template switching can generate false chimeras during reverse transcription and/or amplification (Ozsokol and Milos, 2011; Zaphiropoulos, 2011). Emerging high-throughput sequencing technologies, such as direct RNA sequencing without cDNA synthesis, might reduce the problem of technical artefacts during RNA sequencing (Ozsokol and Milos, 2011; Ozsolak et al., 2009).

Another scenario to consider is the possibility that the chimeric transcripts may have resulted from post-transcriptional events, not from genomic rearrangements. Trans-splicing, the joining of two separate RNA molecules, is known to occur frequently in trypanosomes and nematodes. It has also been suggested in mammals, including humans (Gingeras, 2009; Jividen and Li, 2014; Zaphiropoulos, 2011). The biological role of this mechanism is not known (Gingeras, 2009; Zaphiropoulos, 2011). Some have suggested that trans-splicing events may precede the formation of genomic rearrangements in cancer (Li et al., 2008; Zaphiropoulos, 2011). A recent study reported recurrent chimeric transcripts in chronic lymphocytic leukemia (CLL). Since no corresponding genomic rearrangement could be found by FISH, WGS, or Southern blotting, the authors suggested that the observed chimeras were the result of trans-splicing (Velusamy et al., 2013). However, this finding has later been debated (Vandepoele et al., 2015).
The FISH analyses performed in this study were inconclusive. Only three of six fusion transcripts could be directly tested for using this method (UQCR10-C1orf194, CAPS-DIS3L-CLU, and DPYSL2-GFAP). The numbers of nuclei available for analysis was insufficiently low in all cases (range: 50-74). In order to detect putative abnormalities in small clones, higher numbers of nuclei are needed. It is, for example, generally recommended that at least 200 interphase nuclei be counted in FISH analyses of hematological disorders (Wolff et al., 2007). In addition, the FISH was performed on short-term cultured tumor material. It is possible that the parenchymal tumor cells did not thrive and divide under in vitro conditions. Stromal cells present in the original sample may have ‘outgrown’ the neoplastic cells and skewed the results in favor of normal FISH signals, reflecting a growth selection difference that may be particularly significant in low-grade CNS neoplasms (Fuller and Perry, 2002).

Backing this suspicion is the circumstance that early karyotypic studies of oligodendrogliomas reported very low frequencies of 1p/19q deletions (Griffin et al., 1992; Thiel et al., 1992) whereas later studies, which did not depend on tumor cultures, reported the frequency of this aberration to be 60-70 % (Bello et al., 1995; Reifenberger et al., 1994). Interphase FISH on formalin-fixed, paraffin embedded tumor material, tissue slides or extracted nuclei, is not affected by the ability of neoplastic cells to divide in vitro and could in addition be performed on larger numbers of nuclei as well as on nuclei from morphologically distinct tissue areas.

We found a CLOCK-NMU chimeric transcript in one of our normal control samples (Table 2). In this case, exon 7 of CLOCK was fused to exon 2 of NMU. These two genes are located approximately 50 kb apart on the same strand of chromosome 4 and are both involved in the regulation of circadian rhythm (Martinez and O’Driscoll, 2015; Takahashi et al., 2008). The CLOCK-NMU fusion has previously been detected in a lung adenocarcinoma (Yoshihara et al., 2014). Due to the close proximity of these two genes, this might represent a readthrough fusion, something that typically occurs when two adjacent genes are spliced together (Jividen and Li, 2014; Nacu et al., 2011). However, as the CLOCK and NMU genes are separated by another gene (PDCL2), a readthrough transcription event is perhaps less likely. Other possible explanations could be cDNA-generated artefacts or trans-splicing events as previously discussed. It is also possible that the fusion is caused by a germline microdeletion without major detrimental biological consequences.

In this study, we have described the detection of two cancer-specific fusion genes in a heterogeneous series of ependymal tumors; a reciprocal UQCR10-C1orf194 fusion in an adult ependymoma of the spinal cord and a TSPAN4-CD151 fusion in a pediatric, infratentorial anaplastic ependymomoma. The genomic rearrangements causing these fusion genes are not precisely characterized. Such characterization could potentially be achieved by WGS, which enables structural variant and copy number analysis at base-pair resolution.
ACKNOWLEDGMENTS

The authors thank Lisbeth Haugom for technical assistance and Mads Motrøen for help with R/Bioconductor analyses.
REFERENCES


Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL,


FIGURE LEGENDS

**Figure 1.** Raw RNA-seq reads containing the fusion breakpoints for (A) F02; *C1orf194-UQCR10* (n = 62), (B) F02r; *UQCR10-C1orf194* (n = 3), and (C) F05; *TSPAN4-CD151* (n = 20).

**Figure 2.** Sanger chromatograms of the fusion breakpoint between *C1orf194* and *UQCR10* (F02) on the (A) RNA and (B) DNA levels. The chromatogram at the top shows the reverse transcriptase PCR fragment, whereas the chromatogram at the bottom shows the genomic PCR fragment.

**Figure 3.** Genomic PCR indicating the presence of genomic rearrangements leading to the formation of *UQCR10-C1orf194* (F02r) and *TSPAN4-CD151* (F05) fusion genes. (A) Gel electrophoresis of nested genomic PCR is shown to the left. Lane M: 1 kb DNA ladder; the strongest bands represent 6000, 3000, and 1000 bp fragments. Lane 1: An approximately 3000 bp fragment. The sequence of the 5’ end of the fragment mapped to the 5’ UTR of *UQCR10* (upper UCSC Genome Browser image) whereas the sequence of the 3’ end mapped to intron 2 and exon 3 of *C1orf194* (lower UCSC Genome Browser image). (B) Gel electrophoresis of long-range PCR fragment to the left. Lane M1: 1 kb DNA ladder; lane M2: lambda-HindIII marker; lane 2: an approx. 10 000 bp fragment. The sequence of the 5’ end mapped to intron 1 of *TSPAN4* (upper UCSC Genome Browser image) whereas the sequence of the 3’ end mapped to intron 2 of *CD151* (lower UCSC Genome Browser image).