Molecular and mechanical functions of the intermediate filament protein GFAP
Stassen, O.M.J.A.

Citation for published version (APA):
Stassen, O. M. J. A. (2016). Molecular and mechanical functions of the intermediate filament protein GFAP: Beyond the wickerwork

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
GFAPδ/GFAPα ratio correlates with astrocytoma grade and directs astrocytoma gene expression towards a more malignant profile

Oscar M.J.A. Stassen*
Emma J. van Bodegraven*
Fabrizio Giuliani
Martina Moeton
Regina Kanski
Jacqueline A. Sluijs
Miriam E. van Strien
Willem Kamphuis
Pierre A.J. Robe
Elly M. Hol

* Authors contributed equally

Manuscript ready for submission
Abstract

Astrocytomas are the most common malignant brain tumours and are to date incurable. It is unclear how astrocytomas progress into higher malignant grades. The intermediate filament cytoskeleton is emerging as an important regulator of tumour malignancy. The majority of the astrocytomas express the intermediate filament protein Glial Fibrillary Acidic Protein (GFAP). Several GFAP splice variants have been identified and the main variants expressed in human astrocytoma are the GFAPα and GFAPδ isoforms. Here we show a decrease of only GFAPα in astrocytomas with increasing grade, and a significant downregulation in glioblastoma (WHO grade IV), resulting in an increased GFAPδ/α ratio. Mimicking these changes in astrocytoma cell lines and comparing the transcriptomic changes with the changes in the patient tumours, we have identified a set of GFAPδ/α ratio-regulated tumorigenic and anti-tumorigenic genes. These genes are involved in cell proliferation and protein phosphorylation, and their expression correlated with patient survival. We additionally show that changing the ratio of GFAPδ/α, by targeting GFAP expression, affected expression of tumorigenic genes, suggesting that GFAP splicing is an attractive therapeutic target for astrocytoma.

Introduction

Astrocytomas are the most common malignant brain tumours, with an incidence of 5.9 per 100,000 in the Netherlands (Ho et al., 2014). Astrocytomas are malignant brain tumours that develop from either astrocytes, adult neural stem cells, or glia progenitors (Chen et al., 2012). These tumours are classified into different grades of malignancy based on histological assessment. Grade I (pilocytic astrocytoma) are slowly growing, localized tumours that only very seldom become anaplastic. Grade II (diffuse astrocytoma) are slowly growing tumours that tend to invade the brain more diffusely and sooner or later progress to more anaplastic grade III, or develop hyperplastic neovessels and areas of necrosis and become grade IV (glioblastoma multiforme (GBM)) tumours (Louis et al., 2007). Glioblastomas can also arise de novo (Ohgaki and Kleihues, 2013) and are by far the most common and most malignant variant of astrocytoma. Resection surgery, followed by additional radio- and chemotherapy increases survival, but disease recurrence is inevitable. The current median survival time after diagnosis is between 4 and 20 months, depending on the age, the clinical condition of the patient and the treatment (Ho et al., 2014). One of the key factors in the poor response to treatment is the heterogeneity of the cells within a single tumour, with different signalling pathways active. This makes it extremely challenging to target the entire tumour with a single-pathway drug (Patel
et al., 2014). However, a better knowledge of the molecular mechanisms of tumour cells interacting with their cellular and proteinaceous environment proves to be highly valuable to improve treatment strategies (Verhaak et al., 2010) and patient prognosis (Cancer Genome Atlas Research Network et al., 2015; Charles et al., 2012).

The intermediate filament (IF) protein family is a large family of cytoskeletal proteins that are central in the integration of cellular structure and cell signalling (Coulombe and Wong, 2004). IF proteins are involved in cellular differentiation processes (Yang et al., 2011), and are involved in the tumour biology of various malignancies. For instance, keratin 14 is essential for the metastasizing invasive front of breast cancer (Cheung et al., 2013), keratin 17 is an important transforming protein in Ewing Sarcoma (Sankar et al., 2013), whereas vimentin is a marker of the epithelial to mesenchymal transition and induces the characteristic cellular changes of this transition (Mendez et al., 2010; Thiery et al., 2009). Vimentin regulates lung cancer cell adhesion through focal adhesions and activates Slug, a transcription factor involved in epithelial mesenchymal transition in breast carcinoma (Havel et al., 2015; Virtakoivu et al., 2015). Thus, in several tumours a change in the composition of the IF-network induces the progression towards a more invasive tumour.

A loss of GFAP - the typical astrocyte IF protein - in higher grade astrocytomas was described more than 40 years ago (Velasco et al., 1980). The GFAP locus encodes for multiple GFAP isoforms, and at least 10 different splice variants are now known to be expressed in the human brain (Kamphuis et al., 2014; Middeldorp and Hol, 2011). There are several indications that a change in the stoichiometry of these isoforms correlates with astrocytoma grade. Staining with two different antibodies for either all GFAP isoforms or for GFAPδ shows a negative correlation of astrocytoma grade with total GFAP but a positive correlation with GFAPδ immunoreactivity (Brehar et al., 2014; Choi et al., 2009; Heo et al., 2012). GFAPα and GFAPδ are the most abundantly expressed isoforms in the central nervous system and differ only in their C-terminal 41 amino acids, a signalling hotspot in IF proteins due to the high content of phosphorylatable residues (Moeton et al., 2016). Both isoforms are expressed in several astrocytoma cell lines (Middeldorp et al., 2009; Moeton et al., 2014; Perng et al., 2008). GFAPα encodes the canonical GFAP-isoform and is expressed in astrocytes and highly upregulated in reactive gliosis, whereas GFAPδ is enriched in neural stem cells and subpial astrocytes in the human brain (Kamphuis et al., 2014; van Strien et al., 2014; Roelofs et al., 2005). The two isoforms differ in their 3'UTR and GFAPδ contains the alternative exon 7+, which is part of intron 7 and is not present in GFAPα (Nielsen et al., 2002; Roelofs et al., 2005).

Since various members of the IF protein family are described to be involved
in cell-extracellular matrix (ECM) interactions as well as in signalling and differentiation processes involved in tumour malignancy (Cheung et al., 2013; Ivaska et al., 2007; Leduc and Etienne-Manneville, 2015), we investigated the expression of GFAP isoforms in astrocytoma. We first analysed GFAP isoform expression in a large RNA sequencing dataset of the Cancer Genome Atlas (TCGA) and observed different levels of GFAPα expression in astrocytomas of different grade as well as a different GFAPδ/α ratio. Subsequently, modulation of GFAPα and the GFAPδ/α ratio in an astrocytoma cell model resulted in transcriptional changes, which we related to the observations in patients, and resulted in the identification of a set of tumorigenic and anti-tumorigenic genes that are regulated by GFAP.

Methods

Sample selection from The Cancer Genome Atlas

The cancer genome atlas (TCGA) contains data of 528 astrocytoma grade IV cases (also known as GBM) and 516 lower grade glioma cases. Of the latter group 168 were classified as astrocytoma grade II or grade III cases. RNA sequencing data was available for 165 GBM and the 168 astrocytoma grade II and III cases. We excluded recurrent tumours from this study and we averaged the results of the analysis on the same tumour samples (same vial). The final cohort, included in our analysis, consisted of 55 grade II, 105 grade III, and 150 grade IV astrocytoma samples (Table 1).

RNA sequencing data and statistics

An explorative differential gene expression analysis between astrocytoma patients of different grade was performed in R software for statistical computing using the limma package in bioconductor (http://www.bioconductor.org). Expression data of grade II, III, and IV patients containing normalized gene and RNA isoform expression levels were obtained from TCGA data portal (Level 3 released data downloaded June, 2015), and expression levels were extracted as upper quartile normalized RSEM (RNA-Seq by Expectation Maximization) count estimates (normalized counts). Statistical significance was tested using a Bayesian linear model fit with a Benjamini-Hochberg False Discovery Rate (FDR) corrected criterion level of \( \alpha = .1 \). Genes were considered differentially expressed if the test reached statistical significance and showed at least a 1.5 fold absolute change. Normalized isoform expression level 3 data was used to determine GFAPα and GFAPδ expression levels and for each patient sample, the GFAPδ/α ratio was calculated. Statistical difference in the GFAPδ/α ratio and the GFAPδ or GFAPα expression level between grade II, III,
GFAP isoforms in Astorcytoma malignancy

and IV was tested using the non-parametric independent 2-group Mann-Whitney U Test, since the criteria for normal distribution were not met (Shapiro-Wilk test). The p-values were corrected with an FDR correction.

**GFAP isoform modulation in astrocytoma cell lines**

To explore the transcriptional changes induced by GFAP isoform levels in astrocytoma, we performed *in vitro* GFAP isoform modulation experiments. For these experiments we used two different human astrocytoma cell lines. U251-MG has a lower endogenous level of GFAP and was used to study the effect of inducing GFAPα and GFAPδ expression. U373-MG has a high endogenous level of GFAP and was used to knockdown (KD) specific GFAP isoforms and modulate the ratio between the different isoforms. U251-MG cells were transduced with lentiviral vectors containing recombinant-mCherry (referred to as “control”), recombinant human GFAPα-IRES-eGFP (referred to as “GFAPα”) or recombinant human GFAPδ-IRES-mCherry (referred to as “GFAPδ”) as described before (Moeton et al., 2016). U251-MG cells were maintained in 1:1 Dulbecco’s modified Eagle medium (DMEM) GlutaMAX high glucose : Ham’s F-10 nutrient mix, supplemented with 100 U/ml

<table>
<thead>
<tr>
<th>Table 1. Clinical Characteristics of included astrocytoma patients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Astrocytoma</strong></td>
</tr>
<tr>
<td><strong>N (number of patients)</strong></td>
</tr>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td><strong>Median (LQ, UQ)</strong></td>
</tr>
<tr>
<td><strong>Average</strong></td>
</tr>
<tr>
<td><strong>Survival, days</strong></td>
</tr>
<tr>
<td><strong>Median (LQ, UQ)</strong></td>
</tr>
<tr>
<td><strong>Censored, %</strong></td>
</tr>
<tr>
<td><strong>Karnofsky score</strong></td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>90</td>
</tr>
<tr>
<td>70-80</td>
</tr>
<tr>
<td>&lt;70</td>
</tr>
<tr>
<td>NA</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>male</td>
</tr>
<tr>
<td>female</td>
</tr>
</tbody>
</table>

Age, survival, Karnofsky score and gender specifications of included patients as provided by the The Cancer Genome Atlas (TCGA) database. LQ: Lower quartile, UQ: Upper Quartile, NA: Not Available, i survival data was available for 41 grade II, 91 grade III and 150 grade IV astrocytoma patients
penicillin, 100 µg/ml streptomycin (1% P/S) and 10% (v/v) fetal bovine serum (FBS; all Invitrogen, Bleiswijk NL).

U373-MG human astrocytoma cells were transduced with lentiviral vectors containing shRNAs targeting either GFAPα (referred to as “GFAPα”; shRNA targeted to the 3’UTR in exon 9), GFAPpan (referred to as “GFAPpan”, shRNA targeted exon 2 which is present in all known GFAP isoforms), or with a non-targeting control shRNA construct (referred to as “NTC”) as described by Moeton et al (Moeton et al., 2014). U373-MG cells were maintained in DMEM GlutaMAX high glucose supplemented with 1% P/S and 10% FBS. Transduction efficiency was maintained by culturing cells with puromycin (2 µg/ml).

The serum concentration in the cell culture medium was changed from 10% to 2% FBS three days before an experiment. For the microarray analysis, GFAPα+ and GFAPδ+ cells or GFAPα- or GFAPpan- cells were plated in medium with 2% FBS in 6-well flexible bottom Bioflex plates coated with the laminin derived peptide YIGSR (Flexcell International, Hillsborough NC, US; Dunn Labortechnik, Asbach, DE), at a density of 5E4 cells/cm².

For maintenance, all cells were split twice a week and cultured in uncoated tissue culture plastics at 37°C, under a humidified 5% CO₂ / 95% air atmosphere.

**RNA isolation**

RNA was isolated by adding TRIzol (Bioline, London, UK; GC biotech B.V., Alphen aan den Rijn, NL) to the cell culture wells. RNA was extracted according to the manufacturer’s protocol, and precipitated in isopropyl alcohol in the presence of glycogen at -20°C overnight (Roche; Almere) (Kamphuis et al., 2012). Samples were centrifuged at 4°C and at a speed of 2.0*10⁴ g for 45 minutes (min), washed twice with cold 75% ethanol (in MilliQ (Millipore)), and the RNA pellets were dissolved in MilliQ. The 260/280 absorbance ratio was measured by spectrophotometry (NanoDrop; Thermo Scientific) to assess RNA concentration and purity.

**Real time quantitative PCR**

After treating 500 ng RNA with DNase for 2 min at 42°C (gDNA wipeout buffer), cDNA was made in 10 µl reactions using Quantiscript® Reverse Transcriptase with a mix of oligo-dT and random primers at 42°C for 30 min. The reaction was stopped by an inactivation step at 95°C for 3 min (Quantitect; Qiagen Benelux B.V., Venlo, NL).

The produced cDNA was diluted 20x in MilliQ and was used as a template in the real time quantitative polymerase chain reaction (qPCR) in 96-wells plates and
the reaction was analysed with an ABI7300 (Applied Biosystems, Life Technologies, Bleiswijk, NL). The reaction mix consisted of 1 µl template, 3.5 µl MilliQ, 0.5 µl primer mix (final concentration of 0.1 mM for each primer), and 5 µl SYBR® Green PCR Master Mix (Applied Biosystems, Life Technologies, Bleiswijk, NL). The qPCR program consisted of an incubation of 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 min at 60°C. After the qPCR reaction, a dissociation curve was made by ramping the temperature from 60°C to 95°C. Curve analysis was performed using the Sequence Detection Software version 1.4 (Applied Biosystems), with a standard threshold of 0.2 (determined to be in the log linear part of the derived curve), and an automatic determination and correction of baseline fluorescence. Primers were designed to target intron-spanning exons (when possible) of the transcripts of interest and to generate 50-125 bp long amplicons. Primer efficiencies and dissociation curves were verified using a 1:20-1:320 dilution series of U251-MG and U373-MG cDNA and the amplicon product size was verified by agarose gel electrophoresis. Specifications of primers are given in Sup. Material 1.

A set of 3 reference genes was identified based on the microarray analysis (PPP3CB, CNOT10, CLNS1A), of which the geometric mean was used to normalize the data.

**RNA labelling, hybridization, and scanning**

cRNA was made from 100 ng RNA using a Low Input Quick Amp Labelling Kit, two-colour (Agilent, Amstelveen, NL), as per manufacturer’s protocol. A mix of 825 ng of Cy3 and Cy5 labelled cRNA samples was hybridized to Human GE 4x44K v2 Microarrays (Agilent, Amstelveen, NL) according to the manufacturer’s protocol. The recombinant expression samples (N=8 per condition: control, GFAPα+, GFAPδ+) and the GFAP KD samples (N=5 for GFAPα−, N=5 for GFAPpan−, N=6 for NTC) were hybridized to the microarrays. cRNA fluorescence intensity was scanned using an Agilent DNA Microarray Scanner. Scans were made with 100% and 10% photon multiplier tube intensity, at 5 µm resolution. The microarray data have been submitted to the Gene Expression Omnibus (GSE74567).

**Microarray data processing**

Data processing was performed as described before (Bossers et al., 2010a; Orre et al., 2014a, 2014b). Briefly, data were extracted from microarray images using Feature Extraction software. Using the limma package in bioconductor (http://www.bioconductor.org), spot intensities were imported into R software for statistical computing. Spots labelled by the Feature Extraction software as saturated or non-uniformly distributed foregrounds and backgrounds, as well as visually identified artefacts on the array were omitted. In these cases the statistical
Chapter 4

analysis of differential expression was based on the remaining spots. Intensities were quantile normalized between all arrays, and sample intensities were extracted for an intensity-based analysis (Bossers et al., 2010b). Redundant probes detecting identical transcripts were averaged. All genes that did not have a condition with an average hybridisation intensity of $\log_2 > 6$ were omitted from the analysis.

**Microarray statistics**

Statistical significance was tested using a Bayesian linear model fit with an FDR corrected criterion level of $\alpha = .1$ to prevent type II errors (Smyth, 2004). The analysis of U251 and U373 cells were performed independently of one another. For recombinant expression, all three conditions (control (N=8), GFAP$\alpha^+$ (N=8), and GFAP$\delta^+$ (N=8)) were compared. For knockdown, contrasts between the two KD conditions (GFAP$\alpha^-$ (N=5), GFAPpan$^-$ (N=5)) versus NTC (N=6) were made. Microarray probes were considered to report a differentially expressed gene if the statistical test was significant and the fold change showed at least a 1.5 fold absolute change.

**Linear regression analysis**

*In vitro* differentially expressed genes (our microarray data) induced by changes in GFAP isoforms (FDR <0.1; FC > 1.5) were compared to transcriptomic (TCGA data) differences between astrocytoma of different grade (FDR <0.1; FC > 1.5). A comparison was made between *in vitro* GFAP$\alpha$ regulated genes (observed in the

![Box plots showing (a) GFAP$\alpha$, (b) GFAP$\delta$, and (c) GFAP$\delta/\alpha$ levels in grade II (n=55), III (n=105) and IV (n=150) astrocytomas (whiskers: ±1.5 × IQR; notch: 95% CI; ***: FDR < 0.001). Expression levels of GFAP-isoforms were obtained from RNAseq level 3 released normalized isoform expression data of the TCGA database. The GFAP$\delta/\alpha$ ratio was calculated for each patient. A significant decrease in GFAP$\alpha$ expression in grade IV astrocytoma compared to grade II and III (a) and no difference in GFAP$\delta$ expression between astrocytoma of different grade (b) resulted in a relative increase in GFAP$\delta$ expression compared to GFAP$\alpha$ (c), which we define as the GFAP$\delta/\alpha$ ratio.*

Figure 1. GFAP-isoform expression in astrocytoma
GFAP isoforms in Astocytoma malignancy

GFAPα, GFAPpan, GFAPα conditions) and differentially expressed genes between grade II and grade IV, and grade III and grade IV, analysed as described above. The same comparison was made for in vitro GFAPδ/α ratio regulated genes (observed in the GFAPδ+, GFAPα−, and GFAPα+ conditions). The TCGA gene expression levels of the overlapping genes were tested for a significant correlation to GFAPα, GFAPδ, and GFAPδ/α ratio in patients by fitting a linear regression model to the data. P-values were FDR adjusted. At FDR < 0.01 genes were identified as strongly correlating genes. Gene expression of the final set of (anti)-tumorigenic genes was correlated to GFAPα, GFAPδ and GFAPδ/α ratio within each astrocytoma grade in the same way. Correlation was considered significant at FDR < 0.05.

Gene ontology

To detect gene ontology (GO) clusters that were overrepresented in our datasets we used the topGO R package (Gene Ontology Consortium, 2015). We used a Fisher statistical test to test for overrepresentation, with a weighted algorithm to correct for dependency between parent-child relations between ontology clusters (Alexa et al., 2006). The test for overrepresentation of our genes of interest was performed in the context of all Entrez gene IDs considered in the statistical test for differential expression on the RNAseq data. Databases used to search for overrepresentation were the GO domains ‘Cellular compartment’ (CC), ‘Biological Process’ (BP) and ‘Molecular Function’ (MF). Annotation was performed using an annotation build of 10-Aug-2015 (org.Hs.ea.db, ID = entrez) for the analysis on our RNAseq data set.

Survival Analysis

To test whether survival and progression free survival estimates for patients with below or above median expression of a gene of interest were significantly different, we calculated survival and progression free survival probabilities with a Kaplan-Meier survival analysis using the Survival package in R. Survival and progression free survival data were available for 41 grade II and 91 grade III astrocytoma patients. Survival data was available for 150 grade IV astrocytoma patients (downloaded from TCGA January, 2016). Estimates were compared using the log rank regression analysis and were considered different at an FDR adjusted significance level of 0.05. A trend for a difference between estimates was determined at FDR < 0.1.

Immunocytochemistry

For immunocytochemical stainings, cells were fixed in 4% (w/v) paraformaldehyde (4% PFA) dissolved in phosphate buffered saline, pH 7.4 (PBS) for 15 min. Cells were washed with PBS and incubated overnight at 4°C with the primary antibody diluted in blocking buffer (50 mM Tris pH 7.4, 150 mM NaCl,
Chapter 4

a) GFAP<sup>α+</sup> vs Control

b) GFAP<sup>δ+</sup> vs Control

c) GFAP<sup>pan−</sup> vs NTC

d) GFAP<sup>α−</sup> vs NTC

e) Heatmap of GFAP isoform modulated expression

f) Array data vs qPCR data

GFAP modulation

- GFAP<sup>α+</sup>
- GFAP<sup>δ+</sup>
- GFAP<sup>pan−</sup>
- GFAP<sup>α−</sup>

Relative expression

Array fold change vs qPCR fold change

Heatmap of GFAP isoform modulated expression
GFAP isoforms in Astorcytoma malignancy

0.25% (w/v) gelatine, and 0.5% triton X-100). Cells were then washed three times with PBS and incubated with the secondary antibodies and the nuclear counterstain Hoechst 33258, diluted in blocking buffer, at room temperature for 1 hour. Cells were then washed three times with PBS, dipped in MilliQ and mounted with Mowiol (0.1 M tris-Hcl pH 8.5, 25% glycerol, 10% Mowiol (Calbiochem, Merck Millipore, Darmstadt, Germany)). Micrographs from the fluorescent stainings were taken with a Leica epifluorescent DMRD microscope. The antibody used was polyclonal anti-GFAP (1:2000; DAKO, #Z0334) and the nuclei were counterstained with Hoechst 33258 (1:1000; invitrogen).

Results

**GFAP isoform expression differs between low and high grade astrocytoma**

To get insight into the biological processes that determine the malignancy of astrocytoma and to investigate the potential role of GFAP isoforms in regulating these processes, we performed a differential gene expression analysis on RNA sequencing (RNAseq) data of low- and high-grade astrocytoma obtained from TCGA. The final cohort used in our analysis included 150 grade IV, 105 grade III, and 55 grade II astrocytoma. Normalized gene expression data was used to analyse differential gene expression between low- (II and III) and high-grade (IV) astrocytoma. GFAP expression was significantly decreased in grade IV compared to both grade II (46%, FDR= 1.58E-10) and grade III (58%, FDR= 3.92E-7) astrocytoma. Interestingly, the absolute difference in normalized gene expression of all genes analysed between low- and high-grade astrocytoma was the largest for GFAP. Grade II versus IV showed an absolute difference in GFAP expression of 3.57E5 normalized counts and grade III versus IV of 2.20E5 normalized counts. In order to determine the expression levels of the GFAP isoforms, TCGA derived RNA sequencing data

Figure 2. Differential expression analysis of astrocytoma cell lines with modulated GFAP networks
Volcano plots showing the –log10 of the FDR p-value of each gene plotted against the log2 fold change (a-d). This visualizes the general effect of GFAPα+ (a), GFAPδ+ (b), GFAPpα− (c), and GFAPα− (d) on the transcriptome. Differentially expressed genes selected for analysis with GO are represented with red dots. Names of the genes with the highest fold change have been added in black, and names of genes in GO-clusters “ECM disassembly”, “ECM organization”, “ECM”, “ECM binding” and “Extracellular Region” in blue (a-d). Validation of the microarray results by qPCR of selected genes (THNSL2, LINGO2, A2M, BASP1, OGDHL, PIEZO2, SOX11, PI15, GLUL, LAMA1, GAPDH, GFAPα, SULF1, CNOT10, PPP3CB, CLNS1A, SERP2, RAMP1, and CRYAB) showed that the fold change assessment of microarray and qPCR were correlated, ρ=0.93 (e). Clustering of the scaled expression changes revealed 4 main clusters of genes, dominated by parallel or opposite changes in expression of these genes between GFAPδ+ and GFAPα−, as indicated by different colors in the dendrogram (purple, gray, yellow, black) (f).
After finding that GFAPα and GFAPδ/GFAPα ratio correlate with astrocytoma grade (Fig. 1), all genes that were differentially expressed across astrocytoma grades in patients were identified from the TCGA database (a). To select genes that are regulated by GFAPα and GFAPδ/GFAPα ratio, we modulated the GFAP network in astrocytoma cell lines in vitro and identified differentially expressed genes regulated by GFAP (b). The cell line dataset was used to filter the astrocytoma data (c). From these combined patient/in vitro genes we calculated initial correlations between GFAP-isoforms/GFAP-ratio and target genes in the patient data and identified a set of GFAP-regulated tumorigenic and anti-tumorigenic genes. These sets are GFAP-regulated based on their in vitro regulation by GFAP isoforms. Gene sets are tumorigenic or anti-tumorigenic based on their expression pattern in astrocytoma patients (d). To identify the subset of genes relevant for astrocytoma progression and biology, the filtered genes were analyzed for survival, progression free survival, and gene ontology (e). To further reduce the risk of confounding by astrocytoma grade, a linear regression was performed between the most interesting GFAP-regulated (anti-)tumorigenic genes and GFAP expression within the different astrocytoma grades (f). Finally, 8 genes passed all these criteria and emerged as the most probable GFAP-regulated (anti-)tumorigenic genes with biological or clinical relevance in astrocytoma.
Consisting of normalized isoform expression data was used. The only known GFAP isoforms that were annotated in this dataset were GFAPα and GFAPδ. Interestingly, while canonical GFAPα expression is indeed significantly decreased in grade IV astrocytoma compared to grade II (45%, FDR = 1.20E-4; Fig. 1a) and to grade III (55%, FDR = 8.4E-4; Fig. 1a), the expression of the alternative splice variant GFAPδ is not different between astrocytoma grades (Fig. 1b). This results in a relative increase in GFAPδ compared to GFAPα, which we report as the GFAPδ/α ratio. Importantly, the GFAPδ/α ratio was significantly increased in grade IV astrocytoma compared to both grade II (220%, FDR = 5.87E-08; Fig. 1c) and grade III (177%, FDR = 4.16E-08; Fig. 1c).

Extracellular matrix genes are overrepresented in transcriptome changes due to GFAP modulation in astrocytoma cell lines in vitro

To get insight into the function of these GFAP isoforms in astrocytoma and to investigate a potential role for GFAPα and the GFAPδ/α ratio in astrocytoma malignancy, we modulated GFAP isoform expression in U373 and U251 cells by recombinant expression or silencing with shRNAs of the isoforms. We analysed the GFAP isoform induced transcriptomic changes. We generated microarray-based whole-genome gene expression profiles of GFAPα+, GFAPδ+, GFAPpan−, and GFAPα− astrocytoma cell lines. For validation of the cell lines see Sup. Material 2. The results of the microarray are represented in volcano plots (Fig. 2a-2d). Differentially expressed genes (FC>1.5, FDR<.1) are represented with red dots. GFAPα+ expression induced a differential expression of 57 genes and GFAPδ+ expression induced differential expression of 158 genes compared to control. GFAPpan− and GFAPα− cells were compared to the NTC. These resulted in 48 and 848 differentially expressed genes, respectively. Lists of the top 50 most significant hits (1.5 fold upregulated or downregulated) in the GFAPα+, GFAPδ+, GFAPpan−, or GFAPα− cells are provided in Sup. Materials 3-6.

The array data were validated by qPCR. Targets were selected to verify various fold changes and directions of change of expression patterns: THNSL2, LINGO2, A2M, BASP1, OGDHL, PIEZO2, SOX11, PI15, GLUL, LAMA1, GAPDH, GFAPα, SULF1, CNOT10, PPP3CB, CLNS1A, SERP2, RAMP1, and CRYAB. The original RNA samples that were used to generate the labelled cRNA probes for array hybridisation were used to make cDNA and analysed by qPCR (Fig. 2e). The qPCR confirmed the findings of the microarray. Overall, the array and qPCR data were highly correlated (r = 0.93, p-value<.001).

Genes with an FDR adjusted p-value of <.1 in both the recombinant expression
and the knockdown models were clustered and evaluated in a heatmap. This analysis showed that GFAP\(\alpha\) and GFAP\(\delta^+\), which both had the largest impact on the GFAP\(\delta/\alpha\) ratio, had the highest contribution to the clustering of the heatmap. These two conditions led to four main emerging patterns, that either showed an effect in the same direction or in an opposite direction for GFAP\(\alpha^+\) and GFAP\(\delta^+\) (Fig. 2f). The other experimental groups, GFAP\(\text{pan}^-\) and GFAP\(\alpha^+\) had a lower effect on GFAP\(\delta/\alpha\) ratio, and had a less prominent effect on the clustering. The individual genes and their patterns are shown in Sup. Material 7. We conclude that not the change in GFAP per se but the change in GFAP\(\delta/\alpha\) had the most pronounced effect on the transcriptome in astrocytoma cells.

The differentially expressed genes were tested for overrepresentation in a gene ontology analysis, to reveal potential functional differences related to the changes in the GFAP-network. This analysis of both the recombinant expression and knockdown cells resulted in a recurring overrepresentation of gene ontology clusters related to the cell periphery, such as “extracellular matrix”, “extracellular space” (everything extracellular except the matrix) or “plasma membrane” (Sup. Materials 8-11). Remarkably, the 80 differentially expressed genes between GFAP\(\delta^+\) and GFAP\(\alpha^+\) were also overrepresented in the same clusters (Sup. Material 12). These results show that both GFAP modulation per se, as well as modulation of the specific isoforms individually, affected genes involved in the composition of the extracellular matrix and the extracellular space. These findings corroborate our earlier work, where we showed that silencing GFAP\(\alpha\) in astrocytoma cells led to a strong increase in LAMA1 expression (Moeton et al., 2014).

GFAP\(\delta/\alpha\) ratio correlates to tumour malignancy and regulates tumorigenic genes

To identify those GFAP isoform induced transcriptomic changes that are relevant for astrocytoma malignancy, we compared our in vitro transcriptomic data to the TCGA derived patient transcriptomic data (Fig. 3a-b). In patients, GFAP\(\alpha\) expression and the GFAP\(\delta/\alpha\) ratio was significantly different between grade II and III versus grade IV astrocytoma (Fig. 1). Therefore we compared transcriptomic changes induced by in vitro modulation of GFAP\(\alpha\) (GFAP\(\text{pan}^+\), GFAP\(\alpha^+\), GFAP\(\alpha^-\)) or the GFAP\(\delta/\alpha\) ratio (GFAP\(\delta^+\), GFAP\(\alpha^+\), GFAP\(\alpha^-\)) to transcriptomic differences between grade II and III versus grade IV astrocytoma (Fig. 3c).

Fig. 4a shows a Venn diagram for the comparison of genes regulated by a change in GFAP\(\alpha\) in vitro and genes differentially expressed between astrocytoma of low and high grade. Of the 865 genes that were regulated by a change in GFAP\(\alpha\)
GFAP isoforms in Astocytoma malignancy

In vitro expression (FDR < 0.1, FC > 1.5), 315 genes were differentially expressed between astrocytoma of low and high grade (FDR < 0.1, FC >1.5) and showed a similar direction of change as GFAPα. The comparison between genes regulated by a change in the GFAPδ/α ratio in vitro and differentially expressed genes between astrocytoma of low and high grade is visualized in Fig. 4b. Of the 910 genes regulated by a change in GFAPδ/α, 351 were found to be differentially expressed between astrocytoma of low and high grade and showed a similar direction of change as the

Figure 4. Comparison between GFAP-regulated genes in astrocytoma cells and patients

Venn Diagrams show the overlap of genes differentially expressed between low and high grade astrocytoma in patients (left circles) and genes regulated by a change in the GFAP network in vitro (right circles) (a,b). 315 genes are both differentially expressed between astrocytoma of low and high grade as well as regulated by a change in GFAPα expression in vitro (a). 351 genes are both differentially expressed between astrocytoma of low and high grade as well as regulated in vitro by a change in the GFAPδ/α ratio (b). The intersect of both diagrams represent the genes identified as malignancy related GFAPα or GFAPδ/α regulated genes. For these genes, a linear regression analysis was performed. Correlation to GFAPα and GFAPδ expression, and the GFAPδ/α ratio in patients ranging from grade II to grade IV was determined. The heatmap shows the genes that significantly correlated to either GFAPα or the GFAPδ/α ratio (FDR < 0.01) (c). The colour key indicates the correlation coefficient (ρ). Hierarchical clustering on the absolute correlation coefficient results into the identification of two main clusters. Cluster 1 consists of genes that positively correlate to GFAPα but negative to the GFAPδ/α ratio, cluster 2 shows the opposite pattern with genes negatively correlating to GFAPα and positively to the GFAPδ/α ratio. A larger version of this figure is provided in Supplemental Material 13.
GFAPδ/α ratio.

We classified the overlapping genes between our *in vitro* and patient data set as malignancy related GFAPα regulated genes or malignancy related GFAPδ/α regulated genes. A linear regression analysis was performed on the overlapping genes (Fig. 3d), to test their correlation to GFAPα, GFAPδ and GFAPδ/α within grade II to IV astrocytoma. Genes that significantly correlated to either GFAPα or the GFAPδ/α ratio (FDR < 0.01) were clustered in a heatmap (Fig. 4c). This hierarchical

Figure 5. Kaplan-Meier survival and progression free survival curves grade III astrocytoma

The graphs show survival and progression free survival curves of below and above median expression of GFAP regulated anti-tumorigenic (a, b) and tumorigenic genes (c, d). Survival and progression free survival curves are shown for genes of which a high expression predicts both a significantly shorter survival and a progression free survival compared to low expression, or vice versa (FDR < 0.05). Low expression of the anti-tumorigenic genes AKRIC3, ARHGEF10L, SMAD7 and ST3GAL6 predicts a significant shorter survival (a) and progression free survival (b). High expression of the tumorigenic genes C7orf46, DUSP4, HS3T3B1, ODZ1, TTL, VAV3 and ZDHHC23 predicts a significant shorter survival (c) and progression free survival (d).
clustering resulted in the identification of two large gene clusters, i.e. genes that correlated positively to GFAPα and negatively to the GFAPδ/α ratio and genes that correlated negatively to GFAPα and positively to the GFAPδ/α ratio. These genes were classified as GFAP-regulated, since their expression level was also significantly regulated in the in vitro cells with modulated GFAP-networks. Based on their astrocytoma correlation, genes were classified as tumorigenic or anti-tumorigenic genes (Fig. 3a).

We observed a positive correlation for GFAPα, a negative correlation for GFAPδ/α ratio, and no correlation for GFAPδ between astrocytoma of low and high grade (Fig. 1). Therefore, we selected genes that show a strong correlation to both GFAPα and GFAPδ/α ratio in opposite directions (FDR < 0.01), but do not correlate to GFAPδ (FDR > 0.1) as the most relevant tumorigenic or anti-tumorigenic genes regulated by GFAP (Sup. Material 13 gives detailed information of the gene clusters in Fig. 4c). This resulted in 43 genes that were identified as GFAP-regulated anti-tumorigenic genes due to their positive correlation to GFAPα and negative correlation to GFAPδ/α (Sup. Material 14). The 37 genes that correlated negatively to GFAPα while they correlated positively to GFAPδ/α were identified as GFAP-regulated tumorigenic genes (Sup. Material 15).

GFAP-regulated genes are involved in tumour biology

The GFAP-regulated tumorigenic and anti-tumorigenic genes were tested for overrepresentation in a GO analysis to gain an insight in the function of GFAPα and the GFAPδ/α ratio in astrocytoma (Fig. 3e). Table 2 shows significant GO clusters overrepresented by a minimum of 5 of these genes. Interestingly, three biological processes highly related to tumour malignancy, ‘mitotic cell cycle’, and ‘regulation of cell proliferation’ related to tumour growth and ‘regulation of phosphorylation’ involved in the activation or deactivation of many tumour malignancy related signalling pathways, were significantly overrepresented. As shown in Table 2, both GFAP-regulated anti-tumorigenic and tumorigenic genes were overrepresented within these GO clusters. These results show that the GFAPα and GFAPδ/α ratio levels regulate cellular expression patterns of genes that are involved in regulating biological processes known to be different between astrocytoma of low and high grade. This supports a role for GFAP isoforms and the GFAPδ/α ratio in astrocytoma malignancy.

GFAP-regulated tumorigenic and anti-tumorigenic genes are associated with patient survival

In order to test whether differences in GFAPα and GFAPδ/α levels contribute to
patient outcome, we performed Kaplan-Meier analyses for patient survival within each astrocytoma grade patient group (Fig. 3e). No significant differences were found between survival and progression free survival estimates for grade II, III, or IV astrocytoma patients with below (low) or above (high) median GFAPα expression or a below or above median GFAPδ/α ratio. These results indicate that GFAPα and the GFAPδ/α ratio per se do not directly modulate tumour characteristics that affect survival of patients within one astrocytoma grade. However, the survival estimates for below and above median expression of the GFAP-regulated tumorigenic or anti-tumorigenic genes were significantly different for 32 of these genes in astrocytoma grade III patients. Interestingly, high expression of the tumorigenic genes and low expression of the anti-tumorigenic genes was associated with a worse survival probability (FDR < 0.05) (Sup. Material 16). Moreover, high expression of 7 of the tumorigenic genes (C7orf46, DUSP4, HS3ST3B1, ODZ1, TTL, VAV3, ZDHHC23) or low expression of 4 of the anti-tumorigenic genes (AKR1C3, ARHGEF10L, SMAD7, ST3GAL6) was associated with a lower progression free survival probability of grade III astrocytoma patients (Fig. 5).

To further investigate the regulation of these genes of interest by GFAP, we calculated the strength of correlation to GFAPα and the GFAPδ/α ratio within astrocytoma grade (Fig. 3f). This reduces the influence of astrocytoma grades on the correlation of GFAP with expression levels of a gene. For this analysis, we only considered genes that were overrepresented in one of the gene ontology clusters or that were associated with a worse progression free survival probability. We found that 4 of the genes that we identified as GFAP-regulated tumorigenic genes: i.e. NTRK2, FBXO2, ST3GAL6, and AKR1C3, also correlated with GFAPα or GFAPδ/α ratio within astrocytoma grade (Fig. 6a-h, corresponding statistics in Sup. Material 17). The set of genes identified as GFAP-regulated anti-tumorigenic genes, also contained 4 genes which showed the GFAP-dependent pattern within one or more of the astrocytoma grades (NOS2, SPC24, VAV3, and DUSP4; Fig. 7a-h, corresponding statistics in Sup. Material 17).

The genes that emerged from this final analysis are the genes we consider most likely to be regulated by a change in GFAP-isoforms in astrocytoma. We show that the regulation of their expression could influence astrocytoma malignancy through alterations in biological processes highly involved in the regulation of tumour malignancy (mitosis, proliferation, phosphorylation). In addition, regulation of these genes by a change in GFAP-isoforms could alter the biology of the tumour in such a way that it influences disease progression and survival of astrocytoma grade III patients.
Table 2. Overrepresented GO clusters by GFAP-regulated (anti-)tumorigenic genes

<table>
<thead>
<tr>
<th>GO ID</th>
<th>GO Term</th>
<th># sig gene</th>
<th>p-value</th>
<th>GFAP-regulated</th>
<th>Anti-Tumorigenic Genes</th>
<th>GFAP-regulated Tumorigenic Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Biological Processes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0000278</td>
<td>mitotic cell cycle</td>
<td>8</td>
<td>3.2E-4</td>
<td>TTYH1</td>
<td></td>
<td>SPC24, CENPP, CENPQ, PRIM1, RBBP8, RFC4, E2F8</td>
</tr>
<tr>
<td>GO:0042127</td>
<td>regulation of cell proliferation</td>
<td>9</td>
<td>0.020</td>
<td>EDNRB, SGK3, FBXO2, NTRK2, PURA, AKR1C3</td>
<td></td>
<td>VAV3, IGBP5, NOS2</td>
</tr>
<tr>
<td>GO:0042325</td>
<td>regulation of phosphorylation</td>
<td>10</td>
<td>0.043</td>
<td>DMD, EDNRB, SASH1, SMAD7, NTRK2</td>
<td></td>
<td>VAV3, DUSP4, ECT2, BAR1, PDGF</td>
</tr>
<tr>
<td></td>
<td><strong>Cellular Compartment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0005829</td>
<td>cytosol</td>
<td>20</td>
<td>0.006</td>
<td>DMD, CYS1, SGK3, FBXO2, SMAD7, NTRK2, FMN2, MID11P1, AHNAK, AKR1C3, GTPBP1</td>
<td></td>
<td>VAV3, POLR3G, IGBP2P3, SPC24, ECT2, CENPP, NOS2, CENPQ, DCTPP1</td>
</tr>
<tr>
<td>GO:0031090</td>
<td>organelle membrane</td>
<td>17</td>
<td>0.026</td>
<td>ST3GAL6, CPE, EDNRB, CYS1, TMEM59L, FBXO2, NTRK2, FMN2, SLC44A2, TTYH1, AHNAK, EFHD1</td>
<td></td>
<td>ECT2, MANE, MYO19, PDGF, HS3ST3B1</td>
</tr>
<tr>
<td>GO:0045121</td>
<td>membrane raft</td>
<td>7</td>
<td>5.6E-6</td>
<td>DMD, EDNRB, CYS1, FAIM2, MALAHNAK</td>
<td></td>
<td>CXADR</td>
</tr>
<tr>
<td></td>
<td><strong>Molecular Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0042803</td>
<td>protein homodimerization activity</td>
<td>6</td>
<td>0.006</td>
<td>NTRK2, MYOM1</td>
<td></td>
<td>ECT2, NOS2, BAR1, E2F8</td>
</tr>
</tbody>
</table>

Genes identified as GFAP-regulated (anti-)tumorigenic genes were analyzed for gene ontology overrepresentation, in the Biological Processes, Cellular Compartment, and Molecular Function domain. GO ID: Gene ontology identifier, # sig gene: Number of significant genes in the GO term.
Discussion

Intermediate Filament proteins, such as vimentin and keratin, emerge as important regulators in various malignancies (Cheung et al., 2013; Mendez et al., 2010). GFAP is the signature IF of astrocytoma cells, nevertheless the function of GFAP in astrocytoma biology is still unclear. In order to gain new insights into astrocytoma biology, we studied GFAP-isoform expression in resected astrocytoma from patients (TCGA database) and the downstream transcriptional changes in astrocytoma cells caused by modulation of these GFAP isoforms. With this approach we identified anti-tumorigenic and tumorigenic genes that are regulated by GFAP. These genes regulate cellular processes highly related to tumour malignancy as apparent from our GO analysis, and the expression levels have a prognostic value for astrocytoma grade III patients.

Our earlier work shows that an experimentally induced shift in the GFAPδ/α ratio leads to changes in cellular motility and morphology (Moeton et al., 2014, 2016). Therefore, we hypothesized that the stoichiometry of the GFAP-isoforms determines specific molecular and functional changes in astrocytoma cells related to their malignancy. Indeed the most pronounced transcriptional changes in our study were induced by the largest increase in GFAPδ/α ratio. More importantly, the increase in this ratio correlated with an increase in astrocytoma malignancy. Although GFAP immunoreactivity is decreased with increasing astrocytoma grade (Velasco et al., 1980), it is also known that a complete lack of all GFAP-isoforms, as in the GFAP−/− mice, is not sufficient to increase tumorigenicity (Wilhelmsson et al., 2003). We here show that it is the ratio between different GFAP-isoforms, and not the mere expression level of GFAP, that contributes to malignancy. In fact, a positive correlation between GFAPδ and the degree of astrocytoma malignancy was described earlier in a small number of patients (Brehar et al., 2014; Choi et al., 2009; Figure 6. Linear regression analysis within astrocytoma grades for GFAP-regulated anti-tumorigenic genes

Scatterplots show the distribution of GFAPα normalized isoform expression and the GFAPδ/α ratio (=GFAPδ/GFAPα normalized isoform expression) plotted against the normalized gene expression of each gene of interest. The black line shows the regression line for a correlation when all patients are included in the analysis. Red squares depict astrocytoma grade II patients, green dots astrocytoma grade III, and blue crosses depict astrocytoma grade IV patients. Dotted lines show the regression line for a correlation within grade II (red), grade III (green), and grade IV (blue) astrocytoma. (a-h) GFAP regulated anti-tumorigenic genes that significantly correlated to GFAPα or GFAPδ/α (FDR < 0.05) in either grade II, III or IV, and were related to patient survival and progression free survival (AKR1C3, ST3GAL6) and/or were present in one of the three significantly overrepresented GO Biological Processes clusters (NTRK2, FBXO2, AKR1C3) are shown here. The results of the linear regression analysis are reported in Supplemental Material 17.
Chapter 4

Heo et al., 2012).

**GFAP-regulated tumorigenic genes in tumour biology**

We identified four GFAP-regulated tumorigenic genes, i.e. VAV3, NOS2, DUSP4 and SPC24, that are most relevant for astrocytoma malignancy and that are induced upon an increase in the GFAPδ/α ratio both *in vitro* and in astrocytoma patients (Fig. 6a-h).

VAV3 is a guanine nucleotide exchange factor (GEF) that is more expressed in high grade astrocytoma compared to lower grade tumours (our analysis and (Liu et al., 2014a; Salhia et al., 2008)). High VAV3 expression in grade IV astrocytoma patients has been linked to a worse survival probability (Liu et al., 2014a; Salhia et al., 2008) while we find this link in grade III astrocytoma patients. Interestingly, VAV3 is a marker for glioma-initiating cells (GICs) (Liu et al., 2014a). These cells are therapy resistant, highly invasive, and, because of their self-renewing capacity, they can initiate tumour recurrence stimulated by VAV3 expression (Liu et al., 2014a). In addition, in GICs and in different astrocytoma cell lines, VAV3 modulates cell migration and invasion (Liu et al., 2014a; Salhia et al., 2008), probably by targeting the Rho-GTPase Rac1 (Kwiatkowska et al., 2012; Movilla and Bustelo, 1999; Salhia et al., 2008). In patients, Rac1 protein expression is increased in high-grade astrocytoma (Kwiatkowska et al., 2012; Salhia et al., 2008). Our data suggest that this pathway is regulated by a change in the GFAP-IF-network composition. Interestingly, an interaction between the IF protein vimentin, VAV2, and Rac1 controls the formation of focal adhesions via FAK, and together they regulate the invasive capacity of lung cancer cells (Havel et al., 2015). Such an interaction between focal adhesion components and IFs is a recurring theme in IF biology (Leube et al., 2015). Changes in the expression of GFAP-isoforms regulate expression of genes involved in focal adhesion formation (Moeton et al., 2014) and the size of focal adhesions in

**Figure 7. Linear regression analysis within astrocytoma grade for GFAP-regulated tumorigenic genes**

Scatterplots show the distribution of GFAPα normalized isoform expression and the GFAPδ/α ratio (=GFAPδ/GFAPα normalized isoform expression) plotted against the normalized gene expression of each gene of interest. The black line shows the regression line for a correlation when all patients are included in the analysis. Red squares depict astrocytoma grade II patients, green dots astrocytoma grade III, and blue crosses depict astrocytoma grade IV patients. Dotted lines show the regression line for a correlation within grade II (red), grade III (green) and grade IV (blue) astrocytoma. (a-h) GFAP regulated tumorigenic genes that significantly correlated to GFAPα or GFAPδ/α (FDR < 0.05) in either grade II, III or IV, and were related to patient survival (VAV3, DUSP4) and/or were present in one of the three significantly overrepresented GO Biological Processes clusters (DUSP4, SPC24, VAV3, NOS2) are shown here. The results of the linear regression analysis are reported in Supplemental Material 17
GFAP isoforms in Astrocystoma malignancy

Regression

II

III

IV

All

Grade

II

III

IV

DUSP4

VAV3

NOS2

SPC24
astrocytoma cell lines (Moeton et al., 2016). These findings and the observations in our study indicate that a change in the GFAP-network could influence the invasive capacity of astrocytoma cells by regulation of VAV3.

NOS2 encodes for the inducible nitric oxide synthase 2 that metabolizes L-arginine into the short-lived free radical nitric oxide (NO). NOS2 is higher expressed in brain tumours compared to healthy tissue and in highly malignant brain tumours compared to tumours of lower malignancy, in line with our findings (Ellie et al., 1995; Hara and Okayasu, 2004; Hara et al., 1996; Lin et al., 2008). We confirmed that NOS2 expression in grade IV astrocytoma was higher than in both grade II and III. In addition, we associated high expression of NOS2 in grade III astrocytoma patients to a lower survival probability. Previous analysis using the REMBRANDT dataset has linked high NOS2 expression to a worse survival probability in grade IV astrocytoma patients supporting the tumorigenic role of NOS2 (Eyler et al., 2011). This role is furthermore supported by studies that link NOS2 expression to GICs (Eyler et al., 2011), maintenance of tumour growth (Eyler et al., 2011), invasive capacity of tumour cells (Lin et al., 2008), radiation resistance (Kim et al., 2013), and angiogenesis (Kostourou et al., 2011). More evidence exists for a role of NOS2 in the acquisition of chemoresistance (Yang et al., 2004; Yin et al., 2001) and in the attenuation of the anti-tumour inflammatory response in the brain, thereby preventing tumour cell death (Badn et al., 2007; Hara et al., 1996). Hence NOS2 has a tumorigenic function at different levels of astrocytoma biology. Interestingly, induction of NOS2 expression by another identified GFAP-regulated tumorigenic gene, dual specificity phosphatase 4 (DUSP4) has been previously described (Al-Mutairi et al., 2010).

DUSP4 is part of a larger family of DUSPs that has been implicated in different types of tumours. DUSPs are known to modulate the MAPK pathway by their dephosphorylating activity (Al-Mutairi et al., 2010; Prabhakar et al., 2014) and can have both tumour suppressive and tumour promoting functions depending on the context (Prabhakar et al., 2014). DUSP4 is either up- or down-regulated in different types of tumours (Prabhakar et al., 2014). In our analysis, we found an increased DUSP4 expression in high-grade astrocytoma and a lower survival and progression free survival probability for grade III astrocytoma patients with high DUSP4 expression. This is in line with the reported significantly better clinical outcome of patients with increased methylation of the DUSP4 gene promoter (Waha et al., 2010) and the worse prognoses of oligodendrogliomas with an increased expression of DUSP4 due to a CIC mutation (Gleize et al., 2015). It is however possible that the effect of DUSP4 in astrocytomas could vary between tumours, as forced DUSP4
expression in vitro has been reported to inhibit the growth of an astrocytoma cell line (Waha et al., 2010).

The function of kinetochore-associated Ndc80 complex subunit SPC24 (SPC24) has not been studied in astrocytoma before, but findings in other types of tumours and healthy tissue identify SPC24 as a tumorigenic gene (Kaneko et al., 2009; Zhu et al., 2015) that has a negative prognostic effect in liver tumours (Zhu et al., 2015). We report an increased expression in high-grade astrocytoma compared to low-grade and the presence of SPC24 in our overrepresented GO cluster of regulation of proliferation. SPC24 is a component of the kinetochore complex and is essential for proper chromosome segregation and mitotic checkpoint control (Janke et al., 2001). In vitro, targeting SPC24 with siRNA inhibits cell proliferation and invasion and induces apoptosis in liver tumour cells (Zhu et al., 2015). These studies indeed indicate a tumorigenic role for SPC24, which supports our findings in astrocytoma.

GFAP-regulated anti-tumorigenic genes in tumour biology

NTRK2, FBXO2, ST3GAL6 and AKR1C3 are the GFAP-regulated anti-tumorigenic genes that we identified as most relevant for astrocytoma malignancy (Fig. 6e-h). The expression of these genes is inhibited by an increase in the GFAPδ/α ratio in vitro and in patients with high-grade astrocytoma.

Neurotrophic Tyrosine Kinase, Receptor, Type 2 (NTRK2) encodes for the receptor tropomyosin receptor kinase B (TrkB). A lower expression of NTRK2 in high-grade astrocytoma compared to low-grade has been reported previously (Palani et al., 2014) but inconsistently (Assimakopoulou et al., 2007; Xiong et al., 2015). In tumours, TrkB receptor activation is described to stimulate cell proliferation, cell invasion, and angiogenesis, and to contribute to tumour chemoresistance and metastasis (Desmet and Peeper, 2006; Thiele et al., 2009). Thus, TrkB activation is mainly associated with stimulation of tumour aggressiveness and growth, which is in contrast with our identification of TrkB as an anti-tumorigenic gene. However, as reviewed by Desmet et al., studies on different tumours do suggest a possible anti-tumorigenic function (Desmet and Peeper, 2006). In addition, expression of the truncated TrkB isoform, TrkB-T1, is associated with a more anti-tumorigenic function due to inhibition of full-length TrkB effects (Garofalo et al., 2015; Ohira et al., 2006), an effect that could be mediated by the direct binding of the Rho-GTPase inhibitor GDII to TrkB-T1 (Ohira et al., 2005, 2006). Interestingly, as described above, one of our identified tumorigenic genes, VAV3, is also known to increase Rho-GTPase activity and can induce morphological changes and invasivity in astrocytoma cells (Liu et al., 2014a; Movilla and Bustelo, 1999; Salhia et al., 2008). These results indicate
that similar downstream targets are influenced upon a decrease or an increase in the GFAPδ/α ratio in opposite directions.

The aldo-keto reductase 1 family member C3 (AKR1C3) is thought to protect the cell against aldehyde and oxidative cytotoxicity. We found increased expression in low-grade astrocytoma, although others reported no different expression between tumours of different grade (Park et al., 2010). Increased expression of AKR1C3 has been reported before in low-grade glioma with a specific 1q19p co-deletion (Ducray et al., 2008) which corroborates our findings of a better prognosis for grade III astrocytoma patients with high AKR1C3 expression (Bredel et al., 2009). In contrast, a high AKR1C3 expression has been linked to worse prognosis for grade IV patients (Bredel et al., 2009). In vitro findings in astrocytoma cells and oesophageal carcinomas suggest a tumorigenic function (Ragel et al., 2007; Xiong et al., 2014). Future studies will have to determine how AKR1C3 expression changes the biology of astrocytoma cells and leads to a better outcome for grade III astrocytoma patients.

Finally, we identified the genes encoding the F-box protein 2 (FBXO2) and the ST3 beta-galactoside alpha-2,3-sialyltransferase (ST3GAL6), as GFAP-regulated anti-tumorigenic genes. Both showed an increased expression in low-grade astrocytoma compared to high-grade and high ST3GAL6 expression in astrocytoma grade III patients is associated with a lower survival and progression free survival probability. To our knowledge, these genes have not been studied in astrocytoma. However, their described function in healthy tissue and other tumour types combined with our observations imply a role for them in astrocytoma. ST3GAL6 is a sialyltransferase that modulates glycosylated proteins such as N-glycans by adding sialic acid to their oligosaccharide chain. In different tumours, glycosylation is altered and ST3GAL6 is a modulator of this process in breast cancer cells. Increased ST3GAL6 expression leads to changes in interactions with the tumour cell environment and in cell adhesion and migration (Glavey et al., 2014). Furthermore, ST3GAL6 expression is a negative prognostic factor in a specific subtype of breast carcinoma (Milde-Langosch et al., 2014). To the contrary, decreased ST3GAL6 expression is observed in hepatocellular carcinomas compared to non-tumour liver tissue (Souady et al., 2011). Our data suggest an anti-tumorigenic role of ST3GAL6, but how ST3GAL6 modulation of glycan proteins could lead to a less malignant astrocytoma subtype is unknown. More support has been published for an anti-tumorigenic role of the f-box protein 2 (FBXO2). This gene encodes an F-box protein that is part of a ubiquitin ligase complex that specifically targets N-glycosylated proteins for proteasomal degradation (Yoshida et al., 2002). FBXO2 is thought to be responsible for the elimination of the N-glycosylated integrin β1 precursor protein, important
for adhesion, migration, and invasion (Yoshida et al., 2002). In addition to this, FBXO2 regulates protein degradation of cell cycle regulators. Increased expression of FBXO2 in different cell types induces growth arrest (Erhardt et al., 1998; Wen et al., 2010), related to the presence of FBXO2 in the overrepresented GO cluster of regulation of proliferation. These observations support an anti-tumorigenic role for FBXO2.

*GFAP-isoform stoichiometry and astrocytoma biology*

Our data show that the stoichiometry of GFAP-isoform expression is important for determining astrocytoma malignancy related molecular and functional processes. The gene ontology analysis on the GFAP-regulated (anti-)tumorigenic gene sets suggests a role in the regulation of tumour growth. Previous studies have suggested an important role of the IF network in astrocytoma cell migration and interaction with the ECM (Moeton et al., 2014, 2016). Moreover, a GO analysis on all GFAP-isoform induced transcriptional changes observed in our astrocytoma cell lines revealed the cell’s interaction with its environment through adhesion, ECM composition, and ECM remodelling as the major targeted modules regulated by GFAP modulation (Sup. Material 8-11). For brain tumours to grow and metastasize, ECM remodelling is an essential step. On the one hand tumours have to degrade the ECM barrier to allow for infiltration. On the other hand the ECM environment of the tumour itself has to meet requirements for maintenance of the various heterogeneous populations of tumour cells present in the tumour, ranging from a tumour stem cell population to the vasculature of the tumour (Wang et al., 2010). Remarkably, the 80 differentially expressed genes between GFAPδ+ and GFAPα+ are also overrepresented in the same clusters representing interaction with environment (Sup. Material 12) suggesting a role of the isoforms in these processes. In addition, analysis of annotations and function in physiology and pathology in an Ingenuity Pathway Analysis with a focus on cancer, more specifically indicated a role of GFAP-isoform regulated genes in tumour invasion (*data not shown*). Together, our data and previous observations suggest that GFAP-isoforms in astrocytoma could regulate the malignant phenotype of the cells both by regulating their growth as well as their invasive capacity by modulating the cell’s interaction with its environment and the ECM.

In conclusion, our data supports an important role for the IF-network in astrocytoma malignancy. Astrocytoma of low and high grade consist of a different network composition induced by altered levels of GFAP-isoforms, leading to changed expression of genes that are involved in tumour growth and invasion and
are associated with survival and progression of disease in patients. We identified VAV3, NOS2, DUSP4, and SPC24 as tumorigenic, and NTRK2, FBXO2, ST3GAL6, and AKR1C3 as anti-tumorigenic genes that are all regulated by the GFAP network. Thus, GFAP isoforms have an important function in astrocytoma malignancy and the GFAP-network and GFAP alternative splicing are novel targets that await further exploration as a potential novel treatment for astrocytoma.

Acknowledgements

We would like to acknowledge Andi van der Putten for qPCR and primer verification. The results published here are in part based upon data generated by The Cancer Genome Atlas Research Network: http://cancergenome.nih.gov/. We thank the Lineberger Comprehensive Cancer Center at the University of North Carolina at Chapel Hill for collecting and processing the data. This work was supported by the research programme of the Foundation for Fundamental Research on Matter (FOM), grant 09MMC06, and the Netherlands Organization for Scientific Research [NWO; VICI grant 865.09.003].

Legend Supplemental Materials

Supplemental materials can be found on www.ellyhollab.eu or will be made available upon request.

Supplemental Material 1

Primer pairs used and their qPCR efficiencies for microarray validation.

Supplemental Material 2

Modulation of GFAP-network in U373 and U251 astrocytoma cells

To get insight into the function of these GFAP-isoforms in astrocytoma and to investigate a potential role for GFAPα and the GFAPδ/α ratio in astrocytoma malignancy, we modulated GFAP-isoform expression in U251 (a-e) and U373 (f-l) cells by recombinant expression or silencing with shRNAs of the isoforms and analysed the GFAP-isoform induced transcriptomic changes. U251 cells were transduced with GFAPα (GFAPα+) or GFAPδ (GFAPδ+). Staining the control U251 cells showed some cells positive for endogenous GFAP, but predominantly GFAP-negative cells (a) GFAPα+ cells showed a GFAP-network throughout the cytoplasm (b), whereas GFAPδ+ cells showed the characteristic juxtanuclear accumulation of GFAP (c), as we have described before (Moeton et al., 2016). At the RNA level, using primers that recognize all (pan) GFAP-isoforms, a significant increase of pan-GFAP expression in GFAPα+ (FC = 9.6, FDR<.001) and GFAPδ+ (FC = 16.3, FDR<.001) cells was observed (d). GFAPδ mRNA was only increased in GFAPδ+ cells (FC = 1534, FDR<.001) (e). Depicted is the median (horizontal line) with quartiles (box), N=8. These finding are in agreement with western blot results showing that the amount of GFAP protein is increased in these cell lines.
as shown before (Moeton et al., 2016). U373 cells were transduced with shRNA directed against pan-GFAP (GFAPpan-) or against GFAPα (GFAPα-). Both the GFAPpan- (FC = 0.46, FDR=.063) and the GFAPα- (FC = 0.15, FDR=.027) U373 cells showed a decrease in GFAPα mRNA (f). GFAPα- cells shifted their ratio of GFAPδ/α mRNA more towards GFAPδ, due to a strong decrease in GFAPα (ratio of GFAPδ/α (arbitrary units) the ratio for NTC = 0.42; GFAPpan- = 0.51, FDR=.31; GFAPα- = 1.73, FDR=.077) (g), while the GFAPδ was stable (h) N=5, depicted is the median (horizontal line) with quartiles (box)). The knockdown of GFAP was confirmed by western blot, and was more efficient in the GFAPα- than in the GFAPpan- cells (i) depicted is median with quartiles (box), N=4. Stainings confirmed the less efficient knockdown of GFAP in the GFAPpan- (k) than the GFAPα- (l), compared to NTC cells (j). Scale bars in the micrographs represent 50 um.

**Supplemental Material 3**
Top 50 of most significant probes in GFAPα+ compared to Control cells.
Most significant genes annotated to probes detecting more than 1.5 fold up or down change induced by GFAPα+ expression. t = t-test statistics, fold change: >1 = GFAPα+ induced expression, <1 = GFAPα+ repressed expression, FDR adj.P.Val = False Discovery Rate controlled p-value.

**Supplemental Material 4**
Top 50 of most significant probes in GFAPδ+ compared to Control cells.
Most significant genes annotated to probes detecting more than 1.5 fold up or down change induced by GFAPδ+. t = t-test statistics, fold change: >1 = GFAPδ+ induced expression, <1 = GFAPδ+ repressed expression, FDR adj.P.Val = False Discovery Rate controlled p-value.

**Supplemental Material 5**
Top 50 of most significant probes in GFAPpan- compared to NTC cells.
Most significant genes annotated to probes detecting more than 1.5 fold up or down change induced by GFAPpan-. t = t-test statistics, fold change: >1 = GFAPpan- induced expression, <1 = GFAPpan-repressed expression, FDR adj.P.Val = False Discovery Rate controlled p-value.

**Supplemental Material 6**
Top 50 of most significant probes in GFAPα- compared to NTC cells.
Most significant genes annotated to probes detecting more than 1.5 fold up or down change induced by GFAPα-. t = t-test statistics, fold change: >1 = GFAPα- induced expression, <1 = GFAPα- repressed expression, FDR adj.P.Val = False Discovery Rate controlled p-value.

**Supplemental Material 7**
Heat map of genes with an FDR p-value <.1 in at least one knockdown and one recombinant expression condition.
Four main patterns were distinguished of either upregulation in both GFAPα- and GFAPδ+ or
downregulation in both GFAPα- and GFAPδ+, or an opposite regulation by GFAPα- and GFAPδ+, as indicated by different colors in the dendrogram (purple, grey, yellow, black). Some genes showed similar behaviour in GFAPα+ as well as in GFAPδ+ cells, whereas GFAPpan- in general didn’t show large effects. The observed strongest changes in the GFAPα- and GFAPδ+ condition imply that the change is due to the increase in GFAPδ/GFAPα ratio.

**Supplemental Material 8**
Top Gene Ontology terms per domain overrepresented by GFAPα+ compared to control.
Annotated genes = number of total microarray probes annotated to this term; Sig genes = number of probes detecting significant change in this term; p-val = parent-child weighted p-value of Fisher’s exact test for overrepresentation; GO = Gene Ontology; BP = Biological Process; CC = Cellular Component; MF = Molecular Function.

**Supplemental Material 9**
Top Gene Ontology terms per domain overrepresented by GFAPδ+ compared to control.
Annotated genes = number of total microarray probes annotated to this term; Sig genes = number of probes detecting significant change in this term; p-val = parent-child weighted p-value of Fisher’s exact test for overrepresentation; GO = Gene Ontology; BP = Biological Process; CC = Cellular Component; MF = Molecular Function.

**Supplemental Material 10**
Top Gene Ontology terms per domain overrepresented by GFAPpan- compared to NTC.
Annotated genes = number of total microarray probes annotated to this term; Sig genes = number of probes detecting significant change in this term; p-val = parent-child weighted p-value of Fisher’s exact test for overrepresentation; GO = Gene Ontology; BP = Biological Process; CC = Cellular Component; MF = Molecular Function.

**Supplemental Material 11**
Top Gene Ontology terms per domain overrepresented by GFAPα- compared to NTC.
Annotated genes = number of total microarray probes annotated to this term; Sig genes = number of probes detecting significant change in this term; p-val = parent-child weighted p-value of Fisher’s exact test for overrepresentation; GO = Gene Ontology; BP = Biological Process; CC = Cellular Component; MF = Molecular Function.

**Supplemental Material 12**
Top Gene Ontology terms per domain overrepresented by comparing GFAPα+ to GFAPδ+.
Annotated genes = number of total microarray probes annotated to this term; Sig genes = number of probes detecting significant change in this term; p-val = parent-child weighted p-value of Fisher’s exact test for overrepresentation; GO = Gene Ontology; BP = Biological Process; CC = Cellular
GFAP isoforms in Astorcytoma malignancy

Component; MF = Molecular Function.

Supplemental Material 13
Heatmap of the genes that significantly correlated to either GFAP\(\alpha\) or the GFAP\(\delta/\alpha\) ratio (FDR < 0.01).
The color key indicates the correlation coefficient (R). Hierarchical clustering on the absolute correlation coefficient results into the identification of two main clusters. Cluster 1 consists of genes that positively correlate to GFAP\(\alpha\) but negative to the GFAP\(\delta/\alpha\) ratio, cluster 2 shows the opposite pattern with genes negatively correlating to GFAP\(\alpha\) and positively to the GFAP\(\delta/\alpha\) ratio.

Supplemental Material 14
Linear regression results of the 43 identified anti-tumorigenic genes.
Linear regression results of genes that show a significant positive correlation to GFAP\(\alpha\) (\(\alpha\) columns), a negative correlation to GFAP\(\delta/\alpha\) ratio (\(\delta/\alpha\) columns) and do not correlate to GFAP\(\delta\) (\(\delta\) columns).
FDR = False Discovery Rate controlled p-value.

Supplemental Material 15
Linear regression results of the 37 identified tumorigenic genes.
Linear regression results of genes that show a significant negative correlation to GFAP\(\alpha\) (\(\alpha\) columns), a positive correlation to GFAP\(\delta/\alpha\) ratio (\(\delta/\alpha\) columns) and do not correlate to GFAP\(\delta\) (\(\delta\) columns).
FDR = False Discovery Rate controlled p-value

Supplemental Material 16
Kaplan meier survival analysis of GFAP-regulated (anti-)tumorigenic genes in grade III astrocytoma patients.
Results of log rank regression analysis of survival estimates of below and above median expression of genes of interest.
FDR = False Discovery Rate controlled p-value, AT = anti-tumorigenic, TU = tumorigenic, Direction of expression and survival probability = indication whether low or high expression of the gene of interest results in a lower survival probability for grade III astrocytoma patients.

Supplemental Material 17
Linear regression results of final set of GFAP-regulated (anti-)tumorigenic most likely to be regulated by GFAP in patient tumours and are relevant for astrocytoma malignancy.
Linear regression results of genes that show a significant correlation to GFAP\(\alpha\) (\(\alpha\) columns) or to the GFAP\(\delta/\alpha\) ratio (\(\delta/\alpha\) columns) within astrocytoma grade II, III or IV.
FDR = False Discovery Rate controlled p-value