

Article

Novel Gene Fusions in Glioblastoma Tumor Tissue and Matched Patient Plasma

Lan Wang ^{1,†}, Anudeep Yekula ^{1,†}, Koushik Muralidharan ¹, Julia L. Small ¹ , Zachary S. Rosh ¹, Keiko M. Kang ^{1,2}, Bob S. Carter ^{1,*} and Leonora Balaj ^{1,*}

¹ Department of Neurosurgery, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02115, USA; lanwang.amber@gmail.com (L.W.); ayekula@mgh.harvard.edu (A.Y.); kmuralidharan@mgh.harvard.edu (K.M.); jsmall6@mgh.harvard.edu (J.L.S.); Zrosh@mgh.harvard.edu (Z.S.R.); kmkang@mgh.harvard.edu (K.M.K.)

² School of Medicine, University of California San Diego, San Diego, CA 92092, USA

* Correspondence: bcarter@mgh.harvard.edu (B.S.C.); Balaj.Leonora@mgh.harvard.edu (L.B.)

† These authors contributed equally.

Received: 11 March 2020; Accepted: 7 May 2020; Published: 13 May 2020



Abstract: Sequencing studies have provided novel insights into the heterogeneous molecular landscape of glioblastoma (GBM), unveiling a subset of patients with gene fusions. Tissue biopsy is highly invasive, limited by sampling frequency and incompletely representative of intra-tumor heterogeneity. Extracellular vesicle-based liquid biopsy provides a minimally invasive alternative to diagnose and monitor tumor-specific molecular aberrations in patient biofluids. Here, we used targeted RNA sequencing to screen GBM tissue and the matched plasma of patients ($n = 9$) for RNA fusion transcripts. We identified two novel fusion transcripts in GBM tissue and five novel fusions in the matched plasma of GBM patients. The fusion transcripts FGFR3-TACC3 and VTI1A-TCF7L2 were detected in both tissue and matched plasma. A longitudinal follow-up of a GBM patient with a FGFR3-TACC3 positive glioma revealed the potential of monitoring RNA fusions in plasma. In summary, we report a sensitive RNA-seq-based liquid biopsy strategy to detect RNA level fusion status in the plasma of GBM patients.

Keywords: glioblastoma; gene fusions; liquid biopsy; RNA-seq; extracellular vesicles

1. Introduction

Glioblastoma Multiforme (GBM) is the most common malignant primary central nervous system tumor. It is highly aggressive, with a median overall survival of 15–23 months despite aggressive treatment [1]. A recent large scale genomic and transcriptomic sequencing analysis has shed light on the heterogeneous molecular landscape of GBM, with approximately 30–50% of patients harboring gene fusions [2]. Gene fusions are chromosomal alterations formed as a result of translocation, interstitial deletions, or insertions, resulting in a hybrid of two coding or regulatory sequences between genes [3]. Fusion transcripts are hybrid RNA produced as a result of chromosomal rearrangement (DNA level) or trans-splicing and cis splicing between adjacent genes (RNA level), and have been associated with cancer. These, in turn, generate chimeric proteins with altered functions to drive oncogenic pathways [4].

Emerging studies in cancers have demonstrated the role of oncogenic fusions, such as EML4-ALK in non-small-cell lung cancer [5], BCR-ABL1 fusion in chronic myeloid leukemia [6,7], PML-RARA in acute promyelocytic leukemia [8] and FGFR3-TACC3 in GBM [9] in tumor growth and progression. These fusions have been extensively described in literature from a diagnostic, prognostic and therapeutic standpoint [10]. Therefore, identifying and understanding gene fusions in glioma patients is critical in developing personalized diagnostic, monitoring and therapeutic modalities.

Conventionally, tumor tissue biopsy is used for histological and molecular analysis, including fusion identification [11]. In the context of CNS tumors, this procedure is highly invasive and serial biopsies are not practical to capture temporal and spatial tumor heterogeneity [12–14]. Liquid biopsy offers a minimally invasive platform for the molecular profiling of cancers from biofluids as a potential means to diagnose and monitor cancers over time and therapy [12–15]. Extracellular vesicles (EVs) are small membrane-bound nanoparticles released by all cells, including cancer cells, and have emerged as an attractive liquid biopsy modality due to their stable architecture that encloses tumor-specific genetic cargo, including mRNA, miRNA and other non-coding RNA which are reflective of the cell of origin [16]. Myriad studies over the past decade have demonstrated the utility of tumor specific mutation detection in biofluid-derived EVs for diagnosis and monitoring [17].

Novel sequencing technologies have been used to identify tumor-specific molecular signatures in the EV-derived genetic cargo [18–21], but fusion detection from biofluids has not been previously reported. RNA-seq allows genome-wide transcriptome sequencing to identify novel gene fusions, and the latest advances have enhanced the sequencing capabilities to identify molecular alterations from ultra low-input RNA [22]. Here, we identify gene fusion transcripts in the extracellular vesicle RNA (EV RNA) derived from the plasma of patients with GBM. We used a QIAseq Targeted RNAscan Human Oncology Panel, which screens 232 known fusion-related genes, to identify RNA fusions. We explore the possibility of using RNA-seq-based fusion discovery to identify known and novel fusion transcripts in tumor tissue and the matched plasma of patients with GBM.

2. Results

2.1. Experimental Design and Overview of the Library Preparation Methodologies

A cohort of adult GBM patients with prior molecular characterization ($n = 9$) and healthy controls ($n = 10$) were included in the RNA-seq exploratory fusion discovery study. Our GBM patient population included $n = 3$ patients with gene fusions (P1–P3) and $n = 6$ (P4–P9) patients with no gene fusions, as reported by Massachusetts General Hospital (MGH) Solid Fusion Assay. The demographic, histologic and molecular characteristics of the GBM patient cohort and the demographics of healthy controls are summarized in Table 1. The gene fusions analyzed by the MGH Solid Fusion Assay and RNA-seq are listed in Table S1 and Table S2, respectively. The RNA-seq analysis of the GBM tumor tissue and matched plasma of patients and healthy controls was performed using a QIAseq Targeted RNAscan Human Oncology Panel (Figure 1). Firstly, the RNA was extracted from tumor tissue using an RNeasy Mini kit, with 30 minutes of on-column DNase digestion using the RNase-Free DNase kit to eliminate the genomic DNA (gDNA). The RNA was isolated from 2 mL of plasma using an ExoRNeasy Maxi kit. All the RNA samples were analyzed using an Agilent 2100 Bioanalyzer to assess the integrity of the RNA extracted from the tumor tissue and plasma (Figure S1, Table S3).

The typical workflow of the RNA-seq library preparation and computational analysis is illustrated in Figure 1. In summary, the RNA samples are initially converted to first strand cDNA, and a second strand synthesis is used to generate double stranded cDNA (ds-cDNA). This ds-cDNA is then end-repaired, dA-tailed and ligated to a sequencing platform-specific adapter containing unique molecular tags (MTs) and sample indexes. These MTs enable original targets to be recognized by a unique sequence barcode that allows for accurate identification and quantification. Subsequently, the barcoded adapter-ligated cDNA molecules are subjected to a limited target-barcode enrichment with a single primer extension (SPE) for target enrichment. This enrichment ensures that the intended targets are sufficiently enriched to be represented in the final library. A universal PCR amplification is then performed to further amplify the library and add a second sample index. The sample libraries were generated after the library preparation and analyzed using an Agilent 2100 Bioanalyzer. The average length of the sample libraries of the tumor tissue and plasma is 400–500 bp and 350–400 bp, respectively (Figure S2). The sample libraries were quantified and pooled at equimolar ratio for sequencing.

Table 1. Demographics of Glioblastoma Multiforme (GBM) patients and healthy control cohorts.

| IDs | Age | Sex | Diagnosis | IDH | MGMT | EGFR | MET | TERT | CDKN2A |
|-----|-----|-----|-----------|-----|--------------|-------------|-------------|--------|------------|
| P1 | 72 | F | GBM | WT | Methylated | Unamplified | Unamplified | Mutant | Loss |
| P2 | 41 | F | GBM | WT | Methylated | NK | NK | WT | Maintained |
| P3 | 52 | F | GBM | WT | Methylated | NK | NK | Mutant | Maintained |
| P4 | 73 | F | GBM | WT | Unmethylated | Unamplified | Unamplified | Mutant | Loss |
| P5 | 71 | M | GBM | WT | Unmethylated | Amplified | Amplified | WT | Maintained |
| P6 | 61 | F | GBM | WT | Unmethylated | Unamplified | Unamplified | WT | Maintained |
| P7 | 68 | F | GBM | WT | Unmethylated | Amplified | NK | Mutant | Maintained |
| P8 | 65 | F | GBM | WT | Methylated | Unamplified | Unamplified | WT | Loss |
| P9 | 74 | F | GBM | WT | Methylated | Unamplified | Amplified | WT | Maintained |
| H1 | NK | NK | - | - | - | - | - | - | - |
| H2 | 24 | F | - | - | - | - | - | - | - |
| H3 | NK | NK | - | - | - | - | - | - | - |
| H4 | NK | NK | - | - | - | - | - | - | - |
| H5 | 25 | M | - | - | - | - | - | - | - |
| H6 | 24 | F | - | - | - | - | - | - | - |
| H7 | 24 | F | - | - | - | - | - | - | - |

Abbreviations: GBM, glioblastoma; WT, wildtype; IDH, isocitrate dehydrogenase; EGFR, epidermal growth factor receptor; MET, MET proto-oncogene; TERT, telomerase reverse transcriptase; CDKN2A, Cyclin Dependent Kinase Inhibitor 2A; NK, not known.

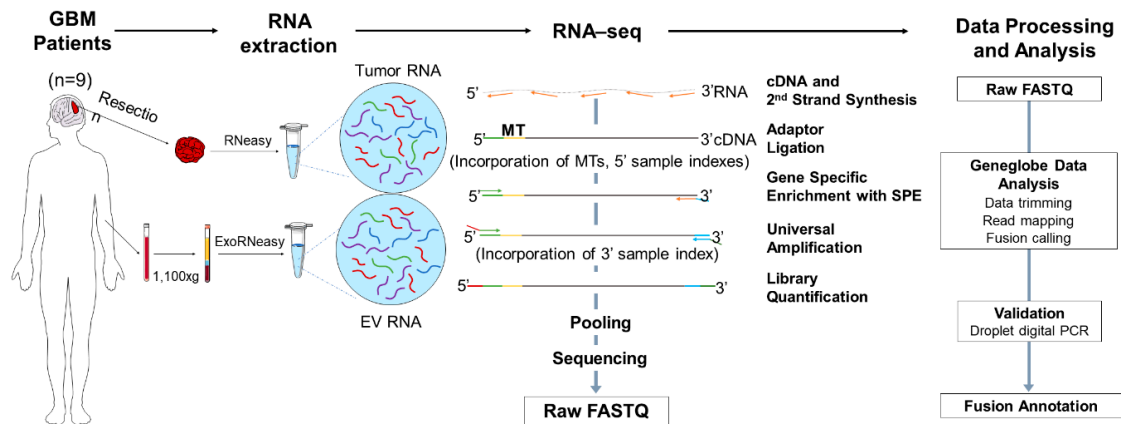


Figure 1. Overview of RNA-seq-based gene fusion discovery workflow. The RNA was extracted from the tumor tissue and matched plasma of patients with glioblastoma (GBM) ($n = 9$) and plasma of healthy controls ($n = 10$) and was sequenced using the QIAseq Targeted RNAscan Human Oncology Panel. Schematic illustration of the workflow depicting the RNA extraction, library preparation, sequencing, data processing, analysis and final validation using droplet digital PCR and fusion annotation. GBM, glioblastoma; EV RNA, extracellular vesicle RNA; MT, molecular tag; SPE, single primer extension.

A sequencing data analysis was performed in the GeneGlobe Data Analysis Center (Figure 1), and mapping metrics are shown in Table S6. The number of mapped reads were comparable between the tumor tissue and plasma samples. The average RNA Control Primers MT counts in the plasma sample libraries were lower than the 300 MT-cutoff defined by GeneGlobe, indicating a low diversity of RNA specimens in the plasma samples. The percentage of overall mapped reads longer than 200 base pairs (mapped reads% > 200 bp) was significantly lower in the plasma sample libraries than in the tumor tissue libraries, which could affect the sensitivity of fusion calling in plasma samples. The fusions detected using RNA-seq were further validated using droplet digital (ddPCR). A summary of the fusion calling and ddPCR validation results is shown in Table S4.

2.2. Detection of Fusions in Tumor Tissue and Matched Plasma of Patients with GBM and Plasma of Healthy Controls

We screened the GBM tumor tissue and matched plasma of $n = 9$ patients to identify the gene fusions identified by the MGH Solid Fusion Assay and additional novel fusion transcripts (Figure 2A).

Patients P1 (FGFR3-TACC3), P2 (FGFR3-TACC3) and P3 (AGK-BRAF) had gene fusions reported by the MGH Solid Fusion Assay. The corresponding fusion transcripts were identified in the tumor tissue of patients P1 and P3. A tumor tissue RNA-seq analysis of P2 was not performed due to inadequate tumor tissue. Patients P4-P9 did not have gene fusions reported by MGH Solid Fusion Assay. Interestingly, the RNA-seq analysis of the tumor tissue revealed no fusion transcripts in patients P4-P9, except in patient P5 (VTI1A-TCF7L2). Novel fusion transcripts were identified in the tumor tissue of P2 (VTI1A-TCF7L2), P3 (SND1-TMEM178B) and P5 (VTI1A-TCF7L2) (Figure 2B). Collectively, there was a 100% concordance between the common gene fusions screened by the MGH Solid Fusion Assay and the RNA-seq. In total, >1 fusion transcripts were detected in the tumor tissue of 22% (2/9) of GBM patients.

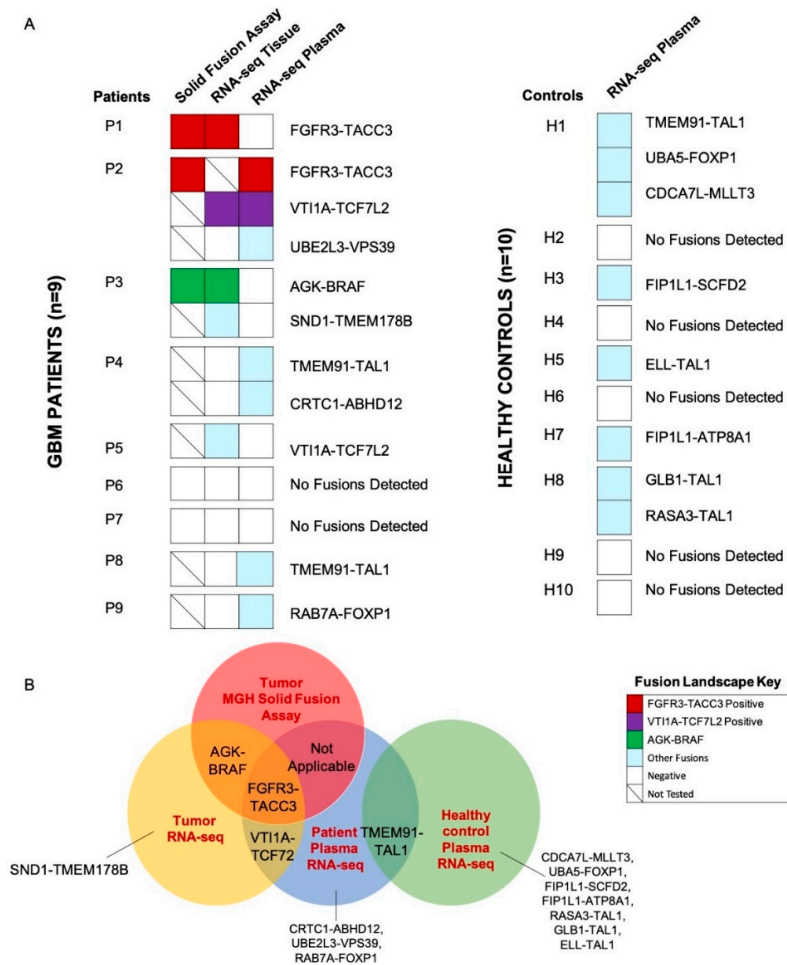


Figure 2. Tissue and plasma RNA fusion landscape. (A) Oncoprint depicting tumor tissue and plasma RNA fusions in the study cohort (GBM patients, $n = 9$, left; healthy controls, $n = 10$, right). (B) Summary of fusions detected by the Massachusetts General Hospital (MGH) Solid Fusion Assay of tumor tissue, RNA-seq of the tumor tissue and plasma of GBM patients and RNA-seq of the plasma of healthy controls.

The RNA-seq analysis of the matched plasma samples of the GBM patients (P1-P3) with gene fusions reported in the MGH Solid Fusion Assay revealed fusions in P2 (FGFR3-TACC3, VTI1A-TCF7L2). No fusion transcripts were detected in P1 and P3. FGFR3-TACC3, VTI1A-TCF7L2 fusion transcripts were detected in both the tissue and matched plasma of patient P2 (Figure 2A). The analysis of the plasma samples of GBM patients (P4-P9) with no gene fusions reported in the MGH Solid Fusion Assay revealed fusion transcripts in P4 (TMEM91-TAL1, CRTC1-ABHD12), P8 (TMEM91-TAL1) and P9 (RAB7A-FOXP1). No fusions were identified in patients P5, P6 and P7. In total, >1 fusion transcripts were detected in the plasma of 22% (2/9) of GBM patients.

Eight novel fusion transcripts were detected by an RNA-seq analysis in the healthy controls H1 (TMEM91-TAL1, CDCA7L-MLLT3, UBA5-FOXP1), H3 (FIP1L1-SCFD2), H5 (ELL-TAL1), H7 (FIP1L1-ATP8A1) and H8 (GLB-TAL1, RASA3-TAL1). TMEM91-TAL1 was the only fusion transcript commonly reported in both the patient and healthy control plasma (Figure 2B). The supporting MTs and ddPCR validation copies of the sample libraries of tumor tissue and plasma samples are reported in Table S4. The hypothetical assembly of fusions at genomic, transcriptomic and proteomic levels is further annotated in Figure 3. Furthermore, 63% (5/8) of the fusion transcripts identified in GBM patients occur in chromosomes 4, 7 and 19, which have been previously described in literature as fusion hotspots (Figure 3). Additionally, 50% (4/8) of fusion transcripts identified in the healthy controls occurred in these regions.

| | Fusion | Genomic Assembly | 5' gene Locus | 3' gene Locus | mRNA Assembly | 5' - Gene 1 Gene 2 - 3' | Fusion/Gene/Gene Product Function |
|----------------|---------------|------------------|--|--|---------------|---------------------------|---|
| GLIOMA COHORT | FGFR3-TACC3+ | | FGFR3-001 (protein_coding) chr4:1,806,934 (+) | TACC3-001 (protein_coding) chr4:1,737,598 (+) | | FGFR3 TACC3 | Oncogenic fusion, enhances glioma proliferation via activation of MAPK, ERK, PI3K, STAT3 pathways (PMID:27469539;27852792) |
| | AGK-BRAF+ | | AGK-001 (protein_coding) chr7:141,555,667 (+) | BRAF-001 (protein_coding) chr7:140,734,467 (-) | | AGK BRAF | Oncogenic fusion, enhances cell proliferation and colony formation via activation of MAPK (PMID: 24131138) |
| | VT11A-TCF7L2 | | VT11A-001 (protein_coding) chr10:112,464,657 (+) | TCF7L2-003 (protein_coding) chr10:113,141,184 (+) | | VT11A TCF7L2 | VT11A gene: Intracellular vesicular trafficking protein (PMID:11057903) TCF7L2 gene: Transcriptional factor involved in Wnt signaling (PMID:19138172) |
| | SND1-TMEM178B | | SND1-001 (protein_coding) chr7:127,721,400 (+) | TMEM178B-001 (protein_coding) chr7:141,212,591 (+) | | SND1 TMEM178B | SND, singly/occasional nuclease and tator domain: component of RNA-induced silencing complex (RISC) and involved in mRNA function (Gene ID: 275044) TMEM178B, transmembrane protein 178B (Gene ID: 106507421) |
| | UBE2L3-VPS39 | | UBE2L3-001 (protein_coding) chr22:21,611,043 (+) | VPS39-001 (protein_coding) chr15:42,199,961 (-) | | UBE2L3 VPS39 | UBE2L3, ubiquitin conjugating enzyme E2 L3; Ubiquitination (Gene ID: 7332) VPS39, VPS39 subunit of HOPS complex; Clustering and fission of late endosomes and lysosomes (Gene ID: 23339) |
| | RAB7A-FOXP1 | | RAB7A-001 (protein_coding) chr3:128,726,359 (+) | FOXP1-001 (protein_coding) chr3:71,015,653 (-) | | RAB7A FOXP1 | RAB7A, vesicular RAS oncogene family; Regulator of vesicular transport (Gene ID: 7879) FOXP1, forkhead box P1; Regulation of gene transcription (Gene ID: 27086) |
| | CRTC1-ABHD12 | | CRTC1-002 (protein_coding) chr19:18,683,828 (+) | ABHD12-001 (protein_coding) chr20:25,323,430 (-) | | CRTC1 ABHD12 | CRTC1, CREB regulated transcription coactivator 1; Transcriptional activation through cAMP response element (CRE) sites (Gene ID: 23375) and ABHD12, aldehyde domain containing 12; metabolites bioactive lysophospholipids (Gene ID: 26090) |
| HEALTHY COHORT | TMEM91-TAL1 | | TMEM91-002 (protein_coding) chr19:41,351,202 (+) | TAL1-001 (protein_coding) chr1:47,229,351 (-) | | TMEM91 TAL1 | TMEM91, Transmembrane protein 91 (Gene ID: 641649) TAL1, TAL MHLH transcription factor 1, erythroid differentiation factor; Transcriptional regulator of hematopoiesis (PMID: 30145780; Gene ID: 6886) |
| | CDCA7L-MLLT3 | | CDCA7L-001 (protein_coding) chr7:21,945,781 (-) | MLLT3-001 (protein_coding) chr9:20,354,879 (-) | | CDCA7L MLLT3 | CDCA7L, cell division cycle associated 7 like; transcriptional regulation as a repressor that inhibits histone acetylase A (Gene ID: 55536) MLLT3, MLLT3 super elongation complex subunit; increase the catalytic rate of RNA polymerase II (Gene ID: 4590) |
| | UBA5-FOXP1 | | UBA5-001 (protein_coding) chr3:132,660,698 (+) | FOXP1-001 (protein_coding) chr3:71,015,653 (-) | | UBA5 FOXP1 | UBA5, Ubiquitin-Like-Modifier Activating Enzyme 5; Post-translational modifier protein (Gene ID: 79876) FOXP1, forkhead box P1; Regulation of gene transcription (Gene ID: 27086) |
| | FIP1L1-SCFD2 | | FIP1L1-009 (protein_coding) chr4:53,399,839 (+) | SCFD2-001 (protein_coding) chr4:53,313,763 (-) | | FIP1L1 SCFD2 | FIP1L1, Factor interacting with PAPOLA and CPSF1; Generates cleavage and polyadenylation specificity factor (CPSF) which helps in mRNA poly(A) tail formation (Gene ID: 81608) SCFD2, sec1 family domain containing 2; protein transport (Gene ID: 152579) |
| | ELL-TAL1 | | ELL-001 (protein_coding) chr19:18,521,921 (-) | TAL1-001 (protein_coding) chr1:47,229,351 (-) | | ELL TAL1 | ELL, Elongation Factor for RNA Polymerase II (Gene ID: 8178) TAL1, TAL MHLH transcription factor 1, erythroid differentiation factor; Transcriptional regulator of hematopoiesis (PMID: 30145780; Gene ID: 6886) |
| | FIP1L1-ATP8A1 | | FIP1L1-002 (protein_coding) chr4:53,428,183 (+) | ATP8A1-003 (protein_coding) chr4:42,616,078 (-) | | FIP1L1 ATP8A1 | FIP1L1, Factor interacting with PAPOLA and CPSF1; Generates cleavage and polyadenylation specificity factor (CPSF) which helps in mRNA poly(A) tail formation (Gene ID: 81608) ATP8A1, ATPase for ion transport (Gene ID: 101396) |
| | GLB-TAL1 | | GLB1-001 (protein_coding) chr3:33,097,011 (-) | TAL1-001 (protein_coding) chr1:47,229,351 (-) | | GLB1 TAL1 | GLB1, beta galactosidase (Gene ID: 2720) TAL1, TAL MHLH transcription factor 1, erythroid differentiation factor; Transcriptional regulator of hematopoiesis (PMID: 30145780; Gene ID: 6886) |
| | RASA3-TAL1 | | RASA3-001 (protein_coding) chr13:114,132,435 (-) | TAL1-001 (protein_coding) chr1:47,229,351 (-) | | RASA3 TAL1 | RASA3, Ras GTPase-activating protein 3 (Gene ID: 22821) TAL1, TAL MHLH transcription factor 1, erythroid differentiation factor; Transcriptional regulator of hematopoiesis (PMID: 30145780; Gene ID: 6886) |

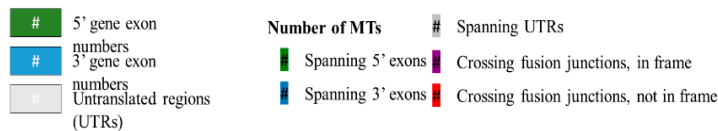


Figure 3. Annotation of gene fusions at genomic, transcriptomic and fusion protein level. Summary of genomic assembly, gene locus (QIaseq RNAscan Analysis Report generated on GeneGlobe), mRNA assembly, hypothetical fusion protein diagram and gene fusion/gene function. +, fusions with documented oncogenic function. Light blue shade indicates the GBM fusions in hotspot regions.

2.3. Longitudinal Follow-Up of Patient 2

Patient 2 presented with seizures secondary to a left frontal mass (Figure 4), with imaging features suggestive of a high-grade glioma. A biopsy was deferred in light of the patient's comorbidities.

The patient was empirically treated with chemoradiation for three months. However, due to the radiographic evidence of tumor progression, the patient underwent a subtotal resection to establish the diagnosis. A histopathologic analysis revealed an IDH wildtype diffuse astrocytoma positive for FGFR3-TACC3 fusion. The patient subsequently underwent the remaining cycles of chemoradiation. We analyzed the plasma collected pre-surgery (t_1) and post-surgery (t_2) and identified no fusions on RNA-seq analysis. There was no sufficient tumor tissue available from this time point for sequencing analysis.

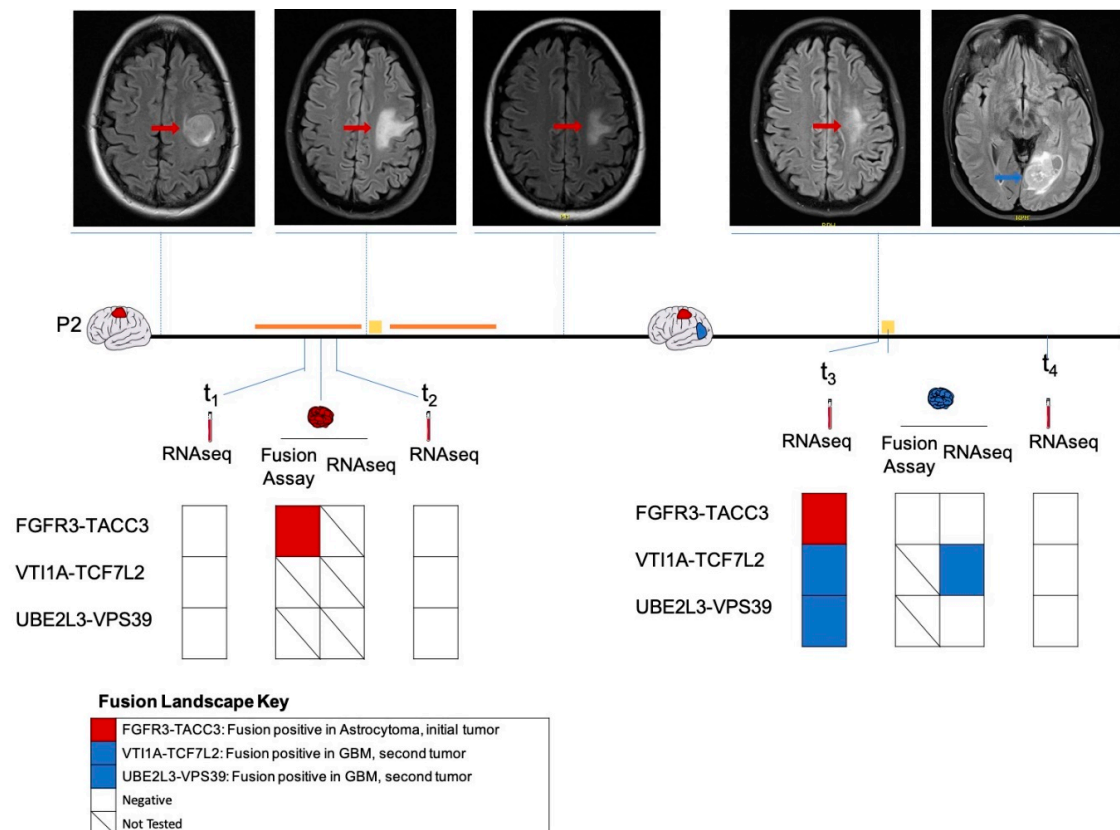


Figure 4. Longitudinal monitoring of fusion profile in patient P2. The fusion profiles of the initial frontal diffuse astrocytoma (indicated in red), plasma samples pre-surgery (t_1) and post-surgery (t_2) as well as the subsequent occipital GBM (indicated in blue), plasma samples pre-surgery (t_3) and one month post-surgery (t_4) as determined by MGH Solid Fusion Assay, RNA-seq were represented. Magnetic resonance images of the corresponding time points along the disease course were shown. Surgical procedures are indicated using a yellow square, administration of chemoradiation is indicated using an orange line.

Approximately 1.5 years following initial diagnosis, the patient developed a new left occipital lesion, unrelated to the previous tumor. MRI revealed an independent new left occipital mass in addition to the residual frontal diffuse astrocytoma. The patient underwent a subtotal resection with pathology revealing a new isocitrate dehydrogenase (IDH) wild-type GBM negative for FGFR3-TACC3 by the MGH Solid Fusion Assay (Figure 4). An RNA-seq analysis of the GBM tumor detected a VT11A-TCF7L2 fusion. Plasma collected pre-surgery (t_3) also revealed a VT11A-TCF7L2 fusion. Additionally, fusions UBE2L3-VPS39 and FGFR3-TACC3 were also detected. A plasma RNA-seq one month later revealed no fusions. We hypothesized that the FGFR3-TACC3 fusion identified in the plasma at t_3 was derived from the initial FGFR3-TACC3-positive diffuse astrocytoma.

3. Discussion

RNA fusion transcripts are major drivers of cancer, and their accurate detection and monitoring is critical to advancing clinical care. Gene fusions occur in approximately 30–50% of patients with GBM [3], and understanding the characteristics of these fusions can be critical in developing personalized diagnostic and therapeutic modalities. Advances in sequencing technologies have opened up the possibility of screening a wide range of fusions, identifying novel oncogenic driver gene fusions or passenger fusions, which can be potential biomarkers [23–25]. Currently, tissue-based genetic tumor profiling is used to classify disease and guide therapy. The standard biopsy is associated with surgical complications and residual comorbidities, and therefore obtaining multiple biopsy sections routinely to monitor the disease over time and therapy is not always feasible. Liquid biopsy is an attractive alternative for the detection and monitoring of genetic alterations in biofluids, including blood, urine, cerebrospinal fluid (CSF), saliva etc., providing a minimally invasive platform by which to monitor brain tumors. Recent studies have shown the possibility of sequencing EV RNA in various biofluids to identify distinct molecular signatures [18–21], but the identification of gene fusions from plasma has not been reported.

Here we used a targeted RNA-seq panel (QIAseq Targeted RNAscan Human Oncology Panel) to screen the GBM tissue and matched plasma samples of nine GBM patients for both previously known and novel RNA fusions, demonstrating the feasibility of using an RNA-seq-based liquid biopsy to profile the transcriptome level gene fusions in EVs derived from the plasma of patients with GBM. There was a 100% concordance between the common gene fusions detected by MGH Solid Fusion Assay and RNA-seq, which was both reassuring and indicative of the specificity of the RNA-seq assay. We observed a clustering of the novel gene fusions VTI1A-TCF7L2 and SND1-TMEM178B along with the commonly reported FGFR3-TACC3 and AGK-BRAF fusions in the GBM tissue of patients P1–P3, who were fusion-positive by the MGH Solid Fusion Assay. Only one fusion, VTI1A-TCF7L2, was identified in the GBM tissue of patients P4–P9, who were fusion-negative by the MGH Solid Fusion Assay. This indicates the increased propensity of certain GBM tumors to harbor multiple fusions in their tissue due to genomic instability. Five novel fusion transcripts TMEM91-TAL1, CRT1-ABHD12, VTI1A-TCF7L2, UBE2L3-VPS39 and RAB7A-FOXP1 were identified in the plasma of GBM patients. Eight novel fusion transcripts TMEM91-TAL1, CDCA7L-MLL3, UBA5-FOXP1, FIP1L1-SCFD2, ELL-TAL1, FIP1L1-ATP8A1, GLB1-TAL1 and RASA3-TAL1 were identified in healthy plasma. Interestingly, TMEM91-TAL1 was the only fusion common in both plasma cohorts.

The detection of the FGFR3-TACC3 fusion transcript in the plasma is of important clinical value. This fusion is formed by the confluence of its fusion partners that code for fibroblast growth factor receptor 3 (FGFR3) and transforming acidic coiled-coil (TACC3). This chimeric protein localizes to mitotic spindles, resulting in chromosomal segregation defects and the induction of aneuploidy, leading to cancer [26]. This oncogenic fusion is known to occur in 1.2% to 8.3% of patients with GBM [3,27], and activates mitogen-activated protein kinase (MAPK), extracellular signal-regulated kinase (ERK), phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) and the signal transducer and activator of transcription 3 (STAT3) pathways to promote tumor proliferation [26,28–31]. Drugs targeting FGFR kinase activity have shown promise in preclinical and clinical trials [9,26,32,33]. A real-time polymerase chain reaction (RT-PCR)-based screening assay has been developed to identify all the possible FGFR3-TACC3 variants in GBM tumor tissue, despite its structural heterogeneity at the genomic level and variability at the mRNA level [29]. Our limited capability of detection of this fusion transcript in plasma could be attributed to its heterogeneity at the genomic and transcriptomic levels, as well as the heterogeneity of plasma samples and EVs. Nevertheless, our preliminary results indicate the possibility of the detection of FGFR3-TACC3 in plasma, opening new avenues of liquid biopsy for the detection of FGFR3-TACC3 fusion transcripts to aid in the diagnosis and monitoring of FGFR3-TACC3-positive GBMs. Additionally, this fusion is reported in the context of other systemic cancers, including bladder, head and neck, lung, nasopharyngeal, esophageal squamous cell and cervical cancers [26]. Thus, an FGFR3-TACC3 liquid biopsy assay would have far-reaching benefits.

The VTI1A-TCF7L2 fusion transcript was identified in the GBM tissue of patients P2 and P3; additionally, this fusion was also identified in the plasma of patient P2 in a longitudinal setting. Although the exact function of the fusion protein has not been described, VTI1A-TCF7L2 has been implicated in several cancers, including lung, colorectal cancer (CRC), breast cancer and glioma [34,35]. The VTI1A (vesicle transport through interaction with t-SNAREs 1) gene encodes a soluble N-ethylmaleimide-sensitive fusion protein-attachment protein receptor that functions in intracellular trafficking [36], while the TCF7L2 (transcription factor 7-like 2) gene product is a high mobility group box-containing transcription factor that plays a key role in the Wnt-signaling pathway [37]. The TCF7L2 gene is frequently mutated in CRC [37] and has been demonstrated to enhance cell proliferation [38]. Considering the role of the VTI1A gene in intracellular vesicular trafficking, this fusion is an attractive candidate for EV-based fusion biomarking. Further functional studies are required to explore the oncogenic mechanisms and clinical relevance of this recurrent fusion in glioma and other cancers.

Shah et al. analyzed the transcriptome of 185 GBM tumors and identified two major genomic hotspots at the chromosomal locations 7p and 12q. Additional regions, including chromosomes 1, 4, 6 and 19, also revealed a high frequency of fusions. Furthermore, they showed that nearly 20% of GBMs have a propensity to harbor more than one fusion, and the majority of the detected fusions were private fusions occurring only in one patient [3]. We detected >1 fusion transcripts in the tumor tissue and plasma of 22% (2/9) of GBM patients and in 20% (2/10) of the healthy controls. Interestingly, 63% (5/8) of the fusion transcripts detected in GBM patients occurred in these hotspots, specifically in chromosomes 4, 7 and 19. These chromosomal locations are likely unstable in GBM, with high a propensity of forming deletion bridges which connect these amplified regions, generating multiple fusion transcripts.

We also detected four novel fusions involving the gene domain TAL1, including TMEM91-TAL1, ELL-TAL1, GLB1-TAL1 and RASA3-TAL1. Interestingly, all of them were detected in the plasma of healthy controls, while TMEM91-TAL1 was detected in the plasma of two patients with GBM and a healthy control. The TAL1 gene is an essential transcriptional regulator of hematopoiesis, in particular for the specification of hematopoietic lineages [39]. TAL1 is also implicated as a key transcriptional factor involved in the development of T-cell acute lymphoblastic leukemia [40]. Similarly, the MLLT3 (mixed lineage leukemia, translocated to 3, also known as AF9) gene is a critical regulator of hematopoietic stem cell self-renewal, and also implicated in leukemias [41–43]. The detection of fusions involving the TAL1 and MLL3 fusion domains in the plasma of healthy controls could be a consequence of their wide roles in hematopoiesis. As none of the healthy controls have a history of leukemia, it is highly likely that these TAL1 fusions are passenger fusions. The detection of passenger somatic mutations in the cell-free DNA or white blood cells of healthy individuals is not uncommon [44–47], and it could be indicative of an evolutionary path to increase protein diversity [48] or tumor formation [46,49]. This strengthens the argument that cancer progression is a fine balance of driver and passenger mutations [50]. The relevance of the fusion transcripts identified in the healthy controls is largely unknown, and further studies are needed to decipher their functional significance in health and disease. Finally, the low frequency of hematopoietic fusion detection in GBM patients compared to healthy controls could implicate the abundance of tumor-specific EVs in the plasma of GBM patients, thereby hindering any passenger fusion detection [51–53]. Additionally, the justification of using normal human plasma as a negative control cohort, in the light of evidence that fusion transcripts also occur in healthy tissue, is subject to question. The assumption that healthy controls to have no fusion transcripts could deviate and undermine our confidence in the true fusions detected in GBM patients [4].

In addition, we report several novel gene combinations (H1: CDCA7L-MLLT3, P3: SND1-TMEM178B), of which individual fusion partners have been previously reported in the context of several cancers, including MLLT3 in leukemia [54], SND1 (staphylococcal nuclease and tudor domain) in lung cancer [55,56]. These are probably coincidental and lack clinical relevance. Further investigation might provide insights into the relevance and the function of these fusions, if any.

Patient 2 initially had a frontal diffuse astrocytoma positive for an FGFR3-TACC3 fusion, and underwent chemoradiation and a subtotal resection. Later, the patient developed an occipital GBM, which had no fusions as reported by a MGH Solid Fusion Assay. RNA-seq of the occipital GBM tumor tissue revealed the presence of a VTI1A-TCF7L2 fusion. A plasma analysis prior to the resection also revealed a VTI1A-TCF7L2 fusion in addition to an FGFR3-TACC3 fusion. The FGFR3-TACC3 fusion likely originated from the initial residual frontal diffuse astrocytoma. Although the FGFR3-TACC3 fusion was not detected at the time points t_1 , t_2 or t_4 , this could probably be due to the heterogeneity in the plasma EV RNA and low sensitivity of the RNA-seq platform for low-input RNA sequencing. Nevertheless, our results indicate the feasibility of RNA-level fusion detection in plasma. This offers new avenues to expand liquid biopsy-based methods for fusion transcript detection in tumors and matched plasma for diagnostic and monitoring purposes.

One limitation of this study includes the small fragment length of EV RNA (<200 nt), which might undermine the sensitivity of the assay, as larger fragments are more favorable for fusion detection. Moreover, the low RNA yield and heterogeneity of the RNA isolated from plasma EVs limits fusion discovery and requires further optimization of the fusion discovery platforms to identify genetic alterations from the low-input EV RNA. This heterogeneity of the EV RNA between biological replicates (multiple plasma samples from the same patient) challenges the reproducibility in fusion discovery. However, this is a common limitation for any EV-based biomarker discovery study [57]. Furthermore, the fusions reported may be influenced by the biofluid in question. For instance, a multitude of fusions identified in plasma could be traced to hematopoiesis, which might mask GBM-specific fusion detection. Our analysis uses a targeted panel which is limited in the range of the fusion transcripts screened. In addition, adopting a GeneGlobe Data Analysis platform with default settings as a tool for the data analysis in this study also limits the flexibility of the data processing and fusion calling. Furthermore, the RNA-seq analysis cannot provide insights into the fusion-generating mechanism [4]. Finally, our study is limited by the small sample size. A large statistically powered study with well-defined cohorts, identifying whole transcriptome RNA fusions and the corresponding whole genome DNA fusions would be required to verify, validate and better depict the plasma GBM fusion landscape in order to better distinguish true fusions from sequencing and bioinformatics artifacts and identify clinically relevant recurrent fusions.

This proof-of-principle study demonstrates the feasibility of utilizing targeted RNA-seq for the fusion transcript profiling of EV RNA derived from the plasma of patients with GBM. The detection of fusion transcripts in biofluids can open avenues to aid in the management of patients with GBM in the contexts of diagnosis, prognosis, monitoring as well as patient stratification for novel clinical trials. Further studies with more sensitive sequencing technologies are required to explore the clinical implications of fusion transcript detection in plasma.

4. Materials and Methods

4.1. Study Subject Characteristics

The study population ($n = 9$) consisted of pathology-confirmed IDH wild-type GBM (WHO grade IV) patients ranging from 41 to 74 years who underwent surgery at Massachusetts General Hospital (MGH). Tumor tissue and matched plasma samples were collected. The patient demographics, pathological diagnosis and molecular characteristics are reported in Table 1. Plasma samples from healthy family members ($n = 3$) as well as unrelated healthy volunteers ($n = 7$) were also collected. The demographics of the healthy controls are reported in Table 1. All the samples were collected with informed consent under appropriate protocols approved by Partners institutional review board (Protocol# 2017P001581).

4.2. Tumor Tissue Processing and RNA Extraction

The tumor tissue was microdissected and suspended in RNAlater (Ambion, Austin, TX, USA) and stored at -80°C before use. The frozen tumor tissue was thawed at room temperature and the RNA was extracted using an RNeasy Mini kit according to the manufacturer's instructions, with 30 min of on-column DNase digestion using the RNase-Free DNase kit (both Qiagen, Hilden, Germany) to eliminate the genomic DNA (gDNA). Each tumor RNA sample was eluted in 40 μL nuclease-free water (Invitrogen, Carlsbad, CA, USA) and stored at -80°C before use.

4.3. Plasma processing

Whole blood samples were collected using BD Vacutainer® blood collection tubes (BD Life Science, Franklin Lakes, NJ, USA). Within 2 h of collection, the samples were centrifuged at $1100\times g$ for 10 min and filtered using 0.8 μm filters prior to being aliquoted into cryovials (Thermo Fisher Scientific, Waltham, MA, USA). All the aliquoted samples were stored at -80°C .

4.4. Plasma EV RNA Extraction

The RNA was isolated from 2 mL of plasma from the patients as well as healthy controls using the ExoRNeasy Maxi kit (Qiagen) according to the manufacturer's instructions. The EV RNA was eluted in 14 μL of nuclease-free water and stored at -80°C before use.

4.5. RNA Quantification and Quality Control

The concentration of the tumor RNA samples was determined using the Qubit RNA HS Assay Kit on a Qubit 4 fluorometer (both Thermo Fisher Scientific). The RNA was assessed for quality and integrity using the profiles generated by an Agilent Bioanalyzer. For the tumor tissue, the RNA was diluted to meet the recommended range and assessed using an Agilent RNA Nano Kit; the plasma RNA was directly assessed using an Agilent RNA Pico Kit according to the manufacturer's instructions (both Agilent Technologies, Santa Clara, CA, USA).

The tumor RNA was also assessed for gDNA contamination by a GAPDH RT-qPCR (cutoff Ct 30) using a 100 ng tumor RNA template, 2 \times Power SYBR™ Green PCR Master Mix (Applied Biosystems, Foster, KY, USA) and 4 μM of forward and reverse primers (Invitrogen, 5'-GAAGGTGAAGGTCGGAGTC-3' and 5'-GAAGATGGTGATGGGATTC-3') in a total reaction mix of 25 μL . The thermocycling conditions were as follows: 95 $^{\circ}\text{C}$ for 10 min, 40 cycles of 95 $^{\circ}\text{C}$ for 15 s and 60 $^{\circ}\text{C}$ for 1 min with data collection. The cycling was performed on a StepOnePlus Real-Time PCR System (Applied Biosystems). The Ct values were analyzed using the StepOne Software (v2.3) with auto settings.

4.6. Library Preparation and Sequencing

A QIAseq Targeted RNAscan Human Oncology Panel and QIAseq 12-Index (both Qiagen) were used for the library prep and sample indexing. An amount of 100 ng of the tumor RNA and 5 μL of the EV RNA were used as an input for the library preparation, as per the manufacturer's instructions. The sample libraries were diluted to meet the recommended concentration range and assessed using a DNA High Sensitivity Chip (Agilent Technologies) on an Agilent Bioanalyzer to determine the size distribution, and were quantified on a Qubit 4 fluorometer using a Qubit dsDNA HS Assay kit (Thermo Fisher Scientific). The molarity of the sample libraries was calculated based on the concentration and median library size. The quantified libraries were normalized to 4 nM and pooled at equimolar ratios. The sequencing libraries were finally denatured and diluted to 6 pM according to the guidelines (Illumina, San Diego, CA, USA). Paired end sequencing was performed on a MiSeq Sequencer to generate FASTQ files only, using a MiSeq Reagent Kit V2, 300 cycles (Illumina), with 231 cycles for read 1 and 71 cycles for read 2, as per the panel instructions.

4.7. Sequencing Data Processing

An integrated Qiagen panel-specific data analysis pipeline was adopted here for data processing (Figure 1). The raw individual FASTQ files were uploaded to GeneGlobe Data Analysis Center using a NGS Analysis Tool designed for the QIAseq Targeted RNAscan Panel with default settings (Qiagen, Hilden, Germany). The MT sequences were first extracted from the R2 reads for quantification and trimmed off along with adapters and other stable sequences using cutadapt software (version 1.9.1). Reads were mapped to the reference genome GENCODE Release 26 (*Homo sapiens*, GRCh38.p10) using STAR software (version 2.5.2a). The R1 reads containing gene-specific primers were first mapped with downstream endogenous bases; the unmapped segments were mapped in the 2nd pass. Quality metrics, including mapping status, RNA control primers, mapped read length and splice distance statistics were generated for a data evaluation.

4.8. Fusion Calling and Reportable Range

The GeneGlobe data analysis pipeline integrates a list of predefined filters to avoid reporting artifacts caused by sequence homology, read-through RNA or signal noise. For a fusion to be called by the GeneGlobe data analysis pipeline, the fusions must (1) contain exons from different genes within one read or read-pair; (2) be at least 12 base pairs long; (3) pass all filters or be flagged with at most one filter; (4) be curated fusions, which are extensively described in the literature (and will be reported even with limited evidence). Our final reporting strategy includes (1) curated fusions and (2) high confidence fusions, which passed all filters defined by GeneGlobe. Low confidence fusions, which have not passed all the GeneGlobe defined filters or are not curated are not reported as true fusions. Fusions COL1A1-COL1A2 and HAC1-COLQ are reported by GeneGlobe as a result of the bioinformatics artifact and are thus excluded from our analysis [58]. The summary of fusion calling results and the number of MTs supporting the fusion are provided in Table S4. Supporting reads for the fusion were used to generate contigs using SPAdes software (version 3.10.1) to obtain a fusion consensus sequence.

4.9. Validation of Fusions by Droplet Digital PCR

We validated all the positive fusions reported by RNA-seq using a droplet digital PCR (ddPCR) assay. The primers (Invitrogen) were designed to target the fusion sequence using Primer-Blast (NCBI) and were shown in Table S5 Supplementary Materials, Dataset. S2. ddPCR amplification was performed in duplicate, using 1 μ L of a 4 nM sequencing libraries template, 10 μ L 2 \times QX200 ddPCR EvaGreen Supermix (Bio-Rad, Hercules, USA) and 1 μ L of each 2.5 nM primer (Invitrogen) in a total reaction mix of 20 μ L. The QX200 AutoDG Automatic Droplet generator (both Bio-Rad) was used to generate droplets. The thermocycling conditions were as follows: 95 °C for 5 min, 40 cycles of 95 °C for 30 s and 60 °C for 1 min, followed by 4 °C for 5 min and 90 °C for 5 min and held at 4 °C until further processing. The droplets were counted and analyzed using the QX200 droplet reader, and a QuantaSoft analysis (Bio-Rad) was performed to acquire absolute quantification data. The average fusion copies per/ μ L were taken and the gates were set based on the fusion wild-type samples in the 1-D amplification plot (Figure S3). The summary of the ddPCR copies/ μ L supporting the validation of fusions is provided in Table S4.

5. Conclusions

In this study, we used targeted RNA-seq to screen tumor tissue and matched plasma samples from GBM patients for the presence of known and novel fusion transcripts. We successfully detected a range of previously reported oncogenic fusions (FGFR3-TACC3, AGK-BRAF) as well as novel fusions, including VTI1A-TCF7L2 and SND1-TMEM178B. The fusion transcripts FGFR3-TACC3 and VTI1A-TCF7L2 were detected in both the tissue and matched plasma. Furthermore, we identified an FGFR3-TACC3 fusion transcript in the patient plasma in a longitudinal setting. Our study demonstrates

the feasibility of RNA-seq-based fusion transcript detection and paves a path for liquid biopsy-based fusion transcript profiling for diagnosis and longitudinal monitoring.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/12/5/1219/s1>. Figure S1: Representative RNA Bioanalyzer profiles; Figure S2: Representative library Bioanalyzer profiles; Figure S3: Representative ddPCR 1-D amplification plot; Table S1: Genes screened by MGH Solid Fusion Assay; Table S2: Genes screened by QIAseq Targeted RNAscan Panel; Table S3: tumor RNA and EV RNA profiles; Table S4: Summary of fusion calling results; Table S5: Primers design for fusion transcript droplet digital PCR validation; Table S6: Overview of GeneGlobe RNA-seq mapping metrics.

Author Contributions: L.B. and B.S.C. conceptualized and supervised the study and designed the experiments; L.W. performed RNA-seq experiments and data analysis. J.L.S., and Z.S.R. consented patients and collected specimens and the relevant clinical data. L.W., A.Y., K.M., K.M.K., L.B. and B.S.C. generated figures and wrote the paper with input from all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by grants U01 CA230697 (BSC, LB), UH3 TR000931 (BSC), P01 CA069246 (BSC), R01 CA239078 (BSC, LB). The funding sources had no role in the writing the manuscript or the decision to submit the manuscript for publication. The authors have not been paid to write this article by any entity. The corresponding author has full access to the manuscript and assumes final responsibility for the decision to submit for publication.

Acknowledgments: We would like to thank all of the MGH Neurosurgery clinicians and staff that assisted with the collection of samples. Additionally, we would also like to thank the patients and their families for their participation in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ostrom, Q.T.; Gittleman, H.; Xu, J.; Kromer, C.; Wolinsky, Y.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009–2013. *Neuro-Oncology* **2016**, *18*, v1–v75. [[CrossRef](#)] [[PubMed](#)]
- Shah, N.; Lankovitch, M.; Lee, H.; Yoon, J.-G.; Schroeder, B.; Foltz, G. Exploration of the gene fusion landscape of glioblastoma using transcriptome sequencing and copy number data. *BMC Genom.* **2013**, *14*, 818. [[CrossRef](#)] [[PubMed](#)]
- Xu, T.; Wang, H.; Huang, X.; Li, W.; Huang, Q.; Yan, Y.; Chen, J. Gene Fusion in Malignant Glioma: An Emerging Target for Next-Generation Personalized Treatment. *Transl. Oncol.* **2018**, *11*, 609–618. [[CrossRef](#)] [[PubMed](#)]
- Kumar, S.; Razzaq, S.K.; Vo, A.D.; Gautam, M.; Li, H. Identifying fusion transcripts using next generation sequencing. *Wiley Interdiscip Rev. RNA* **2016**, *7*, 811–823. [[CrossRef](#)]
- Sabir, S.R.; Yeoh, S.; Jackson, G.; Bayliss, R. EML4-ALK Variants: Biological and Molecular Properties, and the Implications for Patients. *Cancers* **2017**, *9*, E118. [[CrossRef](#)]
- Rowley, J.D. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **1973**, *243*, 290–293. [[CrossRef](#)]
- Kang, Z.-J.; Liu, Y.-F.; Xu, L.-Z.; Long, Z.-J.; Huang, D.; Yang, Y.; Liu, B.; Feng, J.-X.; Pan, Y.-J.; Yan, J.-S.; et al. The Philadelphia chromosome in leukemogenesis. *Chin. J. Cancer* **2016**, *35*, 48. [[CrossRef](#)]
- De Braekeleer, E.; Douet-Guilbert, N.; De Braekeleer, M. RARA fusion genes in acute promyelocytic leukemia: A review. *Expert Rev. Hematol.* **2014**, *7*, 347–357. [[CrossRef](#)]
- Singh, D.; Chan, J.M.; Zoppoli, P.; Niola, F.; Sullivan, R.; Castano, A.; Liu, E.M.; Reichel, J.; Porrati, P.; Pellegatta, S.; et al. Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* **2012**, *337*, 1231–1235. [[CrossRef](#)]
- Schram, A.M.; Chang, M.T.; Jonsson, P.; Drilon, A. Fusions in solid tumours: diagnostic strategies, targeted therapy, and acquired resistance. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 735–748. [[CrossRef](#)]
- Penson, R.T.; Sales, E.; Sullivan, L.; Borger, D.R.; Krasner, C.N.; Goodman, A.K.; del Carmen, M.G.; Growdon, W.B.; Schorge, J.O.; Boruta, D.M.; et al. A SNaPshot of potentially personalized care: Molecular diagnostics in gynecologic cancer. *Gynecol. Oncol.* **2016**, *141*, 108–112. [[CrossRef](#)] [[PubMed](#)]
- Gerlinger, M.; Rowan, A.J.; Horswell, S.; Math, M.; Larkin, J.; Endesfelder, D.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P.; et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **2012**, *366*, 883–892. [[CrossRef](#)]

13. Heitzer, E.; Ulz, P.; Geigl, J.B. Circulating tumor DNA as a liquid biopsy for cancer. *Clin. Chem.* **2015**, *61*, 112–123. [CrossRef] [PubMed]
14. Cheng, F.; Su, L.; Qian, C. Circulating tumor DNA: a promising biomarker in the liquid biopsy of cancer. *Oncotarget* **2016**, *7*, 48832–48841. [CrossRef] [PubMed]
15. Cai, L.-L.; Ye, H.-M.; Zheng, L.-M.; Ruan, R.-S.; Tzeng, C.-M. Circulating tumor cells (CTCs) as a liquid biopsy material and drug target. *Curr. Drug Targets* **2014**, *15*, 965–972. [PubMed]
16. Roy, S.; Lin, H.-Y.; Chou, C.-Y.; Huang, C.-H.; Small, J.; Sadik, N.; Ayinon, C.M.; Lansbury, E.; Cruz, L.; Yekula, A.; et al. Navigating the Landscape of Tumor Extracellular Vesicle Heterogeneity. *Int. J. Mol. Sci.* **2019**, *20*, 1349. [CrossRef]
17. Figueroa, J.M.; Skog, J.; Akers, J.; Li, H.; Komotar, R.; Jensen, R.; Ringel, F.; Yang, I.; Kalkanis, S.; Thompson, R.; et al. Detection of wild-type EGFR amplification and EGFRvIII mutation in CSF-derived extracellular vesicles of glioblastoma patients. *Neuro. Oncol.* **2017**, *19*, 1494–1502. [CrossRef]
18. Huang, X.; Yuan, T.; Tschannen, M.; Sun, Z.; Jacob, H.; Du, M.; Liang, M.; Dittmar, R.L.; Liu, Y.; Liang, M.; et al. Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC. Genomics* **2013**, *14*, 319. [CrossRef]
19. Cheng, L.; Sun, X.; Scicluna, B.J.; Coleman, B.M.; Hill, A.F. Characterization and deep sequencing analysis of exosomal and non-exosomal miRNA in human urine. *Kidney Int.* **2014**, *86*, 433–444. [CrossRef]
20. Miranda, K.C.; Bond, D.T.; Levin, J.Z.; Adiconis, X.; Sivachenko, A.; Russ, C.; Brown, D.; Nusbaum, C.; Russo, L.M. Massively parallel sequencing of human urinary exosome/microvesicle RNA reveals a predominance of non-coding RNA. *PLoS ONE* **2014**, *9*, e96094. [CrossRef]
21. Ogawa, Y.; Taketomi, Y.; Murakami, M.; Tsujimoto, M.; Yanoshita, R. Small RNA transcriptomes of two types of exosomes in human whole saliva determined by next generation sequencing. *Biol. Pharm. Bull.* **2013**, *36*, 66–75. [CrossRef] [PubMed]
22. RNA sequencing: advances, challenges and opportunities.: MGH OneSearch n.d. Available online: <https://phstwlp2.partners.org:3699/eds/pdfviewer/pdfviewer?vid=1&sid=ea4e5203-5f5d-40c0-9f0e-60a37775633b%40sdc-v-sessmgr03> (accessed on 8 October 2019).
23. Pflueger, D.; Terry, S.; Sboner, A.; Habegger, L.; Esgueva, R.; Lin, P.-C.; Svensson, M.A.; Kitabayashi, N.; Moss, B.J.; MacDonald, T.Y.; et al. Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. *Genome Res.* **2011**, *21*, 56–67. [CrossRef] [PubMed]
24. Tanas, M.R.; Sboner, A.; Oliveira, A.M.; Erickson-Johnson, M.R.; Hespelt, J.; Hanwright, P.J.; Flanagan, J.; Luo, Y.; Fenwick, K.; Natrajan, R.; et al. Identification of a disease-defining gene fusion in epithelioid hemangioendothelioma. *Sci. Transl. Med.* **2011**, *3*, 98ra82. [CrossRef]
25. Steidl, C.; Shah, S.P.; Woolcock, B.W.; Rui, L.; Kawahara, M.; Farinha, P.; Johnson, N.A.; Zhao, Y.; Telenius, A.; Neriah, S.B.; et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature* **2011**, *471*, 377–381. [CrossRef]
26. Lasorella, A.; Sanson, M.; Iavarone, A. FGFR-TACC gene fusions in human glioma. *Neuro-Oncology* **2017**, *19*, 475–483. [CrossRef]
27. Na, K.; Kim, H.-S.; Shim, H.S.; Chang, J.H.; Kang, S.-G.; Kim, S.H. Targeted next-generation sequencing panel (TruSight Tumor 170) in diffuse glioma: a single institutional experience of 135 cases. *J. Neurooncol.* **2019**, *142*, 445–454. [CrossRef]
28. Costa, R.; Carneiro, B.A.; Taxter, T.; Tavora, F.A.; Kalyan, A.; Pai, S.A.; Chae, Y.K.; Giles, F.J. *FGFR3-TACC3* fusion in solid tumors: mini review. *Oncotarget* **2016**, *7*. [CrossRef] [PubMed]
29. Di Stefano, A.L.; Fucci, A.; Frattini, V.; Labussiere, M.; Mokhtari, K.; Zoppoli, P.; Marie, Y.; Bruno, A.; Boisselier, B.; Giry, M.; et al. Detection, Characterization, and Inhibition of FGFR-TACC Fusions in IDH Wild-type Glioma. *Clin. Cancer Res.* **2015**, *21*, 3307–3317. [CrossRef]
30. Stransky, N.; Cerami, E.; Schalm, S.; Kim, J.L.; Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **2014**, *5*, 4846. [CrossRef]
31. Gallo, L.H.; Nelson, K.N.; Meyer, A.N.; Donoghue, D.J. Functions of Fibroblast Growth Factor Receptors in cancer defined by novel translocations and mutations. *Cytokine Growth Factor Rev.* **2015**, *26*, 425–449. [CrossRef]
32. Zhang, J.; Zhou, Q.; Gao, G.; Wang, Y.; Fang, Z.; Li, G.; Yu, M.; Kong, L.; Xing, Y.; Gao, X. The effects of ponatinib, a multi-targeted tyrosine kinase inhibitor, against human U87 malignant glioblastoma cells. *OncoTargets and Therapy* **2014**, *2013*. [CrossRef]

33. Tabernero, J.; Bahleda, R.; Dienstmann, R.; Infante, J.R.; Mita, A.; Italiano, A.; Calvo, E.; Moreno, V.; Adamo, B.; Gazzah, A.; et al. Phase I Dose-Escalation Study of JNJ-42756493, an Oral Pan-Fibroblast Growth Factor Receptor Inhibitor, in Patients with Advanced Solid Tumors. *J. Clin. Oncol.* **2015**, *33*, 3401–3408. [[CrossRef](#)] [[PubMed](#)]
34. Lan, Q.; Hsiung, C.A.; Matsuo, K.; Hong, Y.-C.; Seow, A.; Wang, Z.; Hosgood, H.D., 3rd; Chen, K.; Wang, J.C.; Chatterjee, N.; et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat. Genet.* **2012**, *44*, 1330–1335. [[CrossRef](#)] [[PubMed](#)]
35. Wang, N.; Deng, Z.; Wang, M.; Li, R.; Xu, G.; Bao, G. Additional evidence supports association of common genetic variants in VTI1A and ETFA with increased risk of glioma susceptibility. *J. Neurol. Sci.* **2017**, *375*, 282–288. [[CrossRef](#)] [[PubMed](#)]
36. Flowerdew, S.E.; Burgoyne, R.D. A VAMP7/Vti1a SNARE complex distinguishes a non-conventional traffic route to the cell surface used by KChIP1 and Kv4 potassium channels. *Biochemical Journal* **2009**, *418*, 529–540. [[CrossRef](#)]
37. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **2012**, *487*, 330–337. [[CrossRef](#)]
38. Bass, A.J.; Lawrence, M.S.; Brace, L.E.; Ramos, A.H.; Drier, Y.; Cibulskis, K.; Sougnez, C.; Voet, D.; Saksena, G.; Sivachenko, A.; et al. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nat. Genet.* **2011**, *43*, 964–968. [[CrossRef](#)]
39. Porcher, C.; Chagraoui, H.; Kristiansen, M.S. SCL/TAL1: a multifaceted regulator from blood development to disease. *Blood* **2017**, *129*, 2051–2060. [[CrossRef](#)]
40. Tan, T.K.; Zhang, C.; Sanda, T. Oncogenic transcriptional program driven by TAL1 in T-cell acute lymphoblastic leukemia. *Int. J. Hematol* **2019**, *109*, 5–17. [[CrossRef](#)]
41. Prange, K.H.M.; Mandoli, A.; Kuznetsova, T.; Wang, S.-Y.; Sotoca, A.M.; Marneth, A.E.; van der Reijden, B.A.; Stunnenberg, H.G.; Martens, J.H.A. MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. *Oncogene* **2017**, *36*, 3346–3356. [[CrossRef](#)]
42. Chen, X.; Burkhardt, D.B.; Hartman, A.A.; Hu, X.; Eastman, A.E.; Sun, C.; Wang, X.; Zhong, M.; Krishnaswamy, S.; Guo, S. MLL-AF9 initiates transformation from fast-proliferating myeloid progenitors. *Nat. Commun.* **2019**, *10*, 5767. [[CrossRef](#)]
43. Calvanese, V.; Nguyen, A.T.; Bolan, T.J.; Vavilina, A.; Su, T.; Lee, L.K.; Wang, Y.; Lay, F.D.; Magnusson, M.; Crooks, G.M.; et al. MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. *Nature* **2019**, *576*, 281–286. [[CrossRef](#)] [[PubMed](#)]
44. Holstege, H.; Pfeiffer, W.; Sie, D.; Hulsman, M.; Nicholas, T.J.; Lee, C.C.; Ross, T.; Lin, J.; Miller, M.A.; Ylstra, B.; et al. Somatic mutations found in the healthy blood compartment of a 115-yr-old woman demonstrate oligoclonal hematopoiesis. *Genome Res.* **2014**, *24*, 733–742. [[CrossRef](#)] [[PubMed](#)]
45. Razavi, P.; Li, B.T.; Brown, D.N.; Jung, B.; Hubbell, E.; Shen, R.; Abida, W.; Juluru, K.; De Bruijn, I.; Hou, C.; et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* **2019**, *25*, 1928–1937. [[CrossRef](#)] [[PubMed](#)]
46. Liu, J.; Chen, X.; Wang, J.; Zhou, S.; Wang, C.L.; Ye, M.Z.; Wang, X.Y.; Song, Y.; Wang, Y.Q.; Zhang, L.T.; et al. Biological background of the genomic variations of cf-DNA in healthy individuals. *Ann. Oncol.* **2019**, *30*, 464–470. [[CrossRef](#)]
47. Babiceanu, M.; Qin, F.; Xie, Z.; Jia, Y.; Lopez, K.; Janus, N.; Facemire, L.; Kumar, S.; Pang, Y.; Qi, Y.; et al. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res.* **2016**, *44*, 2859–2872. [[CrossRef](#)]
48. Parra, G. Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res.* **2005**, *16*, 37–44. [[CrossRef](#)]
49. Tomasetti, C.; Vogelstein, B.; Parmigiani, G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1999–2004. [[CrossRef](#)]
50. McFarland, C.D.; Korolev, K.S.; Kryukov, G.V.; Sunyaev, S.R.; Mirny, L.A. Impact of deleterious passenger mutations on cancer progression. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2910–2915. [[CrossRef](#)]
51. Evans, S.M.; Putt, M.; Yang, X.-Y.; Lustig, R.A.; Martinez-Lage, M.; Williams, D.; Desai, A.; Wolf, R.; Brem, S.; Koch, C.J. Initial evidence that blood-borne microvesicles are biomarkers for recurrence and survival in newly diagnosed glioblastoma patients. *J. Neurooncol.* **2016**, *127*, 391–400. [[CrossRef](#)]

52. Osti, D.; Del Bene, M.; Rappa, G.; Santos, M.; Matafora, V.; Richichi, C.; Faletti, S.; Beznoussenko, G.V.; Mironov, A.; Bachi, A.; et al. Clinical Significance of Extracellular Vesicles in Plasma from Glioblastoma Patients. *Clin. Cancer Res.* **2019**, *25*, 266–276. [[CrossRef](#)] [[PubMed](#)]
53. André-Grégoire, G.; Bidère, N.; Gavard, J. Temozolomide affects Extracellular Vesicles Released by Glioblastoma Cells. *Biochimie* **2018**, *155*, 11–15. [[CrossRef](#)] [[PubMed](#)]
54. Wang, Z.; Shi, Y.; Liu, H.; Liang, Z.; Zhu, Q.; Wang, L.; Tang, B.; Miao, S.; Ma, N.; Cen, X.; et al. Establishment and characterization of a DOT1L inhibitor-sensitive human acute monocytic leukemia cell line YBT-5 with a novel KMT2A-MLLT3 fusion. *Hematol. Oncol.* **2019**, *37*, 617–625. [[CrossRef](#)] [[PubMed](#)]
55. Gurevich, R.M.; Aplan, P.D.; Humphries, R.K. NUP98-topoisomerase I acute myeloid leukemia-associated fusion gene has potent leukemogenic activities independent of an engineered catalytic site mutation. *Blood* **2004**, *104*, 1127–1136. [[CrossRef](#)] [[PubMed](#)]
56. Jang, J.S.; Lee, A.; Li, J.; Liyanage, H.; Yang, Y.; Guo, L.; Asmann, Y.W.; Li, P.W.; Erickson-Johnson, M.; Sakai, Y.; et al. Common Oncogene Mutations and Novel SND1-BRAF Transcript Fusion in Lung Adenocarcinoma from Never Smokers. *Sci. Rep.* **2015**, *5*, 9755. [[CrossRef](#)]
57. Yekula, A.; Muralidharan, K.; Kang, K.M.; Wang, L.; Balaj, L.; Carter, B.S. From laboratory to clinic: Translation of extracellular vesicle based cancer biomarkers. *Methods* **2020**. [[CrossRef](#)]
58. Li, J.; Ding, Y.; Li, A. Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. *World J. Surg. Oncol.* **2016**, *14*, 297. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).